

Universidad Autónoma Metropolitana  
Unidad Xochimilco  
División de Ciencias Sociales y Humanidades

La evaluación de la investigación y su relación con la producción  
de conocimiento: un análisis estructural para el Sistema  
Nacional de Investigadores de México

Tesis para obtener el grado de doctora en Ciencias Sociales, área de  
concentración Economía y Gestión de la Innovación

Presenta

Natalia Gras Tuñas

Asesora: Dra. Gabriela Dutrénit Bielous

Lectores:

Dra. Claudia González-Brambila

Dr. José Miguel Natera

Dr. Martín Puchet

Dr. Ismael Rafols

Dra. Giovanna Valenti

Dr. Delfino Vargas

Febrero 2018

Ciudad de México

## Agradecimientos

Este trabajo es el resultado de un largo recorrido que he transitado, afortunadamente, muy acompañada, a todos gracias.

Para mis hermanos, mi viejo y mi vieja -que llevo en el corazón- un enorme reconocimiento y agradecimiento porque con su amor, incondicionalidad y contención, me han estimulado y ayudado desde siempre. Para Juan otro, no solamente por el amor, el camino juntos, con las idas, las vueltas y el retorno, sino también por su mirada aguda, crítica y sugerencias metodológicas que enriquecieron y ampliaron las posibilidades de estudio en este trabajo.

Para los amigos un *gracias totales* por todo, por ser como son y por estar y compartir conmigo siempre.

Para Gaby, mil gracias es poco! por su calidad humana, por su ayuda permanente y contención incansable desde que nos conocimos en 2010, cuando me facilitó todo cuanto pudo para que yo pudiera llegar y vivir en México. Además, me enseñó y me dió las herramientas para investigar. Sus lecturas detalladas y comentadas, exigencias, sugerencias y críticas han enriquecido enormemente este estudio.

Para Judith, gracias por instalarme la inquietud, el gusto y pasión por el tema que se aborda en esta tesis. Los diálogos con ella, los trabajos compartidos y sus preguntas me brindaron valiosos espacios de reflexión y aprendizaje. Cómo no mencionar en este agradecimiento a mis compañeros de trabajo por el apoyo, el aguante durante todos estos años y la amistad. Para Mariela, Melissa y Lucía además por la *terapia* y el cariño.

También gracias a la Universidad Autónoma Metropolitana, a su Doctorado en Ciencias Sociales, a sus profesores y estudiantes. Ellos, junto a mis compañeros de generación, enriquecieron mi formación, me brindaron la oportunidad de ampliar mis conocimientos y de profundizar mi comprensión sobre diversos temas.

Un agradecimiento para el Consejo Nacional de Ciencia y Tecnología de México, por otorgarme la beca de apoyo económico para la realización de estudios de doctorado, para la Comisión Sectorial de Investigación Científica de la Universidad de la República de Uruguay, por fomentar y apoyar mi formación en el exterior, y para el Foro Consultivo Científico y Tecnológico por permitirme el acceso a la encuesta que se utiliza en este estudio.

Muchas gracias también a los miembros del tribunal de esta tesis, por la lectura y revisión detallada de este documento, la crítica constructiva y los valiosos comentarios recibidos. Un agradecimiento especial a la Dra. Valenti por la lectura comentada de los múltiples avances entregados y al Dr. Vargas por sus ideas y revisión de la implementación del modelo de ecuaciones estructurales realizado.

## Resumen

En esta tesis se explora la naturaleza de las relaciones entre un sistema de evaluación fuerte (Whitley, 2007) del desempeño individual de investigación y dos modos diferenciales de producción de conocimiento –Modo 1 y Modo 2 (Gibbons et al., 1994)- en el contexto del Sistema Nacional de Investigadores (SNI) de México. Estas relaciones no han sido exploradas con anterioridad y si bien diversos estudios han sugerido la homogeneización y estandarización de los tipos de producción de conocimiento, ninguno de ellos ha reportado evidencia específica sobre las relaciones entre la evaluación y esos dos modos diferenciales de producción de conocimiento científico.

El análisis de las relaciones que se estudian en esta tesis, se justifica y fundamenta en la integración que se hace de las formas diferenciales de producción de conocimiento de Gibbons et al (1994) con la teoría que vincula a la evaluación del desempeño de la investigación y la producción de conocimiento (Whitley, 2007 y 2010; Gläser y Laudel, 2007b; Gläser, 2012), así como en los debates y en hallazgos de investigaciones pasadas respecto a las consecuencias probables de la evaluación sobre la dirección, contenido y características epistemológicas de la producción de conocimiento.

Con base en esas contribuciones, en esta tesis se construyen cinco dimensiones conceptuales y se plantean seis hipótesis sobre la naturaleza de las relaciones entre ellas, que se ponen a prueba conjuntamente a través del desarrollo y validación de un modelo teórico estructural (Kline, 2011; Hair et al., 2014; Brown, 2015). Los datos para validar la teoría que se propone y que se representa formalmente en el modelo teórico estructural, provienen de la encuesta titulada “*Percepciones sobre la evaluación académica*” (FCCyT, 2014).

Los resultados obtenidos proveen evidencia específica de que el SNI no es neutro, por el contrario, introduce sesgos en su evaluación a favor de la producción de conocimiento modo 1 y en contra de la producción de conocimiento modo 2, sugiriendo efectos constitutivos de la evaluación y con ello, reduciendo el impacto potencial de la producción de conocimiento en términos de su contribución en la búsqueda de soluciones a los principales problemas que afectan al desarrollo social y económico de México. Desde el punto de vista de las implicancias de política de los resultados de investigación obtenidos, la evaluación que se implementa en el SNI, más que ayudar al cumplimiento integral de sus objetivos de política, los contradice.

Esta tesis aporta un análisis original, válido y confiable sobre la relación entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento y empíricamente contribuye con evidencia nueva sobre esa relación en el contexto de un sistema de evaluación fuerte como el SNI de un país en desarrollo como lo es México. Conceptualmente, esta tesis introduce por un lado, los modos 1 y 2 como dimensiones relevantes que están siendo afectadas por los sistemas de evaluación fuertes y en ese sentido, aporta evidencia sistemática sobre efectos epistémicos del sistema de evaluación sobre la dirección y contenido de la investigación científica en el país. Por otro lado, la tesis introduce el concepto de *flexibilidad epistemológica de los sistemas de evaluación* del desempeño de la investigación. La relevancia y pertinencia de ese concepto radica en las posibilidades que ofrece como instrumento de evaluación y monitoreo de los efectos epistémicos del sistema de evaluación sobre la producción de conocimiento científico y tecnológico, proveyendo información relevante para su eventual re-diseño. Metodológicamente, en esta investigación se realizaron esfuerzos importantes para garantizar la validez de constructo de todas las medidas que se proveen. La relevancia de estos

esfuerzos está en términos de la reflexión metodológica y técnica que aporta esta investigación al campo de los estudios sobre evaluación del desempeño de la investigación y al ámbito de la política pública de estímulo a la investigación y sus instrumentos de evaluación.

Se espera que el análisis y los resultados de esta investigación contribuyan al debate público sobre el SNI y la importancia de estimular también una agenda de producción de conocimiento propia, contextualizada, de calidad, diversa y plural en México, para que la producción de conocimiento que se realiza en el país pueda contribuir también al alcance de sus objetivos de desarrollo nacional.

## Contenido

CAPÍTULO I: Introducción.....	12
CAPÍTULO II: Los cambios en la organización y orientación del conocimiento: las formas diferenciales de producción y su relación con la evaluación.....	23
II.1 De la <i>ciencia académica</i> a la <i>ciencia posacadémica</i> : características cambiantes de la producción de conocimiento.....	24
II.1.1 La producción colectiva y no disciplinar de conocimiento: un cambio en su organización.....	24
II.1.2 La producción de conocimiento explícitamente orientada por problemas económicos y sociales: la pertinencia social y económica del conocimiento.....	26
II.1.3 La ciencia académica y la ciencia industrial: dos extremos en rol potencial del conocimiento.....	27
II.2 Las características diferenciales del <i>Modo 1</i> y <i>Modo 2</i> de producción de conocimiento.....	31
II.2.1 Producción de conocimiento en el Contexto de Aplicación.....	33
II.2.2 Producción de conocimiento Transdisciplinaria.....	34
II.2.3 Producción de conocimiento Heterogénea.....	35
II.2.4 Producción de conocimiento Socialmente Responsable y Reflexivo.....	36
II.2.5 Producción de conocimiento y el control de su calidad.....	36
II.3 La evaluación de la investigación en el <i>contexto de aplicación</i> : algunos enfoques.....	37
II.4 Conclusión: la racionalidad detrás de la selección de los <i>Modo 1</i> y <i>Modo 2</i> como dos <i>tipos ideales</i> de producción de conocimiento.....	41
CAPITULO III: La teoría que vincula la evaluación con la producción de conocimiento.....	44
III.1 Las consecuencias probables sobre la organización de la investigación y dirección de la producción de conocimiento.....	45
III.2 Adaptación a las presiones del entorno: las respuestas estratégicas de los investigadores y el contenido de la producción de conocimiento.....	47
III.3 Los impactos esperados de los sistemas de evaluación fuertes dependen de las características específicas del contexto en que se implementen.....	49
III.4 La gobernanza de la ciencia: los balances de poder y su relación con la producción de conocimiento.....	52
III.5 Conclusión: las relaciones entre la evaluación y los modos diferenciales de producción de conocimiento: una justificación y el fundamento teórico de las relaciones estudiadas en esta tesis.....	53
CAPITULO IV: Los <i>indicadores bibliométricos</i> y sus <i>efectos constitutivos</i> sobre el contenido, dirección y características epistemológicas de la producción de conocimiento.....	57
IV.1 La difusión de los indicadores bibliométricos en la gobernanza de la ciencia y el cuestionamiento internacional.....	57
IV.2 El problema de la validez de los indicadores y sus consecuencias constitutivas.....	69
IV.3 Las relaciones entre la evaluación del desempeño de la investigación y la producción de conocimiento: los hallazgos de investigaciones pasadas.....	71
IV.4 Conclusión: la relación entre la evaluación del desempeño de la investigación basada en indicadores y la dirección, contenido y características epistemológicas de la producción de conocimiento.....	86

CAPITULO V: El Marco Teórico-Conceptual: las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento.....	89
V.1 Las definiciones conceptuales de las dimensiones relevantes: los constructos.....	89
V.1.1 Propensión modo 1.....	90
V.1.2 Propensión modo 2.....	91
V.1.3 Flexibilidad epistemológica del sistema de evaluación .....	92
V.1.4 Satisfacción con el proceso de evaluación.....	95
V.1.5 Dirección de la reforma del sistema de evaluación.....	97
V.2 La teoría estructural sobre las relaciones entre la evaluación y los modos diferenciales de producción de conocimiento: hipótesis sobre las relaciones entre los constructos .....	102
CAPITULO VI: Los sistemas de ciencia, tecnología e innovación y de evaluación a la investigación individual en México .....	112
VI.1 El sistema de Ciencia, Tecnología e Innovación en México: algunas características importantes.....	113
VI.1.1 La producción de conocimiento y su financiamiento .....	113
VI.1.2 El mercado de trabajo de los investigadores .....	117
VI.2 El sistema de evaluación: creación y evolución del SNI desde su marco regulatorio.....	119
VI.2.1 Acuerdo de creación del SNI y sus modificaciones .....	119
VI.2.2 Reglamentos de operación y funcionamiento del SNI.....	123
VI.2.3 Criterios de evaluación específicos e internos de las Comisiones Dictaminadoras por área de conocimiento .....	131
VI.3 Conclusión.....	138
CAPITULO VII: Antecedentes: el debate sobre el SNI en el ámbito nacional.....	141
VII.1 El SNI y sus logros: algunas consecuencias performativas de la implementación de un sistema de evaluación fuerte .....	142
VII.1.1 Institucionalización de la evaluación, retribución económica según desempeño y prestigio.....	142
VII.1.2 Introducción de estándares internacionales de evaluación .....	144
VII.1.3 Profesionalización de la investigación: una actividad altamente normada.....	146
VII.1.4 Fortalecimiento del sistema científico nacional .....	147
VII.1.5 Retención de investigadores en el ámbito académico nacional.....	149
VII.1.6 Mejoramiento del proceso de evaluación .....	150
VII.2 El SNI y sus desafíos: otras consecuencias constitutivas de la implementación de un sistema de evaluación fuerte .....	151
VII.2.1 El SNI y su flexibilidad epistemológica: la diversidad institucional, de actividades y de la orientación de la producción de conocimiento cuestionada.....	151
VII.2.2 El SNI y su flexibilidad epistemológica: la calidad de la producción de conocimiento cuestionada .....	154
VII.2.3 El SNI y la escasa contribución del conocimiento científico y tecnológico al desarrollo social y económico de México.....	156
VII.3 Dos dimensiones que en interacción sirven para entender las consecuencias de la implementación de un sistema de evaluación fuerte como el SNI.....	157
VII.4 El SNI y la dirección del cambio: algunas propuestas para atender sus desafíos.....	163
VII.5 Conclusión.....	166

CAPITULO VIII: Diseño y estrategia de investigación: métodos, técnicas y fuente de información	169
VIII.1 Estrategia de Investigación .....	172
VIII.2 Etapa 1: Exploración y descripción de las relaciones entre las variables observadas, el análisis factorial exploratorio.....	174
VIII.2.1 Diseño del AFE.....	174
VIII.2.2 Sentido de implementar el AFE.....	175
VIII.2.3 Implementación del AFE: obtención de los factores y evaluación del ajuste global ..	176
VIII.2.4 Interpretación conceptual de los factores e identificación de la estructura factorial más apropiada para los objetivos de este trabajo .....	177
VIII.3 Etapa 2: Medición y validación, el análisis factorial confirmatorio .....	179
VIII.3.1 La teoría de medida, sus hipótesis y la consecuente definición de los constructos o variables latentes .....	180
VIII.3.2 Desde la teoría hacia el modelo: consideraciones para la especificación del modelo de medida.....	181
VIII.3.3 Diseño del AFC .....	185
VIII.3.4 Evaluación de la Validez del Modelo de Medida .....	191
VIII.4 Etapa 3: Explicación de la relación entre la evaluación de la investigación y la producción de conocimiento, el modelo de ecuaciones estructurales .....	195
VIII.4.1 Procedimientos y consideraciones para la especificación del modelo teórico estructural .....	195
VIII.4.2 Evaluación de la Validez del Modelo Teórico Estructural .....	197
VIII.4.3 Modelo teórico estructural condicional.....	200
CAPÍTULO IX: Análisis Descriptivo .....	201
IX.1 Caracterización de los investigadores categorizados en el SNI .....	202
IX.2 Ámbitos y motivaciones de la producción de conocimiento: las propensiones Modo 1 y Modo 2 en el SNI .....	207
IX.2.1 Propensión modo 1 de producción de conocimiento .....	208
IX.2.2 Propensión modo 2 de producción de conocimiento.....	212
IX.2.3 Una síntesis de los resultados .....	216
IX.3 Satisfacción con el proceso de evaluación del SNI.....	220
IX.4 Flexibilidad epistemológica del sistema de evaluación del SNI .....	228
IX.5 Dirección de la reforma del sistema de evaluación .....	236
IX.6 Conclusión .....	245
CAPÍTULO X: Los Modelos de Medición y Estructural: Especificación, Confiabilidad y Validación	250
X.1 Estrategia de Análisis .....	250
X.2 Análisis Factorial Exploratorio: ítems, estructura de correlación y dimensionalidad.....	251
X.2.1 Diseño.....	251
X.2.2 Estimación: selección del número de factores a retener, ajuste e interpretación.....	257
X.3 Análisis Factorial Confirmatorio: El Modelo de Medida .....	266
X.3.1 Hipótesis del modelo de medida y definición de los constructos: las variables latentes .....	267
X.3.2 Modelo de Medida: representación gráfica y formalización matemática.....	273
X.3.3 Validez del Modelo de Medida .....	277

X.4 El Modelo de Ecuaciones Estructurales: evaluación de la investigación y producción de conocimiento .....	287
X.4.1 Modelo Teórico Estructural: representación gráfica y formalización matemática.....	287
X.4.2 Validez del Modelo Teórico Estructural: La relación entre la evaluación y la producción de conocimiento.....	291
X.5 El modelo teórico estructural condicional .....	301
X.5.1 Especificación matemática.....	301
X.5.2 Evaluación de la Bondad de Ajuste del Modelo Teórico Estructural Condicional .....	304
X.5.3 Modelo teórico estructural condicional <i>versus</i> el modelo sin condicionar .....	305
X.6 Conclusión: principales resultados.....	311
CAPÍTULO XI: Conclusiones .....	314
Bibliografía .....	324
ANEXO 1: Análisis Factorial Exploratorio. Salida M-PLUS .....	340
ANEXO 2: Análisis Factorial Confirmatorio. Salida M-PLUS .....	392
ANEXO 3: Prueba de la diferencia de la Ji-cuadrado entre el modelo multifactorial de cinco factores <i>versus</i> un modelo unidimensional. Salida M-PLUS .....	405
ANEXO 4: Modelo de Ecuaciones Estructurales. Salida M-PLUS.....	417
ANEXO 5: Modelo de Ecuaciones Estructurales Saturado. Salida M-PLUS.....	430
ANEXO 6: Modelo de Ecuaciones Estructurales Condicional. Salida M-PLUS.....	443
ANEXO 7: Modelos LOGIT para explicar la probabilidad de datos faltantes como una función de las covariables X .....	457

#### Indice de Tablas

Tabla 1: Atributos diferenciales del Modo 1 y Modo 2: dos tipos ideales de producción de conocimiento .....	43
Tabla 2:: Distribución del Gasto Federal en Ciencia y Tecnología (GFCyT) según participación sectorial (2005-2014).....	117
Tabla 3: Investigadores en México según ámbito de empleo y otras medidas relativas.....	118
Tabla 4: Criterios Generales: actividades y productos centrales considerados en la evaluación del SNI .....	126
Tabla 5: Vigencia y Estímulos: tiempo entre evaluaciones y el incentivo económico .....	129
Tabla 6: Distribución poblacional y muestral según género, nivel de categorización y área de conocimiento .....	187
Tabla 7: Distribución de investigadores SNI por Género .....	202
Tabla 8: Distribución de investigadores SNI por Tramos de Edad .....	203
Tabla 9: Distribución de investigadores SNI por Nivel .....	203
Tabla 10: Distribución de investigadores SNI por ascenso en el SNI .....	204
Tabla 11: Distribución de investigadores SNI según si ha participado como evaluador.....	205
Tabla 12: Distribución de investigadores SNI según área de conocimiento .....	205
Tabla 13: Distribución de investigadores SNI según el tipo de institución en la que trabajan.....	206
Tabla 14: Distribución de investigadores según Efectos del SNI.....	206
Tabla 15: Distribución de investigadores SNI según su percepción sobre reformar el mecanismo de evaluación del SNI .....	207
Tabla 16: Distribución porcentual de investigadores SNI, según su grado de acuerdo sobre los ámbitos y motivaciones de la producción de conocimiento (Propensión Modo 1) .....	210



Tabla 17: Distribución porcentual de investigadores SNI, según su grado de acuerdo sobre los ámbitos y motivaciones de la producción de conocimiento (Propensión Modo 2) .....	213
Tabla 18: Distribución porcentual de investigadores SNI, según su percepción sobre los diversos ítems asociados a la Satisfacción con el Proceso de Evaluación.....	221
Tabla 19: Distribución porcentual de investigadores SNI, según su percepción sobre si el SNI facilita ciertas características epistemológicas de la producción de conocimiento (Flexibilidad Epistemológica del sistema de evaluación del SNI) .....	229
Tabla 20: Distribución porcentual de investigadores SNI, según su interés de reformar el sistema de evaluación según los ítems asociados a la dirección de la reforma (Dirección de la Reforma del sistema de evaluación del SNI).....	237
Tabla 21: Ítems, tamaño de muestra, datos faltantes y propiedades de medida .....	252
Tabla 22: Matriz de correlación (policórica) muestral .....	255
Tabla 23: Valores propios y porcentaje de la varianza explicada por cada factor .....	257
Tabla 24: Medidas de Ajuste Global del AFE.....	259
Tabla 25: Estructura Factorial: Solución del AFE para el modelo con 4 factores.....	260
Tabla 26: Estructura Factorial: Solución del AFE para el modelo con 5 factores.....	262
Tabla 27: Estructura Factorial: Solución del AFE para el modelo con 6 factores.....	265
Tabla 28: Variables Latentes o Constructos: definición conceptual, ítems y escala de medida.....	269
Tabla 29: Estadístico de prueba de bondad de ajuste Ji-Cuadrado .....	277
Tabla 30: Índices complementarios de Bondad de Ajuste .....	277
Tabla 31: Prueba de la diferencia de la Ji-Cuadrado para modelos anidados y medidas de Bondad de Ajuste del modelo de medida unifactorial.....	278
Tabla 32: Estimación del Modelo de Medida para la relación entre la evaluación de la investigación individual y la producción de conocimiento –solución sin estandarizar- .....	280
Tabla 33: Estimación del Modelo de Medida para la relación entre la evaluación de la investigación individual y la producción de conocimiento –solución completamente estandarizada- .....	281
Tabla 34: $R^2$ y Varianzas Residuales .....	282
Tabla 35: Varianza Extraída Promedio y Confiabilidad Compuesta .....	283
Tabla 36: Correlaciones entre las Variables Latentes .....	284
Tabla 37: Medidas de Bondad de Ajuste del Modelo Estructural comparadas con el Modelo de Medida .....	292
Tabla 38: Comparación entre los parámetros de medida del Modelo Estructural y las cargas factoriales del Modelo de Medida .....	295
Tabla 39: Parámetros estructurales del Modelo Teórico Estructural: Evaluación y producción de conocimiento .....	296
Tabla 40: Varianza Explicada ( $R^2$ ) de los constructos endógenos .....	299
Tabla 41: Variables independientes: Características de los investigadores categorizados en el SNI .....	302
Tabla 42: Medidas de Bondad de Ajuste del Modelo Estructural Condicional comparadas con las del Modelo Estructural sin condicionar .....	305
Tabla 43: Comparación entre los parámetros de medida del Modelo Estructural condicional versus el modelo sin condicionar .....	305
Tabla 44: Parámetros estructurales del Modelo Teórico Estructural Condicional versus el Modelo sin condicionar: Evaluación y producción de conocimiento .....	307
Tabla 45: Varianza Explicada ( $R^2$ ) de los constructos endógenos Modelo Teórico Estructural Condicional versus el Modelo sin condicionar.....	307
Tabla 46: Efecto total, directo e indirecto de propensión modo 1 sobre satisfacción con el proceso de evaluación (resultados estandarizados).....	308

Tabla 47: Síntesis de los resultados. Estimaciones estandarizadas del modelo teórico estructural condicional .....	318
--	-----

#### Índice de Gráficos

Gráfico 1: GIDE en relación PIB (Algunos países seleccionados, promedio de OECD y de América Latina y el Caribe).....	114
Gráfico 2:% del GIDE según fuente de financiamiento (año 2014).....	115
Gráfico 3: Participación de Hombres y Mujeres en los niveles del SNI.....	204
Gráfico 4: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Género) .....	216
Gráfico 5: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Tramos de Edad) .....	217
Gráfico 6: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Nivel) .....	218
Gráfico 7: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Área de conocimiento) .....	219
Gráfico 8: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Tipo de Institución) .....	220
Gráfico 9: Distribución del índice de Satisfacción con el Proceso de Evaluación del SNI (Total y por Nivel) .....	226
Gráfico 10: Distribución del índice de Satisfacción con el Proceso de Evaluación del SNI (Total y por Área de conocimiento).....	226
Gráfico 11: Distribución del índice de Satisfacción con el Proceso de Evaluación del SNI (Total y por Tipo de Institución) .....	228
Gráfico 12: Distribución del índice de Flexibilidad epistemológica del sistema de evaluación (Total y por Nivel).....	234
Gráfico 13: Distribución del índice de Flexibilidad epistemológica del sistema de evaluación (Total y por Área de conocimiento) .....	235
Gráfico 14: Distribución del índice de Flexibilidad Epistemológica del sistema de evaluación (Total y por Tipo de Institución).....	236
Gráfico 15: Distribución del índice de Dirección de la reforma del sistema de evaluación (Total y por Nivel) .....	243
Gráfico 16: Distribución del índice de Dirección de la reforma del sistema de evaluación (Total y por Área de conocimiento).....	244
Gráfico 17: Distribución del índice de Dirección de la reforma del sistema de evaluación (Total y por Tipo de Institución) .....	245
Gráfico 18: Gráfico de sedimentación (scree plot) .....	259

#### Índice de Figuras

Figura 1: La evaluación de la investigación y los modos diferenciales de producción de conocimiento .....	56
Figura 2: Marco Teórico Conceptual .....	111
Figura 3: Relación entre $y$ , $y^*$ y $\tau$ .....	183
Figura 4: Diagrama de sendero del Modelo de Medida .....	276
Figura 5:Diagrama de senderos para el Modelo Teórico Estructural: evaluación y producción de conocimiento .....	289

Figura 6: Modelo teórico estructural: evaluación de la investigación y producción de conocimiento ..... 300

## CAPÍTULO I: Introducción

Desde la segunda mitad del SXX son observables, en los países altamente industrializados, diversas transformaciones en la producción de conocimiento. Existe un cuerpo de literatura, en el que se destacan las contribuciones de Ziman (2000) y Gibbons et al. (1994), que discute los cambios en el rol del conocimiento, organización, formas de producción, legitimación y difusión.

El punto de partida o referencia de esa literatura para identificar dichas transformaciones, es la producción de conocimiento en el contexto de la ciencia académica (Polanyi, 1962). Mientras que la principal contribución de Ziman (2000) está en la elaboración de un enfoque sobre la transición desde el sistema de *ciencia académica* hacia un nuevo sistema de *ciencia posacadémica*; la de Gibbons et al. (1994) está en el desarrollo de un enfoque centrado en la caracterización de un nuevo modo de producción de conocimiento que denominan *Modo 2*, para diferenciarlo del modo de producción de conocimiento en el contexto de la ciencia académica –*Modo 1*-. Si bien ambos enfoques, con perspectivas y énfasis diferentes, reflexionan sobre aspectos vinculados con la incidencia de la evaluación de la investigación en la producción de conocimiento, el estudio específico de esa relación no es ni el eje central de sus análisis ni representa su principal contribución.

Desde los primeros años de la década de 1980 se difundieron e institucionalizaron diferentes sistemas de evaluación del desempeño de la investigación en diversos sistemas públicos de ciencia, tecnología e innovación (CTI) de distintos países desarrollados y en desarrollo. Surge otro cuerpo de literatura, en el que se destacan las contribuciones de Whitley (2007 y 2010), Gläser y Laudel (2007b) y Gläser (2012), que se ha focalizado en el estudio de la relación entre los sistemas de evaluación del desempeño de la investigación y la producción de conocimiento.

La implementación de *sistemas de evaluación fuertes* del desempeño de la investigación genera un conjunto de consecuencias probables sobre la organización de la investigación y la dirección de la producción de conocimiento (Whitley, 2007). Estos sistemas de evaluación profundizaron los cambios en la gobernanza de la ciencia, es decir, en las relaciones de autoridad y los (des)balances de poder entre quienes gobiernan la selección de los objetivos de la ciencia y la evaluación de sus resultados (Whitley, 2010). Así son esperables consecuencias notables sobre las características epistemológicas de la producción de conocimiento y su contenido (Gläser y Laudel, 2007b; Gläser, 2012).

Si bien es muy difícil establecer causalidad, es decir, que los cambios observados en la ciencia a escala macro son consecuencia directa y exclusiva de la implementación de ciertos sistemas de evaluación (Gläser y Laudel, 2016; Gläser, 2017), distintos estudios, internacionales y nacionales, realizados en diversos campos de las ciencias sociales –en sociología de la ciencia, gobernanza, evaluación, política científica, educación superior e innovación, entre otros- han aportado reflexión y evidencia que sugiere que los sistemas de evaluación del desempeño de la investigación inciden (se asocian) de manera importante en el contenido y características epistemológicas de la producción de conocimiento. A continuación se sintetizan los principales hallazgos -vinculados a la relación entre la evaluación y la producción de conocimiento- de algunos de estos estudios.

Las evaluaciones del desempeño de la investigación han promovido una cierta estandarización de los tipos y productos de la investigación (De Rijcke et al., 2016; Hicks, 2012; Bensusán et al., 2014).

Esas evaluaciones han jerarquizado la producción de artículos científicos publicados en revistas con alto factor de impacto y/o mejor posicionadas en los *rankings* e indexadas a bases de datos internacionales tales como *Web of Science* (WoS) o SCOPUS (Leisyte y Westerheijden, 2014; De Ibarrola, 2012). En contraste, han desvalorizado otros productos de la investigación (De Rijcke et al., 2016; De Ibarrola, 2012) tales como las monografías (Hammarfelt y De Rijcke, 2014), los libros (Leisyte y Westerheijden, 2014), los reportes para informar la toma de decisiones de política (Hicks, 2004), las publicaciones en revistas científicas de circulación nacional (van Dalen y Henkens, 2012; Hicks, 2004; Heras, 2005), los desarrollos tecnológicos (Padilla, 2010), el desarrollo de nuevas técnicas y recomendaciones para el sector productivo (Rivera et al., 2011), entre otros.

Así, dichas evaluaciones han incidido en que el contenido de la producción de conocimiento se oriente por el interés de las audiencias académicas internacionales (De Rijcke et al., 2016; Chavarro, Tang y Rafols, 2016; Vessuri, Guédon y Cetto, 2014; Hicks, 2012; Esteinou Madrid, 2012). Esto se ha expresado en el incremento de la cantidad de artículos publicados internacionalmente (Gläser y Laudel, 2007b; van Dalen y Henkens, 2012; Hammarfelt y De Rijcke, 2014; Leisyte y Westerheijden, 2014; Gonzalez-Brambila, 2005; Gonzalez-Brambila y Veloso, 2007) y con ello, en el aumento del reconocimiento y la incidencia internacional de tales publicaciones (Gläser y Laudel, 2007b; Leisyte y Westerheijden, 2014; Bensusán et al., 2014; Valenti et al., 2013; Esteinou, 2012) de todos los países en los que se han implementado dichas evaluaciones de desempeño.

Sin embargo, de acuerdo con los hallazgos de investigaciones pasadas, estas evaluaciones, simultáneamente y probablemente, ponen en riesgo y/o limitan la diversidad y ciertos tipos de producción de conocimiento, y con ello también alteran el contenido y la dirección de la producción de conocimiento. Algunos trabajos han sugerido que esos sistemas de evaluación generan dificultades para el desarrollo de investigaciones orientadas por la búsqueda de soluciones a problemas complejos tanto científicos (Rafols et al., 2012; Drucker Colin, 2005) como prácticos – sociales, productivos, energéticos, ambientales, etc.- (Rafols et al., 2012; Bensusán et al., 2014), conducen al abandono de líneas de investigación por su escaso desempeño en los parámetros de evaluación (Gläser y Laudel, 2007b), y estimulan una mayor orientación hacia los temas de investigación de “moda” (Gläser y Laudel, 2007b; Pescador Osuna, 2012).

Distintos estudios han problematizado y sugerido, como una tendencia más generalizada derivada de esos sistemas de evaluación, el deterioro y desincentivo de la realización de agendas de investigación contextualizadas (Bianco, Gras y Sutz, 2016), de gran importancia local y/o regional (Chavarro, Tang y Rafols, 2016; Hicks, 2012; van Dalen y Henkens, 2012), en especial en el contexto de América Latina y otras regiones periféricas (Vessuri, Guédon y Cetto, 2014) o, en otros términos, de los países en desarrollo (Bianco, Gras y Sutz, 2016). En el ámbito nacional, diversos autores han reflexionado sobre la evaluación del desempeño de la investigación –específicamente sobre el Sistema Nacional de Investigadores (SNI)- y su incidencia en el desplazamiento de la investigación orientada a atender los principales problemas que afectan el desarrollo social y económico de México hacia una investigación más básica (Esteinou, 2012; Reséndiz Núñez, 2005; Fortes Besprosvani, 2012; Pescador Osuna, 2012).

Distintos estudios han reflexionado sobre el desincentivo de la investigación colaborativa y/o en interacción con actores externos a la academia (De Rijcke et al., 2016; Hicks, 2012, Hicks, 2013; Becerra, 2014; De Ibarrola, 2012; Rivera et al., 2011; Galaz Fontes y Gil Antón, 2009; Amaro, Corona

y Soria, 2008; Galaz Fontes et al., 2008; Grediaga Kuri, 1998; Fortes Besprosvani, 2012), y la devaluación de la investigación aplicada (De Rijcke et al., 2016; Hicks, 2012; Martin y Whitley, 2010; Valenti et al., 2013; Valenti, 2011; Ruiz Gutiérrez, 2012) como consecuencias probablemente derivadas de esos sistemas de evaluación.

La literatura también ha analizado diversas consecuencias plausibles de esas evaluaciones sobre la calidad de la investigación, su relevancia conceptual, su pertinencia u orientación, su originalidad, los períodos temporales en que se desarrolla la investigación y el grado de heterogeneidad en la combinación de conocimientos. Es decir, sobre las características epistemológicas de los procesos de producción de conocimiento de acuerdo con Gläser y Laudel (2007b).

Así, la literatura destaca cierto deterioro de la calidad de la investigación por la presión sobre la cantidad de publicaciones que exigen esos sistemas de evaluación (Müller y De Rijcke, 2017; De Rijcke et al., 2016; Vessuri, Guédon y Cetto, 2014; Martin y Whitley, 2010; Izquierdo, 2006; Drucker Colin, 2005; Ramírez Amador, 2012).

Diversos estudios han planteado que la preocupación sobre la relevancia conceptual de las contribuciones intelectuales queda relegada en los procesos de generación de conocimiento. La presión por obtener resultados publicables induce al desarrollo de investigaciones poco ambiciosas y genera dificultades para el desarrollo de investigaciones de largo plazo (Müller y De Rijcke, 2017; Wang, Veugelers y Stephan, 2017; Leisyte y Westerheijden, 2014; van Dalen y Henkes, 2012; Martin y Whitley, 2010; Bensusán et al., 2014; Izquierdo, 2006; Drucker Colin, 2005; Ramírez Amador, 2012).

A la vez, los estudios describen ciertas estrategias implementadas por los investigadores para maximizar el número de publicaciones, entre ellas la subdivisión de los temas de investigación en pequeños sub-temas publicables rápidamente (van Dalen y Henkens, 2012; Leisyte y Westerheijden, 2014; Bensusán et al., 2014; Heras, 2005; Ruiz Herrera, 2005), la re-publicación, con cambios relativamente menores, de trabajos ya publicados (Leisyte y Westerheijden, 2014; Bensusán et al., 2014; Heras, 2005; Hernández-Calderón, 2005) y/o la publicación prematura (Martin y Whitley, 2010; Membrillo Hernández, 2005). También induce a que los investigadores seleccionen temas de investigación menos riesgosos y con ello, una relativa disminución de la originalidad de los resultados de investigación (Wang, Veugelers y Stephan, 2017; Müller y De Rijcke, 2017; Chavarro, Tang y Rafols, 2017; De Rijcke et al., 2016; Leisyte y Westerheijden, 2014; Martin y Whitley, 2010; Bensusán et al., 2014).

Los trabajos también problematizan una tendencia marcada por la menor pertinencia social, productiva y/o para la política pública de los resultados de la investigación; la presión por obtener resultados publicables internacionalmente induce a una producción de conocimiento internacionalizada y menos contextualizada (Müller y De Rijcke, 2017; Chavarro, Tang y Rafols, 2017; De Rijcke et al., 2016; Bianco, Gras y Sutz, 2016; Vessuri, Guédon y Cetto, 2014; van Dalen y Henkes, 2012; Hicks, 2012 y 2013; Bensusán et al., 2014; Esteinou, 2012). Finalmente, los estudios también sugieren que la investigación tiende a estar más concentrada en torno a objetivos disciplinares que inter/multi/transdisciplinares (Leisyte y Westerheijden, 2014; Rafols et al., 2012; Martin y Whitley, 2010; Bensusán et al., 2014; Sarukhán Kermez, 2005; Asomoza Palacio, 2005).

El mecanismo central que conecta a la evaluación del desempeño de la investigación con la dirección contenido y características epistemológicas de la producción de conocimiento, es la respuesta estratégica de los investigadores sobre sus decisiones de investigación para adaptarse a las presiones que le impone su entorno (Gläser y Laudel, 2007b; van der Most, 2010). En la medida que dichos sistemas de evaluación del desempeño de la investigación están asociados a incentivos materiales –mejor acceso a recursos para la investigación- y simbólicos –reconocimiento y prestigio-, el impacto probable del sistema de evaluación sobre las decisiones estratégicas de investigación que toman los investigadores es mayor. Esto explica que sean notables las consecuencias probables de la evaluación sobre la dirección, contenido y características epistemológicas de la producción de conocimiento. La evaluación de la investigación orienta la acción de los individuos, fusionando así acciones de política con la conducta y el contenido de la investigación (Gläser y Laudel, 2007b).

La magnitud de esas consecuencias probables sobre la producción de conocimiento, derivadas de la implementación de los sistemas de evaluación del desempeño de la investigación, son contexto-dependientes. Es decir, dependen de las características específicas que asuman el sistema público de CTI, el sistema de evaluación y los diversos campos científicos (Whitley, 2007), en cada uno de los contextos nacionales en que se implementen dichos sistemas de evaluación. Existen marcadas diferencias que caracterizan a los distintos sistemas públicos de CTI entre los países desarrollados y en desarrollo, y en las especificidades en los diseños de sus sistemas de evaluación del desempeño de la investigación. Sin embargo, es posible identificar un patrón común en la implementación de todos esos sistemas de evaluación de la investigación que consiste en el uso de indicadores cuantitativos, en particular bibliométricos, como criterios centrales de evaluación para aproximar la calidad e impacto de los resultados de la investigación y la productividad de individuos, grupos u organizaciones productoras de conocimiento. De acuerdo con Hicks et al. (2015), la evaluación del desempeño de la investigación se ha convertido en una actividad cada vez más rutinaria y centrada en el uso de ciertas métricas derivadas de bases de datos bibliográficas internacionales tales como WoS o SCOPUS. Así, las métricas son cada vez más importantes en la gobernanza de la ciencia (Hicks et al., 2015).

Los indicadores en general y los bibliométricos en particular, cuando están asociados a incentivos como en el caso de los sistemas de evaluación del desempeño de la investigación, generan diversas consecuencias –entre ellas las sintetizadas al inicio de esta introducción- que Dahler-Larsen (2014), con fines analíticos y metodológicos, clasifica en consecuencias performativas y constitutivas (Dahler-Larsen, 2014)<sup>1</sup>. Los indicadores miden aspectos puntuales de un todo más amplio, de aquí que pueden observarse efectos constitutivos de los indicadores. Conforme el indicador retiene menor cantidad de dimensiones del fenómeno complejo que busca medir, disminuye su validez. Es decir, la relación entre la medición y el fenómeno que intenta capturar es baja. Cuando los

---

<sup>1</sup> Desde la perspectiva de la acción racional orientada a fines, Merton (1936) refiere a las *consecuencias no anticipadas* de la acción social. Desde la economía a ese fenómeno se lo refiere como *efectos secundarios –side effects-* de las políticas y sus instrumentos. En la literatura sobre evaluación de la investigación es posible identificar el uso de diversos términos para referir a ese fenómeno, entre ellos *efectos perversos*, *efectos no intencionados*, *efectos no deseados*, etc. como consecuencias de los indicadores de desempeño utilizados en los instrumentos de política de evaluación de la investigación. Esos términos, de acuerdo con Dahler-Larsen (2014), tienen problemas teóricos y metodológicos para su observación porque dependen de la identificación de las intenciones detrás de los indicadores y, a la vez o en ese sentido, no atienden el problema de la validez que es intrínseco a los indicadores. De aquí el aporte en términos de la precisión teórica y metodológica del concepto *efecto constitutivo* introducido por Dahler-Larsen (2014). Debido a esa precisión, en el contexto de esta tesis se prefiere y utiliza la conceptualización propuesta por este autor.

indicadores, que buscan medir desempeño, carecen de validez y se utilizan regularmente para categorizar individuos, grupos u organizaciones y con ello, otorgar reconocimiento, mayor prestigio o mejor acceso a recursos, se generan las condiciones para *la paradoja del rendimiento* (Dahler-Larsen, 2014).

En resumen, los sistemas de evaluación fuertes del desempeño de la investigación, centrados en indicadores, generan efectos constitutivos sobre el contenido, dirección y características epistemológicas de los procesos de producción de conocimiento. Este tema, a pesar de su importancia para la ciencia, ha sido relativamente poco explorado (Müller y De Rijcke, 2017; De Rijcke et al., 2016; Bianco, Gras y Sutz, 2016; Gläser, 2012). Tal como señalan De Rijcke et al. (2016), la agenda de investigación que explora esas relaciones es relativamente nueva y si bien se conocen algunas de las implicancias de las evaluaciones cuantitativas, el conocimiento que se tiene en la actualidad todavía es escaso.

Esta tesis se inserta en estas lagunas de conocimiento respecto a las relaciones entre la evaluación del desempeño de la investigación y la producción de conocimiento. En particular, se estudian las relaciones entre la evaluación del desempeño individual de investigación y los modos diferenciales de producción de conocimiento –Modo 1 y Modo 2- caracterizadas por Gibbons et al. (1994). Estas relaciones no han sido analizadas por ninguno de los dos cuerpos de la literatura antes mencionados. Además, si bien diversos estudios han sugerido la homogeneización y estandarización de los tipos de producción de conocimiento, ninguno de ellos ha reportado evidencia específica sobre las relaciones entre la evaluación del desempeño de la investigación y los modos 1 y 2 de producción de conocimiento.

El objetivo general de investigación es explorar la naturaleza de las relaciones inmersas entre un sistema de evaluación fuerte del desempeño individual de investigación y los dos modos diferenciales de producción de conocimiento –Modo 1 y Modo 2- caracterizados por Gibbons et al. (1994) en el contexto del SNI de México.

La pregunta general de investigación de esta tesis es: ¿qué relaciones hay entre un sistema de evaluación fuerte del desempeño individual de investigación como el SNI y los Modos 1 y 2 de producción de conocimiento que predominantemente orientan la investigación que realizan los investigadores categorizados en ese sistema de evaluación y cuál es la naturaleza de esas relaciones?

Para poder explorar la naturaleza de las relaciones y así dar respuesta a la pregunta general de investigación planteada, se construyeron –con base en la teoría y los hallazgos de investigaciones pasadas- cinco dimensiones conceptuales, dos de ellas vinculadas a los modos diferenciales de producción de conocimiento y las restantes tres asociadas a diversos aspectos del sistema de evaluación, saber:

- i) *propensión modo 1*, como una aproximación para observar, de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento de los investigadores categorizados en el SNI, el fenómeno caracterizado por el concepto Modo 1 de producción de conocimiento;
- ii) *propensión modo 2*, como una aproximación para observar, de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento de los



- investigadores categorizados en el SNI, el fenómeno caracterizado por el concepto el Modo 2 de producción de conocimiento;
- iii) *flexibilidad epistemológica del sistema de evaluación*, como una aproximación para observar, de acuerdo con la percepción que tienen los investigadores categorizados en el SNI, en qué medida el sistema de evaluación permite ciertas características epistemológicas de la producción de conocimiento -producción de conocimiento de calidad, diversa, original, de largo plazo y colectiva o de colaboración-;
  - iv) *satisfacción con el proceso de evaluación*, como una aproximación para observar el nivel de satisfacción, que tienen los investigadores categorizados en el SNI, respecto a si el proceso de evaluación es capaz de reconocer el desempeño, es decir, si cuenta con la información necesaria para elaborar juicios académicos apropiados sobre el desempeño de los investigadores e incluye todas sus contribuciones importantes, si los criterios y mecanismos de selección de los evaluadores garantizan evaluaciones académicas apropiadas en las diversas áreas de conocimiento y si la composición de las comisiones dictaminadoras respetan los equilibrios regionales e institucionales y dan garantías respecto a la objetividad de la evaluación; y,
  - v) *dirección de la reforma del sistema de evaluación*, como una aproximación para observar el interés, de acuerdo con la percepción de los investigadores categorizados en el SNI, de reformar ese sistema de evaluación en una dirección que cambia la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, mejor consideración de la investigación inter/multi/transdisciplinaria y que valore más los resultados de la investigación aplicada.

Desde el punto de vista teórico-conceptual, el análisis de las relaciones que se estudian en esta tesis, se justifica y fundamenta en la integración que se hace de las formas diferenciales de producción de conocimiento de Gibbons et al (1994) con la teoría que vincula a la evaluación del desempeño de la investigación y la producción de conocimiento (Whitley, 2007 y 2010; Gläser y Laudel, 2007b; Gläser, 2012), así como en los hallazgos de investigaciones pasadas respecto a las consecuencias probables de la evaluación sobre la dirección, contenido y características epistemológicas de la producción de conocimiento documentadas en la literatura. Los conceptos Modo 1 y Modo 2 se seleccionaron y retuvieron como tipos ideales (Weber, 1992 [1922]) de producción de conocimiento para comprender los rasgos sobresalientes de los ámbitos y motivaciones que orientan la acción de los investigadores en la generación de conocimiento y, con ello, aproximar una medida para observar sus modos diferenciales de producción. Importa resaltar que aunque se trata de dos formas diferenciales de producción de conocimiento en términos de su organización, legitimación, orientación y contextos de producción, ambas implican la producción de conocimiento científico. Sin embargo y a pesar de lo anterior, los resultados de la producción de conocimiento científico Modo 2 no siempre son igualmente reconocidos y valorados como los resultados que se derivan de producción Modo 1.

Los objetivos específicos de investigación son explorar la naturaleza de las relaciones (significancia estadística, magnitud, dirección y su implicancia conceptual) entre las cinco dimensiones involucradas en el análisis, a saber:

- i) la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en SNI y su *propensión modo 1* de producción de conocimiento;
- ii) la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en SNI y su *propensión modo 2* de producción de conocimiento;
- iii) la *satisfacción con el proceso de evaluación* del SNI que tienen los investigadores categorizados en ese sistema y su *propensión modo 1* de producción de conocimiento;
- iv) la *satisfacción con el proceso de evaluación* del SNI que tienen los investigadores categorizados en ese sistema y su *propensión modo 2* de producción de conocimiento;
- v) la *satisfacción con el proceso de evaluación* del SNI que tienen los investigadores categorizados en ese sistema y la *flexibilidad epistemológica del sistema de evaluación* que perciben; y,
- vi) la *dirección de la reforma del sistema de evaluación* del SNI según el interés de los investigadores categorizados en ese sistema y la *satisfacción con el proceso de evaluación* del SNI que dichos investigadores tienen.

De acuerdo con la teoría que vincula a la evaluación del desempeño de la investigación y la producción de conocimiento, las contribuciones de Gibbons et al. (1994) y los hallazgos de investigaciones pasadas respecto a las consecuencias probables de la evaluación sobre la dirección, contenido y características epistemológicas de la producción de conocimiento documentadas en la literatura, a continuación se plantean seis hipótesis, sobre cada una de las relaciones planteadas en los objetivos específicos, que se ponen a prueba en esta investigación:

- ✓ H1: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *flexibilidad epistemológica del sistema de evaluación*.
- ✓ H2: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.
- ✓ H3: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *flexibilidad epistemológica del sistema de evaluación*.
- ✓ H4: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *satisfacción con el proceso de evaluación*.
- ✓ H5: la *flexibilidad epistemológica del sistema de evaluación* se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.
- ✓ H6: la *satisfacción con el proceso de evaluación* se relaciona significativamente y negativamente con la *dirección de la reforma del sistema de evaluación*.

Combinando H1, H2, H3, H4, H5 y H6 se hipotetiza que la flexibilidad epistemológica del sistema de evaluación del SNI percibida por los investigadores categorizados en ese sistema y la satisfacción que tienen con el proceso de evaluación que allí se realiza dependen simultáneamente, significativamente y en direcciones opuestas de los modos diferenciales de producción de conocimiento de los investigadores. A la vez, la satisfacción con el proceso de evaluación que tienen dichos investigadores, también depende significativamente y positivamente de la flexibilidad epistemológica del sistema de evaluación que perciben; y, además y simultáneamente, su interés de reformar el sistema de evaluación del SNI en la dirección planteada depende significativamente y negativamente de la satisfacción con el proceso de evaluación que tienen.

Desde el punto de vista de las implicancias conceptuales de las hipótesis planteadas, se argumenta que si no se encuentra evidencia que contradiga dichas hipótesis, entonces los resultados muestran que el SNI no es igualmente flexible desde el punto de vista de las características epistemológicas de la producción de conocimiento que permite, ni es capaz de reconocer y valorar de igual manera los modos diferenciales de producción de conocimiento. En este sentido y en coincidencia con la literatura, se habrá encontrado evidencia que sugiere que el SNI, en tanto sistema de evaluación fuerte e informado fundamentalmente por indicadores bibliométricos, está alterando la dirección y el contenido de la producción de conocimiento. Específicamente, altera el balance entre los modos diferenciales de producción de conocimiento, ya que al tiempo que recompensa cierto modo de producción de conocimiento desconoce –o tiene dificultades para reconocer- otro. Es decir, el SNI tiene sesgos en su evaluación que promueven o incentivan más la producción de conocimiento Modo 1, que la que predominantemente se orienta en Modo 2. Dicho en otros términos, estos resultados sugieren consecuencias constitutivas del SNI sobre el Modo 2 de producción de conocimiento. Así y desde el punto de vista de las implicancias de política de esos resultados de investigación, se puede argumentar que la evaluación que se implementa en el SNI, más que ayudar al cumplimiento integral de sus objetivos de política, los contradice. En particular, compromete el alcance de aquellos objetivos de política del SNI orientados a incentivar la producción de conocimiento científico y tecnológico de calidad para contribuir al desarrollo social y económico del país.

Para alcanzar el objetivo de investigación, en este trabajo se desarrolla y valida un modelo teórico estructural (Kline, 2011; Hair et al., 2014; Brown, 2015) sobre las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento, que permite contrastar con la realidad –representada por los datos- cada una de las hipótesis planteadas de forma simultánea. Con este modelo, y en base a la evidencia que provee el caso mexicano, se busca, por un lado, dar una respuesta rigurosa a la pregunta planteada y, por otro lado, aportar conocimiento relevante, original, válido y confiable sobre las relaciones estudiadas. Importa destacar que, hasta donde se sabe, no existe ningún trabajo, ni en el ámbito nacional ni en el internacional, que haya estudiado con anterioridad estas relaciones.

Los datos para validar la teoría que se propone y que se representa formalmente en el modelo teórico estructural, provienen de la encuesta titulada “*Percepciones sobre la evaluación académica*” (FCCyT, 2014). La encuesta contiene información sobre las percepciones de los investigadores categorizados en el SNI al año 2013 en relación a un conjunto de aspectos vinculados a ese sistema de evaluación y sobre sus ámbitos y motivaciones para la producción de conocimiento. La encuesta contó con una tasa de respuesta de 43% (N=8055) y es representativa según género, nivel y área de conocimiento (Bensusán et al., 2014). Las estimaciones se realizaron con el paquete estadístico M-PLUS versión 7.4 (Muthén y Muthén, 1998-2015).

La elección por el modelado estructural se fundamenta en la complejidad y multidimensionalidad que encierra el fenómeno estudiado en esta investigación y en las posibilidades que ofrece en términos de la construcción de medidas válidas y confiables. El desarrollo y validación del modelo teórico estructural requiere, como paso previo, el desarrollo y validación de un modelo teórico de medida, por lo que se implementan un conjunto de procedimientos, métodos y técnicas, para garantizar la validez de constructo de la medición (la validez de las cinco dimensiones planteadas,

relevantes para el objetivo de este trabajo). Sin medidas (constructos o variables latentes) válidas y confiables, difícilmente puedan extraerse resultados rigurosos.

Así, en esta investigación se realizaron esfuerzos importantes para garantizar la validez de constructo de todas las medidas que se proveen. La relevancia de estos esfuerzos está en términos de la reflexión metodológica y técnica que aporta esta investigación al campo de los estudios sobre evaluación del desempeño de la investigación y al ámbito de la política pública de estímulo a la investigación y sus instrumentos de evaluación. De aquí también la relevancia conceptual y pertinencia práctica de esta investigación.

Finalmente, se espera que el análisis y los resultados de esta investigación contribuyan al debate público sobre el SNI y la importancia de estimular también una agenda de producción de conocimiento propia, contextualizada, de calidad, diversa y plural en México. Conjuntamente con ello, también se espera contribuir a la discusión informada en el ámbito de la política pública de CTI. En particular, sobre las dimensiones (flexibilidad epistemológica del sistema de evaluación) y los parámetros a re-considerar (criterios de evaluación y su validez) en el diseño e implementación de su sistema de reconocimiento y recompensa de la ciencia, para que la producción de conocimiento que se realiza en México pueda contribuir también al alcance de los objetivos del desarrollo nacional.

Este documento se estructura en once capítulos. Le sigue a esta introducción, el capítulo II que revisa los aportes teóricos realizados desde la sociología de la ciencia, en particular las contribuciones de Ziman (2000) y Gibbons et al. (1994) sobre la organización y orientación de la producción de conocimiento, así como sus formas diferenciales de producción y sus requerimientos específicos de evaluación. Este capítulo concluye con una fundamentación que justifica la selección que se hace en esta tesis de los conceptos Modo 1 y Modo 2 como tipos ideales de producción de conocimiento. Así, este capítulo representa el primer paso para avanzar hacia la construcción del marco teórico conceptual de esta tesis.

El capítulo III se centra en la revisión de las contribuciones sobre la teoría que vincula a la evaluación del desempeño de la investigación y la producción de conocimiento. Siguiendo a Whitley (2007) se describen cinco grandes consecuencias probables de los sistemas de evaluación de la investigación sobre la organización y la dirección de la producción de conocimiento. De acuerdo con Gläser y Laudel (2007b) se describe el mecanismo que conecta a la evaluación de la investigación con el contenido y características epistemológicas de la producción de conocimiento. Se argumenta la relevancia del sistema público de ciencia, del sistema de evaluación de la investigación y de los diversos campos científicos como dimensiones contextuales importantes para entender el impacto esperado de la evaluación sobre el contenido y la dirección de la producción de conocimiento (Whitley, 2007) y se revisan las contribuciones de Whitley (2010) y Gläser (2012) que destacan que los cambios que se observen sobre la dirección y el contenido de la producción de conocimiento dependen, en última instancia, de la gobernanza de la ciencia que tiene lugar en cada contexto específico. El capítulo concluye con una reflexión en torno a los elementos teóricos que permiten integrar las formas diferenciales de producción de conocimiento (Gibbons et al., 1994) con la teoría que vincula la evaluación de la investigación con la producción de conocimiento, y que justifican y fundamentan teóricamente el análisis de las relaciones estudiadas en esta tesis. El capítulo finaliza con la figura 1 que sintetiza las relaciones teóricas entre la evaluación del desempeño de la

investigación y los modos diferenciales de producción de conocimiento. Así este capítulo representa el segundo paso para avanzar hacia la construcción del marco teórico conceptual de esta tesis.

El capítulo IV ofrece una revisión bibliográfica amplia sobre diversos aspectos vinculados con la evaluación del desempeño de la investigación y la producción de conocimiento. Por una parte se revisa la evolución y amplia difusión de los indicadores bibliométricos en la gobernanza de la ciencia, el debate internacional sobre su uso como criterios centrales de evaluación, sus limitaciones y se plantea el problema de la validez de los indicadores en general y de los bibliométricos en particular. Por otra parte, se revisa la literatura especializada internacional y nacional que documenta diversos efectos constitutivos sobre la dirección, contenido y características epistemológicas de la producción de conocimiento como consecuencia probable de los sistemas de evaluación del desempeño de la investigación centrados en indicadores bibliométricos. Este capítulo concluye con una síntesis de los hallazgos de investigaciones que inspiran la construcción de las dimensiones relevantes para el análisis que se hace y que dan sentido a las hipótesis que se ponen a prueba en esta tesis. Así, este capítulo representa el tercer paso hacia la construcción del marco teórico conceptual de esta tesis.

En el capítulo V se construye el marco teórico conceptual que justifica y provee los fundamentos teóricos sobre los que se elabora tanto las definiciones conceptuales de las dimensiones relevantes para el análisis que se hace en esta investigación, así como las relaciones que se exploran entre ellas (hipótesis). Para ello se retoman e integran ciertos conceptos y las relaciones entre ellos que emergen de la extensa sistematización realizada de teoría, debate internacional y hallazgos de investigaciones pasadas revisadas en los capítulos II, III y IV. Este capítulo finaliza con la figura 2 que representa el marco teórico conceptual construido, sintetizando las dimensiones conceptuales y las hipótesis que las vinculan. Así, la figura 2 representa gráficamente la teoría estructural sobre las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento, que se propone y que se ponen a prueba en esta investigación.

En el capítulo VI, de acuerdo con el marco teórico conceptual construido, se presenta una caracterización del sistema público de CTI de México y de su Sistema Nacional de Investigadores. Con base en la revisión de distintos estudios y diversas fuentes de información estadística, la caracterización que se realiza del sistema de CTI de México se centra en dos dimensiones contextuales, a saber: i) la naturaleza de su régimen de financiamiento de la investigación; y ii) las características específicas del mercado de trabajo de los investigadores en México. Asimismo, con base en la sistematización de la información contenida en el Acuerdo de creación del SNI y sus modificaciones, los Reglamentos de operación y funcionamiento del Sistema y en los Criterios Internos de evaluación de las Comisiones Dictaminadoras por áreas de conocimiento, la caracterización que se realiza del sistema de evaluación del SNI se enfoca en dos dimensiones contextuales, a saber: i) su gobernanza y estructura; y ii) las implicancias del sistema de evaluación en la asignación de recursos para la investigación. Este capítulo concluye con los rasgos sobresalientes que caracterizan al sistema público de CTI de México y a su Sistema Nacional de Investigadores y con ello una reflexión sobre cuáles podrían ser los impactos esperados del SNI sobre la dirección y contenido de la producción de conocimiento. Así este capítulo es útil no solamente para contextualizar y entender mejor los resultados de esta investigación, sino también para contextualizar los análisis sobre los logros y desafíos del SNI y las tensiones sobre ese sistema, que emergen de dos grandes debates públicos y que se sistematizan en el siguiente capítulo.

En el capítulo VII se revisan y sistematizan dos debates públicos sobre el SNI, uno realizado en 2005 a 20 años de su creación, y el otro, en 2012, cerca de una década después. En ellos participaron hacedores de políticas públicas, gestores de la ciencia, autoridades, asesores, evaluadores, investigadores y especialistas de diversas disciplinas, quienes debatieron sobre los *logros, desafíos y direcciones de cambio* para el SNI. Con fines analíticos, la rica información que emerge de estos debates fue organizada de manera tal que facilitara la identificación de las principales consecuencias performativas y constitutivas probables del SNI y, simultáneamente, fuera congruente con las dimensiones conceptuales de interés a los objetivos de esta investigación, sus relaciones y condicionantes. El capítulo concluye con una reflexión sobre los fundamentos que resultan del ámbito mexicano y que justifican la coherencia, adecuación y pertinencia del marco teórico conceptual construido y las hipótesis que de éste se derivan sobre las relaciones entre la evaluación del desempeño de la investigación –en particular del SNI- y las formas diferenciales de producción de conocimiento observables en México.

El capítulo VIII se concentra en el diseño y la estrategia de investigación seguida en este trabajo. Se describen en profundidad los diversos métodos, procedimientos, técnicas de análisis multivariado implementados y la fuente de información utilizada en las estimaciones y en la validación de los modelos de medida y estructural.

En el capítulo IX, con base en los datos contenidos en la encuesta, se describe estadística y analíticamente las percepciones que tienen los investigadores categorizados en el SNI sobre: i) los ámbitos y motivaciones de la producción de conocimiento que realizan (sus propensiones modo 1 y modo 2 de producción de conocimiento); ii) la satisfacción que tienen con el proceso de evaluación del SNI; iii) la flexibilidad epistemológica que perciben del sistema de evaluación al que están sometidos; y iv) la dirección de la reforma del sistema de evaluación. Los resultados del análisis descriptivo realizado se presentan también según ciertas variables de control (género, tramos de edad, nivel de categorización en el SNI, área de conocimiento y tipo de institución en la que trabajan los investigadores).

En el capítulo X se pone a prueba el marco teórico conceptual construido a través de su formalización en un modelo de ecuaciones estructurales. Se presentan, analizan e interpretan todos los resultados de las estimaciones de los Modelos de Medida y Estructural, destacando la especificación, confiabilidad y validación de los resultados obtenidos. Así, en base a la teoría y la evidencia empírica obtenida, se analiza la naturaleza de las relaciones existentes entre las dimensiones conceptuales asociadas con la evaluación del desempeño individual de investigación (*flexibilidad epistemológica del sistema de evaluación, satisfacción con el proceso de evaluación y dirección de la reforma del sistema de evaluación*) y los modos diferenciales de producción de conocimiento que orientan la investigación de los investigadores categorizados en el SNI mexicano (*propensión modo 1 y propensión modo 2*).

Finalmente, en el capítulo XI se concluye con una síntesis de los resultados de investigación obtenidos y se reflexiona sobre sus implicancias para la política pública de CTI, así como sobre las principales contribuciones conceptuales, metodológicas, empíricas de la investigación realizada, su alcance y limitaciones, y, finalmente, se plantean las preguntas y líneas de investigación a futuro que se abren a partir de esta investigación.

## CAPÍTULO II: Los cambios en la organización y orientación del conocimiento: las formas diferenciales de producción y su relación con la evaluación

En la literatura diversos autores discuten las transformaciones en la organización de la producción de conocimiento internacional, observables desde las últimas décadas del SXX en los países desarrollados altamente industrializados. Más específicamente, dichos autores desarrollaron distintos enfoques que emergen como marcos conceptuales analíticos y descriptivos para entender los cambios que se generan en las sociedades contemporáneas respecto al rol del conocimiento, y sus formas de producción, legitimación y difusión. Algunos de estos enfoques, en el contexto de los países en desarrollo o economías industrialmente menos avanzadas, más que descriptivos de la realidad son normativos, para la orientación en el diseño e implementación de sus políticas públicas de CTI.

Como punto de partida o base de referencia -tomada por los autores para identificar las transformaciones en la organización de la producción de conocimiento-, se puede ubicar un primer enfoque clásico y tradicional, que se asocia con lo que Polanyi (1962) denominó como la *República de la Ciencia* para caracterizar la organización de la *ciencia académica*. Desde finales del SXVII, la organización de la *ciencia académica* estuvo vinculada con la universidad moderna como expresión ideal de la *Iluminación* o de la *República de la Ciencia* (Delanty, 1998). En ese marco, los científicos buscaron constituirse como una comunidad autónoma y autodirigida (Albornoz, 2007). El conocimiento se caracterizó por su autonomía, y su rol era la emancipación de la humanidad. El modelo de la ciencia académica era coherente, autónomo, trascendente y auto-referenciado (Delanty, 1998).

Las normas de la *ciencia académica* las planteó Merton en 1942, en su trabajo titulado *The Normative Structure of Science*. En ese trabajo presenta los cuatro principios básicos en los que se basa el *ethos* de la ciencia. El *ethos* de la ciencia refiere a las normas y valores éticos compartidos por todos los científicos. Muy sintéticamente, los cuatro principios son: i) *Communism*: los descubrimientos científicos son de propiedad común. Los científicos renuncian a los derechos de propiedad intelectual a cambio de reconocimiento y estima de la comunidad científica; ii) *Universalism*: la verdad se evalúa en términos de criterios universales e impersonales y no en función de la raza, clase, género, religión o nacionalidad; iii) *Disinterestedness*: los científicos son recompensados por actuar desinteresadamente; y iv) *Organized Skepticism*: el método de la ciencia supone que todas las ideas deben ser probadas, están sujetas a una estructura rigurosa y al escrutinio de la comunidad. Más sintéticamente, los cuatro principios son referidos como *CUDOS*.

Este capítulo tiene por objetivo revisar algunas de las contribuciones teóricas de la sociología de la ciencia tanto sobre los cambios en la organización y orientación del conocimiento, así como sus formas diferenciales de producción y sus requerimientos específicos de evaluación. Este capítulo representa el primer paso para avanzar hacia la construcción del marco teórico conceptual que justifica y provee los fundamentos teóricos sobre los que se construye el análisis que se realiza en esta tesis.

Para ello, el capítulo se organiza en cuatro secciones. Le sigue a esta introducción, la sección II.1 que sistematiza el enfoque de Ziman (2000) sobre las transformaciones en la producción de

conocimiento derivadas de la transición de la ciencia académica a lo que denominó como *ciencia posacadémica*. En la sección II.2 se revisan las contribuciones de Gibbons et al. (1994) sobre las formas diferenciales de producción de conocimiento y su relación con la gobernanza de la ciencia. En particular, se plantean las características diferenciales de los Modo 1 y Modo 2 de producción de conocimiento y los requerimientos específicos vinculados al control de calidad de la producción de conocimiento Modo 2. En la sección II.3 se profundiza, a partir de la revisión de varios enfoques, sobre la evaluación de la investigación en el contexto de aplicación. Finalmente, la sección II.4 concluye con una discusión de los principales conceptos que emergen de esta revisión teórica y su relación con la evaluación, que permite fundamentar las razones por las que en esta tesis se seleccionan y retienen los conceptos Modo 1 y Modo 2 como tipos ideales de producción de conocimiento.

## II.1 De la *ciencia académica* a la *ciencia posacadémica*: características cambiantes de la producción de conocimiento

Ziman (2000) refiere a los cambios en la producción de conocimiento internacional, como “*transformaciones radicales e irreversibles a nivel global de las formas en que la ciencia se organiza, gestiona y desarrolla*” (Ziman, 2000: 67). Su enfoque trata sobre el cambio desde la *ciencia académica* hacia la *ciencia posacadémica*. Esas transformaciones tienen lugar en los países altamente industrializados y, para el autor, se pueden percibir desde fines de la década de 1960. Responden a diversos factores sociales: i) los externos al ámbito académico -tales como presiones políticas, económicas e industriales- pero que inciden en la comunidad científica; y ii) los internos, en tanto la ciencia es un sistema dinámico que debe adaptarse socialmente a las tensiones que se acumulan y generan en su interior, derivadas del progreso científico y tecnológico.

A continuación, en las sub-secciones II.1.1 y II.1.2 se describen los factores sociales internos. Estos factores están vinculados a ciertas características diferenciales de la forma en que se da la producción de conocimiento (su organización y orientación) en el marco de -lo que Ziman denomina como- la *ciencia posacadémica*. Ziman (2000) refiere a estos factores como la *colectivización y explotación del conocimiento*. Para el autor, estos factores internos tienen incidencia en la transformación de la ciencia académica en posacadémica, porque la ciencia es un sistema social dinámico que se adapta a los cambios derivadas del avance científico y tecnológico.

En la sub-sección II.1.3 se presentan los factores sociales externos analizados por Ziman (2000) que, según este autor, también inciden en la transformación de la ciencia académica hacia la posacadémica. El autor refiere a ellos como: i) los *límites del crecimiento*, ii) *la política científica y tecnológica*; y iii) *la industrialización de la ciencia*. En el contexto de este trabajo, la *industrialización de la ciencia* se entiende como un caso extremo del rol potencial del conocimiento, que se deriva de una respuesta específica de la política científica y tecnológica de algunos países altamente industrializados ante los límites del financiamiento público para la ciencia.

### II.1.1 La producción colectiva y no disciplinar de conocimiento: un cambio en su organización

Ziman (2000) usa el concepto de *colectivización* para referirse a los procesos de investigación más colectivos, lo que se expresa en el uso compartido de instrumentos y aparatos, en diversas formas de colaboración entre investigadores -grupos de trabajo, redes, etc.-, y en el aumento de artículos científicos con dos o más co-autores. Según el autor, esas expresiones son las consecuencias sociales de la acumulación del conocimiento y la técnica. “*La ciencia ha progresado a un nivel cuyos*



*problemas no pueden ser resueltos por individuos que trabajan independientemente”* (Ziman, 2000: 70).

Además, agrega que la amplia mayoría de los *problemas* no emergen compartimentados en las diversas especialidades existentes de investigación, por el contrario, su abordaje requiere de enfoques transdisciplinarios. Para ejemplificar, el autor se cuestiona sobre el significado científico de la enfermedad de la *vaca loca*, y su reflexión lo conduce a buscar respuesta desde distintas disciplinas (patología veterinaria, neurofisiología, microbiología, etc.). Es decir, para él es necesario un gran esfuerzo colectivo –incluido el ámbito de la política pública- para establecer equipos de investigación multidisciplinarios que coordinen sus esfuerzos y combinen sus hallazgos.

De acuerdo con Ziman (2000), la ciencia posacadémica no se preocupa exclusivamente de la resolución de problemas prácticos, porque aún, los problemas científicos más fundamentales, también requieren de abordajes transdisciplinarios. Actualmente, y como resultado del progreso de la ciencia y la técnica, los problemas de investigación –ya sean fundamentales o prácticos- son cada vez más desafiantes y mayores. Por lo tanto, su abordaje –necesariamente- tiene que estar sostenido en el trabajo colectivo de investigadores especializados en las diversas disciplinas involucradas en la resolución de esos problemas (Ziman, 2000).

El cambio está en que, a diferencia de lo que sucede en la actualidad –dice Ziman-, la ciencia académica responde a una cultura individualista, sostenida por una poderosa estructura disciplinar pero a la vez altamente dispersa. Además, sostiene que si bien eso no es incompatible con el crecimiento de los grupos de investigación, los equipos de trabajo verdaderamente multidisciplinarios desafían la estructura tradicional respecto a la autonomía personal, las perspectivas de carreras académicas, los criterios para ponderar el desempeño o rendimiento, los liderazgos, y los derechos de propiedad intelectual, entre otros (Ziman, 2000).

En síntesis, Ziman (2000) está describiendo un cambio en la organización de la producción de conocimiento. En el marco de lo que él denomina *ciencia posacadémica*, la producción de conocimiento aborda problemas fundamentales y prácticos cada vez más complejas, y por ello necesariamente debe sostenerse en procesos de investigación colectivos y transdisciplinarios o multidisciplinarios<sup>2</sup>.

---

<sup>2</sup> La transdisciplinariedad, interdisciplinariedad y multidisciplinariedad pueden definirse como los grados en que el conocimiento disciplinar se modifica. La transdisciplinariedad representa el grado máximo de integración disciplinar, donde se transgrede y transforma lo disciplinar, creando nuevos marcos conceptuales que proveen una nueva síntesis de ideas y métodos. La interdisciplinariedad representa el grado intermedio de integración disciplinar, donde se integra y mezcla lo disciplinar o donde interactúan las disciplinas existentes. La multidisciplinariedad representa el grado mínimo de integración disciplinar, donde se coordina y pone en secuencia el conocimiento disciplinar, representa la combinación de distintas disciplinas para entender un fenómeno o problema. (Bocco et al., 2014). En el contexto de esta tesis, la multidisciplinariedad, interdisciplinariedad y transdisciplinariedad se entiende como grados de heterogeneidad en la combinación de conocimientos.

### II.1.2 La producción de conocimiento explícitamente orientada por problemas económicos y sociales: la pertinencia social y económica del conocimiento

Otro factor social interno importante destacado por Ziman (2000), en la transición hacia la ciencia posacadémica, se relaciona con la *explotación del conocimiento*. De acuerdo con el autor, cada vez es mayor la presión por la *utilidad* del conocimiento. Desde su perspectiva, dicha presión es posible, por un lado, porque los desarrollos recientes en términos del avance conceptual y práctico muestran que “(...) se están haciendo progresos notables en la combinación imaginativa de conceptos y técnicas provenientes de disciplinas y sub-disciplinas que anteriormente se pensaban distantes (...)”, y por otro lado, debido a que “[m]uchos campos especializados de la ciencia también han entrado en una fase de “finalización”, es decir, una fase donde hay un marco general de comprensión confiable para guiar estratégicamente la investigación hacia los fines deseados” (Ziman, 2000: 73).

Böhme et al. (1983) desarrollan el concepto de “*finalisation science*” para caracterizar las fases de desarrollo de las disciplinas. La fase final de desarrollo teórico de una disciplina es determinada por factores externos. Es decir, cuando la disciplina llega a su madurez teórica, se abre a que su orientación sea determinada de acuerdo con objetivos externos. Así, los desarrollos teóricos futuros –de esa disciplina– seguirán la trayectoria que esos objetivos establezcan. Asimismo, esos autores señalan que debido a que cada vez son más las disciplinas que se encuentran en esta fase, la relación entre ciencia y sociedad está cambiando. Además, sostienen que en esa relación, la sociedad se vuelve un actor mucho más activo y participativo en la orientación y legitimación de la producción de conocimiento.

En el mismo sentido, Ziman (2000) argumenta que diversos actores (gobiernos, empresas, grupos de ciudadanos y el público en general) demandan *arreglos* más sistemáticos orientados a la identificación, estímulo y explotación del conocimiento *potencialmente* útil. Desde su perspectiva, la ciencia posacadémica se ve presionada *explícitamente* a generar valor económico, riqueza para la nación y la economía, pero también para que genere soluciones a problemas de salud, ambientales, sociales, etc. Es decir, que demuestre explícitamente su pertinencia social y/o económica.

Es así que para Ziman (2000), la diferencia entre ciencia académica y posacadémica no está en la pertinencia del conocimiento en sí, sino que lo nuevo está en el requerimiento de que la investigación debe estar explícitamente orientada a problemas prácticos reconocibles (fundamentalmente económicos señala el autor). Es decir, en el carácter estratégico de la ciencia en general y de la investigación en particular. Según Ziman (2000), se espera que los científicos sean conscientes de las aplicaciones potenciales de su trabajo.

Irvine y Martin (1984) definieron el concepto “*investigación estratégica*” como el desarrollo de investigación básica con la expectativa de producir una base de conocimiento lo suficientemente amplia para generar un marco que sostenga la solución actual y futura de problemas prácticos. Similarmente, el concepto “*ciencia estratégica*” (Rip, 2004) internaliza las presiones por la pertinencia. Si bien mantiene la autonomía o la libertad académica en la definición de las líneas de investigación, los investigadores sí tienen en consideración la pertinencia de su trabajo como una forma de legitimación importante.

De acuerdo con Ziman (2000), la *utilidad* es un concepto moral y no puede estar determinado sin una referencia a fines y valores humanos más generales. De aquí la importancia de la discusión que

plantea Hess (2007), sobre quién decide qué ciencia se hace, y por lo tanto también, para quiénes. Según este autor, para comenzar a responder esas preguntas, primero hay que centrarse en el análisis del sistema de recompensas de la ciencia. Este tema se retoma más adelante en este capítulo.

En síntesis hasta aquí, la producción de conocimiento es cada vez más compleja y enfrenta mayores presiones respecto a su pertinencia social y económica. Por estas razones, cada vez con mayor frecuencia, dicha producción se sostiene en procesos de investigación colectivos y transdisciplinarios o multidisciplinarios.

Importa señalar que, las afirmaciones de Ziman (2000) están fundamentalmente asociadas a lo que observa en los países altamente industrializados, y que las *presiones*, en su mayoría, provienen de las políticas públicas implementadas en esos contextos y de las dinámicas de mercado propias de esas economías. En la siguiente sub-sección se aborda este tema.

En el contexto de un país en desarrollo o relativamente menos industrializado como México, la transición hacia una producción de conocimiento con esas características -orientada por su pertinencia social y económica, y organizada colectiva y transdisciplinariamente o multidisciplinariamente- si bien es observable, todavía hoy, enfrenta diversas dificultades para difundirse como una forma de producción de conocimiento válida y legítima. En otras palabras, la producción de conocimiento científico y tecnológico en México, encuentra dificultades para convertirse en el motor del desarrollo social y económico del país. Si bien las razones son múltiples, este trabajo se enfoca en el sistema de recompensas de la ciencia, en particular, en el sistema de evaluación del desempeño individual de los investigadores, implementado a través del Sistema Nacional de Investigadores (SNI); este sistema de reconocimiento al desempeño individual de investigación en México se describe en el capítulo VI.

### II.1.3 La ciencia académica y la ciencia industrial: dos extremos en rol potencial del conocimiento

En esta sub-sección se presentan los factores sociales externos analizados por Ziman (2000) que, según este autor, inciden en la transformación de la ciencia académica hacia la posacadémica, y a los que refiere como: i) los *límites del crecimiento*, ii) *la política científica y tecnológica*, y iii) *la industrialización de la ciencia*.

Según Ziman (2000) la actividad científica no puede mantener su alta tasa de crecimiento exponencial histórica porque los recursos que la financian no aumentan con la misma velocidad. A esta constatación empírica la refiere como los *límites del crecimiento*, y la ubica como un factor social externo para dar cuenta de la transición de la *ciencia académica* hacia la *ciencia posacadémica*.

Para el autor esto es así, porque la transición a un régimen de “*estado estacionario*” respecto al financiamiento de la ciencia, genera tensiones al *ethos* académico y, también, la necesidad de cambios culturales para adaptarse al nuevo entorno académico. El entorno académico está marcado por palabras como *eficiencia y rendición de cuentas* (“*efficiency*” y “*accountability*”), hasta entonces poco familiares en el ámbito de la ciencia académica. (Ziman, 2000)

Estas palabras, junto con otras tales como *competencia* y *transparencia*, provienen de la nueva gestión dominante en el sector público, que tiene lugar desde la década de los ochenta. Específicamente del enfoque conocido como la *nueva gestión de la política pública –new public management-* (Hood, 1991). Power (1996) analiza el creciente papel de las auditorías en las sociedades industriales modernas, y es crítico de la nueva gestión de la política pública porque se ha basado en un estilo dominante de auditoría que no guarda relación con el objeto auditado. En las secciones IV.1 y IV.3 del capítulo IV se retoma la incidencia del *new public management* en el advenimiento de los sistemas de evaluación en la ciencia y en la difusión de cierto tipo de indicadores para la evaluación del desempeño individual de la investigación.

El siguiente factor social externo importante destacado por Ziman (2000) está vinculado con la *política científica y tecnológica*. Desde que una parte importante del financiamiento para la investigación es público, inevitablemente –dice Ziman- se introduce la política en la ciencia y la ciencia en la política. Sin embargo, a los hacedores de políticas les falta conocimiento para hacerlo directamente, por lo que han recurrido a la comunidad científica.

Argumenta que en el pasado, lo han hecho a través de transferencias de dinero directas “en bloques”, desde el gobierno hacia las organizaciones de investigación (universidades, institutos, centros de investigación). Ese mecanismo de asignación de recursos no afectaba significativamente la autonomía de la comunidad científica. En la actualidad, la tendencia es otra. Según Ziman (2000) la asignación de recursos se basa en lo que denomina como un sistema de “*soft money*”, inventado en Estados Unidos después de la Segunda Guerra Mundial.

Este sistema se basa en la formulación –por parte de los investigadores- de proyectos de investigación, que son presentados a fondos competitivos de los organismos de financiación. Los proyectos son evaluados mediante revisión por pares y los recursos se asignan en función de los méritos científicos. Así, los científicos académicos titulares –*tenured-* pueden realizar sus investigaciones, sujetas únicamente a la condición razonable de que hayan sido juzgadas como “buena ciencia” por los expertos calificados –*pares-* (Ziman, 2000). Este sistema proporciona los recursos para la investigación mediante un mecanismo social que es consistente con el *ethos* académico. A la vez, la competencia por esos recursos refuerza la competencia por el reconocimiento científico, porque tanto la asignación de recursos como el reconocimiento dependen de una evaluación que se realiza esencialmente con los mismos criterios. Con este mecanismo, los gobiernos dan libertad a los científicos para establecer sus agendas de investigación y publicar sus resultados, lo que a la vez, les da la sensación de que tienen el control de sus vidas intelectuales -sostiene el autor-.

Sin embargo, los científicos se vuelven muy vulnerables a las demandas de quienes pagan sus investigaciones. “*La política científica revela discrepancias evidentes entre lo que la ciencia puede producir y lo que la sociedad realmente obtiene*” (Ziman, 2000: 76). Para superar esas discrepancias, los Estados establecen prioridades e instruyen a las organizaciones de investigación y de financiamiento a incorporarlas en sus planes y plasmarlas en programas específicos de financiamiento. Esto reduce significativamente la autonomía de los científicos. Simultáneamente y a cambio, ese mismo mecanismo proporciona a los gobiernos de comunidades de científicos activos, bien reconocidos y agradecidos. (Ziman, 2000)

De acuerdo con el autor, esa transformación en el *contrato social* de la ciencia -expresada en una mayor dependencia de los investigadores del financiamiento externo para la investigación, ahora basada en proyectos- refuerza el fenómeno conocido en la literatura como *Efecto Mateo* (Merton, 1988).

En el contexto de la política científica y tecnológica de la segunda mitad del SXX, la nueva fuerza motora de la ciencia, sostiene Ziman, es la *competencia* por recursos para la investigación que, a la vez, tiene prioridad sobre la competencia por la credibilidad científica. “*Con tantos investigadores que dependen completamente de estas subvenciones para la investigación o contratos para su sustento personal, ganar[los] se convierte en un fin en sí mismo. Los grupos de investigación se transforman en pequeñas empresas de negocios. El foro metafórico de la opinión científica se convierte en un verdadero mercado de servicios de investigación*” (Ziman, 2000: 76).

Finalmente, el último factor social externo destacado por Ziman (2000) en la transición hacia la ciencia posacadémica –muy vinculado con los dos anteriores- está asociado con lo que denomina como la *industrialización de la ciencia*. Desde su perspectiva, la *ciencia industrial* es la antítesis de la *ciencia académica*. Según el autor, la política de arriba hacia abajo –*top-down policy*- desconoce esa diferencia. Los gobiernos por razones enteramente políticas –dice Ziman-, reducen su apoyo financiero a la ciencia, –entre otras acciones- privatizando sus instituciones dedicadas a la investigación. Es decir, los institutos y laboratorios públicos que realizan investigaciones aplicadas en temas de salud, defensa, transporte o en política social son vendidos a grandes empresas o se convierten en empresas independientes.

En ese marco, sostiene, los científicos no tienen influencia, con suerte mantienen sus empleos y continúan trabajando en los mismos problemas. En el pasado, los institutos públicos de investigación hacían investigación académica con objetivos estratégicos. Conforme pasaron los años, eso ha cambiado debido a que el financiamiento depende de la investigación “contractual”. Los científicos, son científicos industriales. (Ziman, 2000)

Según el autor, la misma tendencia se observa en la insistencia creciente por estrechar vínculos entre la academia y la industria. Pero no como resultado del estímulo a que las empresas realicen investigación básica, sino que la expectativa está puesta en las instituciones académicas. Se espera que estas últimas emprendan más investigaciones con financiamiento industrial y que produzcan resultados con valor comercializable.

Slaughter y Leslie (1997) estudian diversos cambios en las carreras académicas y en las universidades públicas de investigación de finales del SXX en Estados Unidos, Canadá, Reino Unido y Australia. En ese trabajo, desarrollan el concepto “*capitalismo académico*” para describir y explicar la creciente cantidad de actividades de mercado que se observan en las universidades analizadas. Entre los resultados destacan, por un lado, la creciente competencia -como si fuera de mercado- por el acceso a recursos externos, tales como: becas, contratos, fondos de investigación, asociaciones universidad-industria, inversiones en *spin-off*, etc, y por otro, la diferenciación de esas actividades de mercado, en actividades con fines de lucro. El conocimiento se capitaliza, genera ganancias y beneficios económicos a través de patentes, acuerdos de regalías y licencias, empresas *spin-off* y asociaciones con industrias.

Según Etzkowitz (1998) esa capitalización del conocimiento está en el corazón de una nueva misión para la universidad. Esta tercera misión, la actividad empresarial como una nueva función académica, es integrada a las funciones académicas tradicionales de enseñanza e investigación de las universidades. La *universidad de investigación* se transforma en la *universidad empresarial*, como un actor económico por derecho propio, más vinculada con los usuarios del conocimiento. (Etzkowitz, 1998).

Más precisamente, se podría señalar que la universidad empresarial es un nuevo actor económico por derecho propio en ciertos contextos económicos. En esos contextos es que establece vínculos de mercado con ciertos usuarios del conocimiento. No se trata de cualquier contexto, ni tipo de vínculo, ni tampoco de cualquier usuario. En la universidad empresarial, el conocimiento se pone al servicio de quienes tienen la capacidad de pagar por él, con las desigualdades que ello genera en términos de exclusión social (Sen, 2000).

Arocena, Bortagaray y Sutz (2008) cuestionan el enfoque de la *universidad empresarial* tanto en sus términos normativos como empíricos. Por un lado y desde el punto de vista empírico, los autores refieren al caso de América Latina y la desvinculación histórica entre universidad-empresa que en ese contexto se observa. Por otro lado, los autores son críticos del enfoque de la universidad empresarial por sus implicancias normativas para la educación superior en el desarrollo de América Latina. Específicamente porque en el enfoque de la universidad empresarial, el concepto de crecimiento económico está equiparado con un concepto mucho más amplio, con el de desarrollo. Arocena, Bortagaray y Sutz (2008) se alejan de las conceptualizaciones de desarrollo como sinónimo de crecimiento económico, porque lo entienden como *desarrollo humano auto-sustentable*. Es decir, como “*el conjunto de procesos que procuran mejorar la calidad material y espiritual de la vida colectiva, bajo formas ambientalmente sustentables y socialmente auto-sustentables [inclusivas]*” (Arocena, Bortagaray y Sutz, 2008: 239).

A diferencia del enfoque de la *universidad empresarial*, en el enfoque de la *universidad para el desarrollo* (Arocena, Bortagaray y Sutz, 2008), el conocimiento es un instrumento central para el desarrollo económicamente sostenible, ambientalmente sustentable y socialmente inclusivo. En el contexto de esta tesis, cuando se habla de ciencia o producción de conocimiento para el desarrollo, se alude a esta conceptualización de desarrollo.

Hecho el paréntesis y regresando a Ziman (2000), el objetivo de la política científica y tecnológica – que analiza como factor social externo- es cerrar la brecha entre los modos académico e industrial de producción de conocimiento. Según el autor, los dos sistemas –ciencia académica y ciencia industrial- siempre han estado conectados y han sido dependientes uno del otro. La existencia de cada uno de ellos siempre fue entendida como esencial para la vitalidad y continuidad del otro, pero rara vez han coexistido bajo el mismo techo. Más allá de sus convergencias técnicas y de gestión - sostiene Ziman-, la ciencia académica y la industrial difieren significativamente en sus objetivos sociales.

Tal como se planteó en la introducción de este capítulo, la organización de la ciencia académica puede entenderse en términos del CUDOS o las normas mertonianas. Según Ziman (2000), si bien esas normas han sido altamente idealizadas, no son del todo irreales. Sin embargo y desde la perspectiva del autor, la ciencia industrial contraviene todas esas normas. Las prácticas sociales características de la ciencia industrial se edifican sobre principios que niegan la existencia de ese

*ethos*. A saber, la ciencia industrial es *propietaria, local, autoritaria, encargada y experta* (Ziman, 2000).

En el contexto de esta tesis, la *industrialización de la ciencia* se entiende como un caso extremo del rol que potencialmente podría tener el conocimiento, cuando conjuntamente operan diversos factores, entre ellos económicos y políticos. Entre los económicos, el rol que históricamente ha desempeñado el conocimiento científico y tecnológico como instrumento para el crecimiento de las economías industrialmente más avanzadas. Entre los políticos, el papel asignado al conocimiento por la respuesta específica de la política científica y tecnológica implementada en los países altamente industrializados. Es decir, no se trata de un fenómeno universal sino que depende de su contexto social e histórico.

Aunque América Latina toda -y México no es la excepción-, hayan sido importadores destacados de las políticas científicas y tecnológicas de los países industrializados, su ciencia sigue siendo ciencia académica. Sin embargo, la región y México en particular, tienen un desafío mayor por delante, el desarrollo de políticas científicas y tecnológicas propias y a medida, para que la ciencia (incluidas la tecnología y la innovación) efectivamente se convierta en uno de los pilares relevantes para dinamizar su desarrollo social y económico. Siendo además éste último, un objetivo explícito de política pública, plasmado en el Plan Nacional de Desarrollo de México 2013-2018.

Finalmente, Ziman (2000) reduce la conceptualización del *Modo 2* de producción de conocimiento de Gibbons et al. (1994), cuando dice: *“el nuevo «modo 2» de producción de conocimiento es prácticamente idéntico a la manera en que ahora las empresas más actualizadas organizan sus actividades de investigación. Lo que podría llamarse «ciencia postindustrial», [que] difiere del estereotipo anterior de la ciencia industrial al sustituir la «competencia de mercado» por la gestión «de mando».* (Ziman, 2000: 80-81).

En la siguiente sección se describen las formas diferenciales de producción de conocimiento, es decir, las conceptualizaciones *Modo 1* y *Modo 2* de Gibbons et al. (1994), y se fundamenta por qué es una reducción equiparar el concepto *Modo 2* con *ciencia industrial* o *posindustrial*.

## II.2 Las características diferenciales del *Modo 1* y *Modo 2* de producción de conocimiento

Gibbons et al. (1994) en su libro *The new production of knowledge: the dynamics of science and research in contemporary societies* caracterizan un nuevo modo de producción de conocimiento, que emerge junto con la forma tradicional de producción de conocimiento. De acuerdo con los autores, este nuevo modo de producción de conocimiento afecta las características sociales de la producción de conocimiento, su gobernanza. Es decir, qué conocimiento se produce, cómo se produce, en qué contextos, cómo se organiza, sus sistemas de recompensa y los mecanismos para controlar la calidad de lo que se produce. Históricamente, la gobernanza de la producción de conocimiento ha estado articulada con las ciencias disciplinares, la física, la química, la biología, etc.; y ha sido considerada como forma paradigmática para producir conocimientos robustos en las ciencias, por lo que ha sido imitada por otras áreas de conocimiento, entre ellas las ciencias sociales y las humanidades (Gibbons et al., 1994).

Estos autores, para describir los cambios que observan en la producción de conocimiento, caracterizan dos tipos diferenciales o *modos* de producción de conocimiento. Al nuevo modo de

producción lo denominan *Modo 2*, para distinguirlo de la forma tradicional de producción de conocimiento que denominan como *Modo 1*. Desde su perspectiva, el Modo 2 interactúa con el Modo 1 de producción de conocimiento, pero no puede reemplazarlo. Por un lado, quienes producen conocimiento en Modo 1 pueden hacerlo también en Modo 2 y *viceversa*; y por otro lado, algunos productos derivados de la producción Modo 2 de conocimiento, pueden servir como dinamizadores de la producción Modo 1 de conocimiento (por ejemplo: el desarrollo de *instrumentos*). En el Modo 2 el conocimiento se produce en el contexto de aplicación y por lo tanto, a diferencia del Modo 1, fuera de las estructuras disciplinares. Esto no quiere decir, que la producción de conocimiento Modo 2 no genere conocimiento teórico fundamental, lo genera sí, pero fuera de las estructuras disciplinares. Para ejemplificar el desarrollo teórico fundamental en el contexto de aplicación y que ese conocimiento contribuye al avance intelectual fuera de las estructuras disciplinarias, los autores señalan a las ciencias de la computación, los materiales, las ciencias biomédicas y ciencias medioambientales. (Gibbons et al., 1994)

A la vez, estos autores sostienen que la tendencia ha sido no reconocer el Modo 2 como una nueva forma emergente de producción de conocimiento, sino que ha sido reducida a términos comúnmente utilizados, tales como: *investigación estratégica, investigación aplicada, investigación pre-competitiva, mission-oriented research*, etc. Para los autores esto se debe, por un lado, a las interacciones entre esos modos de producción de conocimiento, y, por otro lado, al predominio de la idea de que la *ciencia disciplinaria* es la fuente de las aplicaciones futuras. Además de esto, esos autores consideran que la persistencia del modelo lineal de innovación en los debates políticos, se puede explicar por la profunda creencia de que el conocimiento teórico fundamental no puede ser producido y sostenido fuera de las estructuras disciplinares. (Gibbons et al., 1994)

Quizás sea por esta profunda creencia, que en México, en la actualidad, y más allá de la retórica plasmada en los planes de CTI, todavía persiste el enfoque lineal de la innovación en el diseño e implementación de dichos planes de política pública (ver capítulos VI -sección VI.1- y VII -sección VII.3-).

Según Gibbons et al. (1994), el concepto *Modo* de producción de conocimiento intenta resumir las *normas cognitivas y sociales* -ideas, métodos, valores y normas- que debe seguir dicha producción de conocimiento, su legitimación y difusión. Para el Modo 1, estas normas sociales y cognitivas, provienen de las ciencias disciplinares –en particular de la física, del *modelo newtoniano*- pero se difundieron a diversos campos científicos como dominantes para determinar lo que se considera una *práctica científica sólida*. Es decir, las normas cognitivas y sociales del Modo 1 de producción de conocimiento determinan cuáles son los problemas significativos, quiénes pueden hacer ciencia y qué es buena ciencia. Según los autores, el Modo 1 es generalmente asociado con la *ciencia*, y todo lo que se aparte de las normas que lo rigen, no merece ser llamado como ciencia. Sin embargo, y como señalan los autores, si bien el Modo 2 supone una nueva forma de producción, legitimación y difusión del conocimiento, eso no significa que quienes trabajan en este modo de producción se aparten del método científico. (Gibbons et al., 1994)

En el Modo 2 –a diferencia del Modo 1- el conocimiento se genera en contextos económicos y sociales más amplios, en el *contexto de aplicación*. En ese contexto y a diferencia del Modo 1, los problemas no se definen en un marco estructurado disciplinariamente. Por el contrario, dicho marco es *transdisciplinar* más que *monodisciplinar*. A la vez, esa generación de conocimiento se desarrolla



en formas no jerárquicas, *organizadas heterogéneamente*, que son esencialmente transitorias. Es decir, el Modo 2 de producción de conocimiento a diferencia del Modo 1, no está principalmente institucionalizado dentro del ámbito de las estructuras académicas, e implica que el proceso de producción de conocimiento depende de la *interacción estrecha con otros actores* externos. Para los autores esto significa que la producción de conocimiento en el Modo 2 es más *responsable socialmente*. Como consecuencia de estos cambios, el Modo 2 recurre a un *conjunto más amplio de criterios para juzgar la calidad de la producción de conocimiento*. Así y en términos generales, el proceso de producción de conocimiento en el Modo 2 es más *reflexivo* y afecta profundamente lo que se considera como *buena ciencia*. (Gibbons et al., 1994)

En resumen, mientras que en el Modo 1 de producción de conocimiento los problemas son definidos y resueltos en el *contexto académico* según los intereses disciplinares de una comunidad específica –la académica-, en el Modo 2 los problemas y sus soluciones tienen lugar en el *contexto de aplicación* según los intereses de diversos actores. Así, mientras que el Modo 1 es *disciplinario, homogéneo, jerárquico y estable*; el Modo 2 es más *transdisciplinario, heterogéneo, heterárquico y transitorio*. A la vez, cada uno de estos modos de producción de conocimiento requiere de formas diferenciales de control de calidad. Finalmente, en términos comparados, el Modo 2 es más responsable y reflexivo socialmente, involucra un conjunto más amplio, temporal y heterogéneo de actores que colaboran en la búsqueda de soluciones a problemas específicos localizados y contextualizados. (Gibbons et al., 1994)

A continuación, se describen detalladamente las características distintivas de estos modos diferenciales de producción de conocimiento.

### II.2.1 Producción de conocimiento en el Contexto de Aplicación

Siguiendo a Gibbons et al. (1994), la primera característica distintiva entre el Modo 1 y el Modo 2 tiene que ver con el *contexto* que guía la resolución de problemas. Mientras que en el Modo 1 la orientación en la resolución de un problema está dada por lo que es considerado relevante conceptualmente en una disciplina particular, en el Modo 2 la resolución de problemas está orientada en torno a una aplicación específica para dar solución a un problema dependiente del contexto.

Así, en el primer caso el *contexto* se define de acuerdo con las normas cognitivas y sociales que rigen a la investigación básica –fundamental- o *ciencia académica*. En el segundo caso, la producción de conocimiento se da en el *contexto de aplicación*. Es decir, es resultado de un conjunto más amplio de intereses y busca, desde el inicio, ser útil para algún actor (del ámbito del gobierno, o de la industria, o de la sociedad en general). En este caso, la producción de conocimiento responde a procesos de negociación continuos y dicha producción no sucede hasta que no se contemplan los intereses de los distintos actores involucrados. (Gibbons et al., 1994)

A diferencia de Ziman (2000), para Gibbons et al. (1994) la aplicación no es exclusivamente el desarrollo de productos para la industria, y los determinantes de qué conocimiento se produce, no se restringe únicamente a los mercados. Por el contrario, dichos determinantes implican procesos mucho más amplios y difundidos. Si bien es cierto que en la producción Modo 2 de conocimiento, pueden operar factores de oferta y de demanda de mercado, sus fuentes de dinamización son cada vez más diversas. La producción de conocimiento en el Modo 2 tiene en consideración elementos que van más allá del valor económico y/o comercial del conocimiento generado, y su difusión a la

sociedad es a través de procesos que exceden a los mercados. Es por esto que los autores hablan de *conocimiento socialmente distribuido*. De aquí que sea una reducción equiparar el concepto de *Modo 2* con el de *ciencia industrial o posindustrial*, como lo hace Ziman (2000), y que Hessels y van Lente (2008) reproducen.

Entre las críticas que Hessels y van Lente (2008) le hacen al Modo 2 de Gibbons et al. (1994), sostienen que ese modo de producción de conocimiento en realidad no es una característica generalizada sino que solamente sirve para caracterizar algunas ciencias más aplicadas. Sin embargo, de acuerdo con Gibbons et al. (1994), ciertas disciplinas como por ejemplo, la ingeniería química, la ingeniería aeronáutica o la informática, consideradas generalmente como *ciencias aplicadas*, en realidad deben ser entendidas como formas genuinamente nuevas de conocimiento porque la fuente inicial para su nacimiento fue, justamente, la falta de ciencia relevante. Sin embargo, estos autores no las llaman como formas de producción de conocimiento Modo 2 porque rápidamente se convirtieron en espacios de producción de conocimiento disciplinar –Modo 1-. Los autores sostienen que si bien es cierto que comparten algunas características del Modo 2, el conocimiento que producen no ocurre estrictamente en el contexto de aplicación. En el Modo 2 el contexto es más complejo que el de las *ciencias aplicadas* y reúne demandas intelectuales y sociales diversas (Gibbons et al., 1994)

### II.2.2 Producción de conocimiento Transdisciplinaria

De acuerdo con Gibbons et al. (1994), el Modo 2 va más allá del involucramiento de distintos especialistas que trabajan en equipo para la búsqueda de soluciones y su aplicación, orientada a atender los problemas de un entorno complejo. El Modo 2 de producción de conocimiento también requiere del consenso sobre las normas cognitivas y sociales que orientan la acción. Según los autores, este consenso está condicionado por el contexto de aplicación y co-evoluciona con él. Así, los determinantes que dan lugar a una solución potencial a un problema dependiente del contexto, dependen de la integración de diversas habilidades en un marco de acción transitorio, cuya duración estará sujeta a la capacidad que tenga de adaptarse a las exigencias que impone el contexto específico de aplicación.

La transdisciplinaria tiene cuatro características sobresalientes que, siguiendo a Gibbons et al (1994), se detallan a continuación.

La producción de conocimiento transdisciplinar se desarrolla en un marco –distinto del disciplinar-, que evoluciona para orientar la búsqueda de soluciones a problemas complejos. Su ámbito es el contexto de aplicación, pero no es que primero se crea conocimiento y luego se aplica. Es decir, la solución no es un resultado lineal de la aplicación del conocimiento existente. Por el contrario, y si bien el conocimiento existente puede desempeñar un rol importante, el desarrollo de una solución implica una creación de conocimiento genuinamente nueva que, cuando se alcanza, no puede ser reducida en partes disciplinares.

La segunda característica de la transdisciplinaria es que contribuye al conocimiento, aunque no necesariamente al conocimiento disciplinar. La búsqueda de soluciones a problemas complejos involucra elementos empíricos y teóricos, y si bien el conocimiento se crea en un contexto específico de aplicación, desarrolla sus estructuras teóricas para la comprensión, tiene métodos propios y es acumulativo. A la vez, las direcciones que asuma esa acumulación dependerán de las búsquedas de soluciones futuras.

La tercera característica se vincula con las formas en que se comunican y difunden los resultados de investigación. En el Modo 1 los resultados de investigación se comunican y difunden a través de ciertos canales formales, tales como su publicación en revistas científicas especializadas, su presentación en conferencias, etc. En el Modo 2, a diferencia de eso, los resultados son comunicados a los involucrados en la búsqueda de soluciones a problemas dependientes del contexto. Así, la difusión de los resultados de investigación, comienza inicialmente en el propio proceso de producción de conocimiento. Posteriormente, esa difusión continúa, principalmente, a través de la movilidad de los involucrados a nuevos contextos de aplicación. De ese modo se conforman redes de comunicación que persisten, y el conocimiento allí contenido –las soluciones encontradas-, están disponibles como piezas para atender nuevos problemas en configuraciones contextuales diferentes.

Finalmente y vinculado con lo anterior, la cuarta característica es que la transdisciplinariedad es dinámica. Representa la capacidad de resolver problemas complejos cambiantes. Así una solución específica puede ser el punto de partida para nuevos avances, nuevas soluciones o aplicaciones difíciles de predecir *ex ante*. Al igual que en el Modo 1, los avances se pueden basar en el conocimiento previo. La diferencia del Modo 2, es que dichos avances caen afuera de las fronteras del conocimiento de cualquier disciplina particular. Por lo mismo, puede suceder que los nuevos conocimientos producidos en Modo 2, no encajen fácilmente como contribuciones significativas al avance del conocimiento de las disciplinas particulares que, contribuyeron en la exploración de soluciones. De ahí la dificultad de registrar estos avances en los canales formales de comunicación, ya que -en general- éstos se centran en la difusión de contribuciones intelectuales disciplinares. En el Modo 2, la difusión de conocimiento se puede dar a través de canales formales pero principalmente, se da a través de canales informales como los descritos en el párrafo anterior.

### II.2.3 Producción de conocimiento Heterogénea

Tal como se señaló párrafos más arriba, los determinantes que dan lugar a una solución a un problema complejo dependiente del contexto, dependen de la integración de diversas habilidades en un marco de acción transitorio, cuya duración estará sujeta a la capacidad que tenga de adaptarse a las exigencias que impone el contexto específico de aplicación.

En este sentido, Gibbons et al. (1994) destacan que la producción Modo 2 de conocimiento es heterogénea por las habilidades y experiencias que los involucrados aportan a la búsqueda de soluciones. A la vez, ese marco de acción es transitorio, porque la composición del grupo que trabaja en la búsqueda de soluciones cambia conforme avanza en la resolución de los problemas, y con ello también, cambia el contexto de aplicación. A medida que se avanza en la resolución de problemas, y al igual que en el Modo 1, emergen nuevos problemas en contextos distintos y en formas igualmente inesperadas.

Así, los espacios potenciales de creación de conocimiento aumentan –se diversifican- porque los problemas, objetos de soluciones del Modo 2 no emergen exclusivamente de las universidades, sino de los contextos de aplicación. Por lo tanto, en el Modo 2 los espacios para la producción de conocimiento son múltiples -centros sectoriales de investigación, dependencias de gobierno, institutos no universitarios, laboratorios industriales, grupos de reflexión, etc-, cuyos actores interactúan entre sí para dar lugar a soluciones específicas. Las formas de interacción también son

variadas, pueden ser formales o informales y ocurren a través de redes. A diferencia del Modo 1, en el Modo 2, actores heterogéneos participan en la producción de conocimiento.

#### II.2.4 Producción de conocimiento Socialmente Responsable y Reflexivo

Según Gibbons et al. (1994), la producción Modo 2 de conocimiento se ha visto estimulada por el creciente interés público respecto a las consecuencias del avance científico y tecnológico sobre el medioambiente, la salud, las comunicaciones, la privacidad, etc. Dicho de otro modo, la conciencia de que la ciencia y la tecnología no son neutrales, ha conducido a que cada vez sean más los grupos interesados en influir en los procesos de investigación.

Debido a la complejidad de los problemas, se configuran equipos de trabajo heterogéneos, integrados por científicos sociales que trabajan junto con científicos naturales, ingenieros, abogados, empresarios, etc. Así, la responsabilidad social permea todo el proceso de producción de conocimiento, desde el establecimiento de las prioridades de investigación, pasando por la definición del problema, hasta la interpretación y difusión de los resultados obtenidos. Esto, según los autores, es trabajar en el contexto de aplicación. (Gibbons et al., 1994)

Trabajar en el Modo 2 hace que sus participantes sean más reflexivos, básicamente porque los problemas que se abordan en este nuevo modo de producción de conocimiento no pueden ser entendidos sólo en términos científicos y técnicos, las soluciones requieren también de opciones de implementación. De ahí que los procesos de investigación necesariamente también tienen que atender los valores y preferencias de los interesados en las soluciones. Es decir, de los individuos y grupos que tradicionalmente fueron considerados externos al sistema científico y tecnológico, pero que ahora se constituyen en agentes activos tanto en la definición y solución de problemas, como en la evaluación de su desempeño (Gibbons et al., 1994). Finalmente y siguiendo a estos autores, conforme aumenta la reflexividad de los procesos de investigación, también se expande la demanda del conocimiento que generan las ciencias sociales y las humanidades.

#### II.2.5 Producción de conocimiento y el control de su calidad

De acuerdo con Gibbons et al. (1994), los criterios para evaluar la calidad del trabajo derivado de la producción Modo 2 de conocimiento son distintos de los utilizados en el ámbito de la ciencia disciplinar más tradicional –Modo 1 o *ciencia académica*–.

Siguiendo a estos autores, en el Modo 1, la *calidad* es valorada y determinada por los juicios realizados por los pares en la revisión de las contribuciones intelectuales de los individuos. Así, el *control* es mantenido por la selección de los individuos que han sido juzgados como competentes para desempeñarse como iguales. De este modo, calidad y control se refuerzan mutuamente. A la vez, la revisión por pares está normada por dimensiones cognitivas y sociales, porque evalúa la relevancia conceptual de los problemas abordados y la adecuación de las técnicas utilizadas, al tiempo que selecciona a quienes están calificados para atenderlos. En otras palabras, la revisión por pares orienta a los individuos a trabajar en los problemas considerados esenciales para el avance del conocimiento en la disciplina, y que -en buena medida- reflejan los intereses intelectuales y las preocupaciones de la disciplina y de sus *guardianes* –se podría decir: de sus *élites*–. (Gibbons et al., 1994)

Sin embargo, y de acuerdo con esos autores, determinar la *buena ciencia* en el Modo 2 es más difícil y para hacerlo, son necesarios criterios de evaluación adicionales. Estos criterios son incorporados

a través del contexto de aplicación. En dicho contexto, los intereses intelectuales son más amplios e incluyen además preocupaciones sociales, económicas y políticas. Esto no quiere decir que en el Modo 2 la calidad sea menor, significa que la calidad está normada por un conjunto más amplio de dimensiones sociales y cognitivas, con criterios adicionales y que reflejan la ampliación de la composición social del sistema de revisión (Gibbons et al., 1994). Es decir, incluye también a los *impares* (Avalos, 1997); a éstos, con un conjunto de criterios distintos y adicionales a los criterios académicos, les corresponde juzgar las implicancias del uso del conocimiento producido y su eventual impacto social y/o económico (Bianco, Gras y Sutz, 2014). En la siguiente sección se presentan –brevemente- algunos enfoques sobre la evaluación de la investigación en el *contexto de aplicación*.

Así y siguiendo a Gibbons et al. (1994), mientras que en el Modo 1 el avance del conocimiento depende esencialmente de la creatividad individual y del control de calidad orientado por los intereses disciplinares y que se realiza a través de la revisión por pares; en el Modo 2, el avance del conocimiento depende de la creatividad colectiva y del control de calidad orientado por intereses múltiples, y que se realiza como parte de un proceso social ampliado, que incluye a los *pares* y a los *impares* en un contexto de aplicación específico.

De esa manera, en el Modo 1 el conocimiento avanza y se acumula a través de la profesionalización especializada de la actividad de investigación y un tipo de evaluación –control de calidad- altamente institucionalizado en el contexto de la ciencia académica –disciplinar-. En el Modo 2, por el contrario, el conocimiento avanza y se acumula a través de la configuración de equipos de trabajo transitorios que se movilizan –total o parcialmente, de acuerdo con los requerimientos de los diversos contextos de aplicación-, y por lo mismo, requiere de formas flexibles, adaptativas y contextuales de control de calidad. (Gibbons et al., 1994)

Históricamente, y de acuerdo con Gibbons et al. (1994), el Modo 1 ha sido reconocido socialmente como la forma exitosa de producir conocimiento. Los investigadores saben que la manera más eficaz para alcanzar el éxito es respetando la especialización disciplinar, profesionalizándose en su actividad y cumpliendo con los controles de calidad altamente institucionalizados. Sin embargo, cuando el conocimiento se produce en Modo 2, generalmente se considera –sostienen Gibbons et al. (1994)- que el conocimiento generado no puede ser calificado como científico porque fue generado fuera de las estructuras de legitimación de la ciencia académica disciplinar. En respuesta a esa consideración generalizada, Gibbons et al. (1994) afirman que el conocimiento generado en Modo 2 sí es científico, el problema está en la falta de adaptación de las normas que históricamente han regido a la producción de conocimiento. Para estos autores, dichas normas tienen que adaptarse a esta nueva forma de producción de conocimiento, porque de lo contrario, se pone en juego el propio desarrollo de la ciencia.

### II.3 La evaluación de la investigación en el *contexto de aplicación*: algunos enfoques

Diversos autores -entre ellos de Jong et al. (2011); Regeer et al. (2009); Spaapen, Dijstelbloem y Wamelink (2007); Hemlin y Barlebo Rasmussen (2006)- coinciden con Gibbons et al. (1994) en que cuando el conocimiento se realiza en el contexto de aplicación, requiere de enfoques de evaluación diferenciales.

Hemlin y Barlebo Rasmussen (2006) sostienen que los cambios en las formas de producción de conocimiento exigen nuevos criterios de evaluación de la investigación. Desde la perspectiva de esos

autores, la ciencia es cada vez más reflexiva, tanto en términos de su vinculación con lo social como respecto a la utilidad de los productos y/o resultados que genera. Para estos autores, en coincidencia con Gibbons et al. (1994), esa reflexividad tiene consecuencias importantes para el control de calidad de la producción de conocimiento. Señalan que nuevas formas organizacionales están emergiendo en el sistema de producción de conocimiento, que están apareciendo nuevos criterios de calidad y que cada vez son más los actores involucrados e interesados en los resultados de investigación.

En particular, argumentan que se está produciendo una transformación en la evaluación de la investigación. Destacan que se está transitando de un sistema basado en el control de calidad – *quality control* (QC)- a un sistema de evaluación de seguimiento – *quality monitoring* (QM)-. La idea central es que en el QM la evaluación de la investigación -a diferencia de lo que ocurre en el QC- se centra en la calidad del proceso de investigación, y no solamente en sus resultados o productos. El QM se caracteriza por utilizar nuevos criterios, tener otros objetivos y focos, ser desarrollada en cooperación con actores externos al ámbito de la ciencia académica y ser realizada en distintos momentos a lo largo del proceso de investigación. Así, las diferencias esenciales entre el QC (orientada por resultados) y el QM (orientada por procesos) pueden sintetizarse a través de cinco dimensiones de la evaluación: criterio, foco, objetivo, evaluadores y tiempos de evaluación. (Hemlin y Barlebo Rasmussen, 2006).

El objetivo de la evaluación en el QM no es solamente garantizar la confiabilidad y validez de la producción de conocimiento como en el QC, sino que además tiene que garantizar que dicha producción sea socialmente robusta y que aporte insumos relevantes para el aprendizaje organizacional. Mientras que en el QC el criterio que rige la evaluación es exclusivamente la excelencia científica, en el QM -además de la excelencia científica- importa la relevancia social de la producción de conocimiento. Es decir, el criterio de legitimación del conocimiento no responde únicamente al criterio científico interno sino que también depende de factores externos, tales como: la contribución de ese conocimiento producido al desarrollo productivo, social, o al desarrollo de políticas públicas, etc. Adicionalmente, el foco o unidad de evaluación también es diferente. La evaluación ya no se focaliza solo en los investigadores individuales (homogéneo) como en el QC, sino que en el QM importan las redes y organizaciones (heterogéneo) como agregados relevantes para la producción de conocimiento. Además, mientras que en el QC el evaluador es un par –revisión por pares-, en el QM ingresan nuevos actores en el proceso de evaluación –usuarios, consultores, partes interesadas, etc.-. El QM requiere de la participación de diversos actores para juzgar la calidad, relevancia, factibilidad y pertinencia del conocimiento, debido a que los científicos no cuentan con todos los elementos necesarios para evaluar todos los aspectos sociales y organizativos del nuevo conocimiento generado. Finalmente, desde el punto de vista de los tiempos de la evaluación, el QC es de tipo *ex-post* y valora los resultados de la investigación, en el QM la evaluación es durante el proceso de producción de conocimiento y de tipo continua. Es decir, refiere al seguimiento y mejora continua durante el proceso de producción de conocimiento, más que al control de su productividad. (Hemlin y Barlebo Rasmussen, 2006).

De alguna manera similarmente Spaapen, Dijstelbloem y Wamelink (2007) proponen un enfoque y desarrollan un método comprensivo para evaluar la investigación en su contexto. Se trata de la evaluación de la calidad de la investigación respecto a su relevancia para la ciencia y pertinencia o “*valor para la sociedad*”, su unidad de análisis –de evaluación o foco- es el grupo de investigación.

El enfoque que proponen Spaapen, Dijstelbloem y Wamelink (2007) se basa en estudios previos realizados entre 1998 y 2002 en dos campos, ciencias agrícolas y ciencias farmacéuticas; y representa la respuesta de una red independiente de investigadores y hacedores de políticas científicas de los Países Bajos a la demanda realizada por el *Dutch Consultative Committee of Sector Councils for research and development (COS)* –Comité Consultivo Holandés de los Consejos Sectoriales para la investigación y desarrollo-.

El punto de partida de este enfoque, es el reconocimiento de que el éxito de ciertos programas de investigación depende de la forma en que los investigadores logran conectarse con los temas que le demanda su entorno, y del modo en que ese entorno, absorbe o utiliza los resultados de la investigación. Entre los temas de investigación que demanda el entorno, los autores señalan el calentamiento global, la migración, la identidad cultural, enfermedades como el SIDA, entre otros. A la vez destacan que la producción de conocimiento en torno a esos temas requiere de abordajes *MIT (multi/inter/transdisciplinarios)*. En general, dicha producción se realiza en contextos en los que participan expertos de diversos orígenes, con conocimientos y experiencias también distintas, y en los que se deben negociar múltiples intereses y demandas. De acuerdo con los autores, esa complejidad requiere de otros enfoques para evaluar la calidad de la investigación que allí se realiza. En particular y en coincidencia con Hemlin y Barlebo Rasmussen (2006), resaltan la necesidad de enfoques de evaluación que vayan más allá de la revisión por pares que enfatiza la excelencia científica según el desempeño de algunos indicadores, como por ejemplo: la cantidad de publicaciones en revistas de alto impacto. (Spaapen, Dijstelbloem y Wamelink, 2007)

Spaapen, Dijstelbloem y Wamelink (2007), teniendo presente las formas diferenciales de producción de conocimiento, distinguen diversos dominios sociales en los que operan los investigadores. La variedad de interacciones que ocurren entre los investigadores y los diversos actores que conforman dichos dominios –*actores o partes interesadas o audiencias*-, son la base para la construcción de su método de evaluación. Siguiendo a esos autores, los principales dominios o ámbitos para la producción de conocimiento son: la comunidad científica internacional, la industria, la política, el sector público y la sociedad en general (usuarios finales). Para cada uno de esos ámbitos, los autores desarrollan diversos criterios de evaluación e indicadores sobre el desempeño de los grupos de investigación.

Proponen un enfoque con una metodología que involucra cuatro fases como elementos básicos, para alcanzar un procedimiento de evaluación que reconozca la posición contextualizada de la investigación. Asimismo destacan el carácter reflexivo de su metodología, por lo que su ejecución y operacionalización variarán dependiendo de las circunstancias locales específicas, demandas y limitaciones. Muy sintéticamente, las fases que proponen son: i) *Misión*: para establecer el propósito del grupo de investigación; ii) *REPP (Research Embedment and Performance Profile)*: para aproximar una medida cuantitativa y cualitativa de la producción de conocimiento del grupo y su interacción con el entorno; iii) *Partes interesadas*: para ponderar el impacto del trabajo del grupo de investigación; en esta fase se consulta a los actores involucrados o interesados en los usos de la producción de conocimiento sobre el impacto de los resultados de investigación; y, iv) *Retroalimentación*: para re-organizar la estrategia del grupo de investigación. Los resultados obtenidos de las fases dos y tres se confrontan con los de la fase uno. En última instancia con este enfoque de evaluación, los autores buscan una mejor representación y valoración de la variedad de

contribuciones que puede realizar la investigación para el desarrollo de la sociedad. (Spaapen, Dijstelbloem y Wamelink, 2007).

En base a este enfoque de cuatro fases, las principales organizaciones de política científica de los Países Bajos iniciaron, en 2006, un proyecto conocido como “*Evaluating Research in Context –ERIC*” que explora la implementación de esas nuevas formas de evaluación en los diversos sistemas nacionales –e incluso internacionales- de evaluación.<sup>3</sup>

Posteriormente, en 2009, la Comisión Europea continuó estos esfuerzos, a través de su séptimo programa marco (FP7), financiando el proyecto titulado “*Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society –SIAMPI-*”<sup>4</sup>.

Regeer et al. (2009), en su trabajo titulado “*Six Guiding Principles for Evaluating Mode-2 Strategies for Sustainable Development*” proponen un marco para la evaluación de programas públicos de investigación orientados por la búsqueda de soluciones a problemas sociales complejos, en particular los vinculados con el desarrollo sustentable. Específicamente, proponen una extensión del marco de evaluación que se basa en el enfoque *ILA*<sup>5</sup> -*Interacting, Learning and Action-* (Bunders, 1990; Bunders y Broerse, 1991; citados en Regeer et al., 2009).

Regeer et al. (2009), en sintonía con Gibbons et al. (1994), sostienen que la búsqueda de soluciones a problemas sociales complejos se caracteriza por diseños de investigación emergentes (*Modo 2*), con procesos de aprendizajes interactivos entre diversos actores y gestión adaptativa. Destacan que esas características, su complejidad y la intervención sistémica que suponen los programas de investigación para el desarrollo sustentable desafían a la evaluación. En ese sentido, el marco de evaluación que proponen ofrece seis principios guía para asistir a los evaluadores en la evaluación de esos programas complejos. (Regeer et al., 2009)

Los principios guía propuestos por los autores se centran en: i) el *foco de la evaluación*: la evaluación del Modo 2 focaliza en el funcionamiento interno de los intermediarios, específicamente a través del apoyo para el aprendizaje continuo sobre estrategias de intervención para el desarrollo sustentable; ii) el *rol de la evaluación*: el seguimiento y la evaluación están íntimamente vinculados con la intervención porque forman parte de un proceso interactivo orientado a la definición, implementación y eventuales ajustes a dichas intervenciones; iii) el *proceso de creación de un marco para la evaluación*: enfatiza el uso de un enfoque participativo para el desarrollo de la evaluación Modo 2 y sus indicadores; iv) la *naturaleza de los indicadores*: los indicadores no evalúan sino que sensibilizan, son dependientes del contexto y dinámicos; y los últimos dos principios guía, se sintetizan en la *contribución de los indicadores al aprendizaje*: los indicadores proveen una medida de las discrepancias entre las intenciones (orientadas por el Modo 2) del programa de investigación sobre desarrollo sustentable y la práctica de investigación, hacen visibles las condiciones contextuales (por ejemplo: las sociales, políticas, económicas e institucionales) que limitan al

---

<sup>3</sup> Ver Spaapen, Dijstelbloem y Wamelink (2007) para una descripción detallada.

<sup>4</sup> Ver <http://www.siampi.eu/19/546.bGFuZz1FTkc.html>, fecha de consulta 07/09/17.

<sup>5</sup> El objetivo del enfoque *ILA* es apoyar procesos de innovación inclusivos en los campos de la agricultura, la salud y la biotecnología (Regeer et al., 2009).



programa de investigación, y en ese sentido, *iluminan* la búsqueda de estrategias para superar tanto las discrepancias como las limitaciones del programa. (Regeer et al., 2009).

El trabajo realizado por de Jong et al. (2011), retomando la iniciativa *ERiC* y dentro del proyecto *SIAMPI*, ponen a prueba un enfoque para la *evaluación de la investigación en su contexto*. Desde el marco conceptual de las *interacciones productivas* y con base en dos estudios de caso – uno en el campo de la arquitectura y el otro en derecho-, analizan la conformación de agendas de investigación, las colaboraciones para el desarrollo de la investigación, la difusión del conocimiento que se genera y su uso (impacto).

El enfoque para la *evaluación de la investigación en su contexto*, está muy alineado con el de Spaapen, Dijstelbloem y Wamelink (2007), que ponen a prueba, y consta de cuatro etapas de análisis: i) la naturaleza del campo del conocimiento; ii) la misión del grupo de investigación; iii) las audiencias relevantes; y iv) los productos relevantes, tanto en términos de la cantidad y tipos de productos de investigación como en términos de su calidad e impacto. (de Jong et al., 2011)

Sus resultados sugieren que la evaluación de la investigación, debe considerar otros factores más allá de los criterios propios del ámbito de la ciencia académica. En particular, los autores sostienen que dicha evaluación tiene que considerar los múltiples objetivos de investigación, los diversos productos que genera, las características propias del contexto en el que se desarrolla la producción de conocimiento, sus usos y las *audiencias* para las cuáles ese conocimiento es producido. Finalmente, concluyen que la producción científica tiene muchas *audiencias* (sociedad civil, políticos, empresarios, otros científicos, etc) y que las formas en que se produce ese conocimiento, su difusión e impacto, no siempre son apreciables a través de la cantidad de *artículos científicos* publicados en revistas de alto impacto. En ese sentido, sostienen que la evaluación de la investigación debería considerar también toda esa diversidad como resultados de investigación, impacto y desempeño. (de Jong et al., 2011)

#### II.4 Conclusión: la racionalidad detrás de la selección de los Modo 1 y Modo 2 como dos tipos ideales de producción de conocimiento

Hasta este punto se han desarrollado y establecido ciertos conceptos teóricos -y algunas de sus relaciones- que son claves para los objetivos de este trabajo. Estos conceptos provienen de la teoría, fundamentalmente de la sociología de la ciencia. En el contexto de esta tesis, algunos de estos conceptos se toman como estructuras conceptuales referenciales, para alcanzar una caracterización analítica y facilitar cierta aproximación empírica sobre algunas formas diferenciales de producción de conocimiento, observables en los investigadores categorizados en el SNI en México. Como primer paso para la construcción del marco teórico conceptual de esta tesis (capítulo V), a continuación se discuten estos conceptos y se argumenta las razones de la selección de algunos de ellos y sus relaciones con la evaluación.

Algunos de los conceptos descriptos a lo largo de este capítulo, tales como: *ciencia académica*, *ciencia posacadémica*, *ciencia industrial*, *ciencia posindustrial*, *Modo 1* y *Modo 2*, podrían – eventualmente- ser retomados como *tipos ideales* (Weber, 1992 [1922]) de producción de conocimiento. Sin embargo, en el contexto de esta tesis, se retienen únicamente los conceptos *Modo 1* y *Modo 2* como *tipos ideales* para comprender los rasgos sobresalientes de los ámbitos y motivaciones que orientan la acción de los investigadores del SNI en la producción de conocimiento.

La elección por retener estos conceptos como tipos ideales de producción de conocimiento se justifica en dos razones, la primera sustantiva y la segunda práctica.

La razón sustantiva está en el alcance conceptual explicativo de los Modo 1 y Modo 2 de producción de conocimiento y la *conexión de sentido* (Weber, 1992 [1922]) que hacen para los propósitos de este trabajo.

Mientras que el concepto *Modo 1* (Gibbons et al., 1994) puede ser equiparado al de *ciencia académica* (Ziman, 2000) porque ambos refieren al mismo fenómeno y representan el punto de referencia a partir del cual, esos autores, construyen sus respectivas conceptualizaciones sobre las transformaciones en la producción de conocimiento; el concepto *Modo 2* (Gibbons et al., 1994) no es equiparable con ninguno de los restantes conceptos -*ciencia industrial, ciencia posindustrial, ciencia posacadémica*- de Ziman (2000).

Como se argumentó, la conceptualización *Modo 2 de producción de conocimiento* refiere a un fenómeno mucho más amplio que a los fenómenos que señala Ziman (2000) con los conceptos de *ciencia industrial* y *ciencia posindustrial*. Por lo mismo, el concepto *Modo 2* tiene un alcance explicativo potencial de la realidad mayor que cualquiera de los otros dos conceptos.

Asimismo, *Modo 2* y *ciencia posacadémica* tampoco pueden ser considerados como conceptos similares. Básicamente porque el primero refiere a una forma específica de organización, legitimación y difusión de la producción de conocimiento orientada a la resolución de problemas dependientes del contexto (productivos, sociales, de política pública, ambientales, en salud, etc.), que si bien comparte algunas características con el segundo, éste último refiere a un fenómeno mucho más general. *Ciencia posacadémica* es la denominación que propuso Ziman (2000) para describir lo que para él es un nuevo sistema de ciencia, que resulta de las transformaciones que observa en la organización, gestión y desarrollo de la ciencia.

La conexión de sentido refiere a que los Modo 1 y Modo 2 de producción de conocimiento se vinculan con la gobernanza de la ciencia. En particular, y por una parte, respecto a qué conocimiento se produce, cómo se produce, en qué contextos y cómo se organiza. Es decir, al hecho de que cada uno de esos modos puede asociarse con distintos contenidos y direcciones de la producción de conocimiento. Por otra parte, porque también cada uno de esos modos de producción requieren de evaluaciones diferenciales para controlar la calidad del conocimiento que se produce en cada uno de ellos (de Jong et al., 2011; Regeer et al., 2009; Spaapen, Dijstelbloem y Wamelink, 2007; Hemlin y Barlebo Rasmussen, 2006; Gibbons et al., 1994). En ese sentido, es esperable que dependiendo de las características específicas del sistema de evaluación, es decir, de las características del sistema de reconocimiento y recompensa de la “buena ciencia”, varíe la dirección y contenido de la producción de conocimiento (Whitley, 2007; Gläser y Laudel, 2007b) o qué ciencia se hace (Hess, 2007) o, en los términos de esta tesis, el balance entre las propensiones Modo 1 y Modo 2 de producción de conocimiento. En el siguiente capítulo se retoma y profundiza sobre estos temas.

La razón práctica que justifica la elección que se hace en esta tesis de retener los conceptos Modo 1 y Modo 2 como tipos ideales de producción de conocimiento, se basa en las posibilidades de operacionalización que ofrecen y su, consecuente, observación empírica. Dichas posibilidades se materializaron, por una parte, a través del enfoque teórico-metodológico-empírico desarrollado en

este trabajo y, por otra parte, gracias al acceso a la encuesta “Percepciones sobre la evaluación académica” diseñada en el marco del estudio realizado por Bensusán et al. (2014).

Esta encuesta contiene información sustantiva sobre los ámbitos y motivaciones que orientan la producción de conocimiento de los investigadores del SNI que permitió, dado el tratamiento metodológico realizado en esta tesis, una aproximación empírica válida y confiable de los conceptos Modo 1 y Modo 2 desarrollados por Gibbons et al. (1994).

Hasta donde se sabe, en la literatura especializada -nacional e internacional- no existe ningún estudio que haya aproximado empíricamente esos conceptos con características similares a como se hace en este trabajo. Además, la revisión de la literatura realizada por Hessels y van Lente (2008), resalta la necesidad de realizar mayores esfuerzos orientados a aumentar la validez empírica de la conceptualización Modo 2 de Gibbons et al. (1994).

En la tabla 1 se presentan los dos tipos ideales de producción de conocimiento -como conceptos inobservables directamente- y sus atributos -como conceptos potencialmente observables-.

*Tabla 1: Atributos diferenciales del Modo 1 y Modo 2: dos tipos ideales de producción de conocimiento*

<b>Conceptos inobservables directamente</b>	<b>Características/conceptos potencialmente observables</b>
<b>Modo 1</b>	Producción de conocimiento en el <i>contexto de la ciencia académica</i> Producción de conocimiento <i>orientado por la disciplina</i> Producción de conocimiento <i>disciplinario</i> Actores involucrados en la producción de conocimiento <i>homogéneos</i> Control de calidad/legitimación depende del <i>juicio de los pares</i> , de la comunidad científica Difusión del conocimiento a través de <i>canales formales</i> (publicaciones, conferencias, etc.)
<b>Modo 2</b>	Producción de conocimiento en el <i>contexto de aplicación</i> Producción de conocimiento <i>orientado por la búsqueda de soluciones a problemas dependientes del contexto, aplicaciones.</i> Producción de conocimiento <i>transdisciplinario</i> Actores involucrados en la producción de conocimiento <i>heterogéneos</i> Control de calidad/legitimación depende del <i>juicio de los pares y los impares, del desempeño de las soluciones, reflexivo</i> Difusión del conocimiento a través de <i>canales informales</i> (redes, soluciones, etc.)

Fuente: Elaboración propia en base a Gibbons et al. (1994)

En el capítulo V –marco teórico conceptual- se retoman estos dos conceptos y se definen las dos dimensiones conceptuales: *propensión modo 1* y *propensión modo 2*, destacando tanto los atributos que retoman de los conceptos Modo 1 y Modo 2 caracterizados por Gibbons et al. (1994), así como las relaciones teóricas inmersas entre ellas y de éstas con las restantes dimensiones conceptuales relevantes para los propósitos de esta tesis. En el capítulo X se operacionalizan las dimensiones conceptuales asociadas a estos dos conceptos y se presentan los resultados sobre su validez de constructo. En el capítulo VIII se explican todos los procedimientos y técnicas implementadas para lograr su validación.

A continuación, en el siguiente capítulo, se revisan las contribuciones de los autores que han vinculado teóricamente a la evaluación de la investigación con la dirección, contenido y características epistemológicas de la producción de conocimiento.

## CAPITULO III: La teoría que vincula la evaluación con la producción de conocimiento

La teoría que vincula a la evaluación con la producción de conocimiento sostiene que la evaluación de la investigación afecta la producción de conocimiento, ya que orienta a los investigadores y sus investigaciones. Más específicamente, es esperable que ciertos sistemas de evaluación impacten sobre la organización de la investigación y la dirección de la producción de conocimiento. Es decir, sobre las características epistemológicas de los procesos de investigación y el contenido de la producción de conocimiento. A la vez, esta teoría sostiene que la intensidad o reforzamiento de tales efectos probables, derivados de la implementación de esos sistemas de evaluación, son dependientes del contexto. En particular, dependen de las características específicas del sistema de CTI, del sistema de evaluación y de los campos científicos, y que, simultáneamente, reflejan los balances en las distribuciones de poder, que se derivan de las relaciones de autoridad -o la gobernanza- que allí tienen lugar.

En la literatura existen diversos antecedentes que dan cuenta de la relación entre la evaluación de la investigación y la producción de conocimiento. Elzinga (1988) observa diversos impactos de la evaluación de la investigación sobre: i) la dirección de la producción de conocimiento; ii) la conformación de las agendas de investigación y la importancia relativa que en ellas ocupan los temas de investigación que son centrales para la corriente principal; y, iii) la adaptación de la investigación según las fuentes de financiamiento y sus esquemas de evaluación. De ese modo, el autor sostiene que la evaluación está vinculada a los resultados de la investigación. (Elzinga, 1988). En sintonía, Hess (2007) destaca que para dar respuesta a la pregunta sobre qué ciencia se hace, hay que empezar por estudiar su sistema de evaluación y recompensa. Estudios muy recientes analizan los impactos sobre la producción de conocimiento que se derivan de la implementación de sistemas de evaluación centrados en medidas bibliométricas. Entre ellos Dahler-Larsen (2014) señala que el uso de indicadores bibliométricos en la evaluación académica o en los sistemas de evaluación de la investigación genera diversos *efectos constitutivos*. Bianco, Gras y Sutz (2014) reflexionan y alertan sobre las posibles *consecuencias no anticipadas* de las políticas de estímulo a la investigación y sus instrumentos sobre la producción de conocimiento. Rijcke et al. (2016) presentan una revisión de la literatura sobre los posibles efectos de los sistemas de evaluación de la investigación que se basan en indicadores cuantitativos, las respuestas estratégicas de los investigadores a ellos y sus implicancias sobre la producción de conocimiento. Bianco, Gras y Sutz (2016) plantean la centralidad que ha ganado un *tipo universal de indicadores* para medir la productividad de la producción de conocimiento y discuten sus efectos en el contenido de la investigación. Müller y Rijcke (2017) estudian los impactos de los indicadores cuantitativos –que buscan evaluar el desempeño de la investigación- sobre las características epistemológicas de los procesos de producción de conocimiento en el campo de las *ciencias de la vida*. El tema específico asociado al uso de indicadores cuantitativos en la evaluación de la investigación y sus efectos sobre la producción de conocimiento se trata en profundidad en el capítulo siguiente. Sin embargo, el objetivo de referenciar estos trabajos aquí es simplemente dar cuenta de que no faltan estudios que reportan evidencia relacionada con las consecuencias probables de la evaluación de la investigación sobre el contenido, dirección y características epistemológicas de la producción de conocimiento.

Este capítulo tiene dos objetivos. El primero es revisar las contribuciones de los autores que han vinculado teóricamente a la evaluación de la investigación con la dirección, contenido y características epistemológicas de la producción de conocimiento. El énfasis está puesto en los conceptos, sus relaciones y los mecanismos subyacentes a ellas, ya que representan el fundamento teórico sobre el que se construye el análisis que se hace en esta tesis. En el contexto de esta tesis se refiere a esas contribuciones genéricamente como la teoría que vincula a la evaluación de la investigación con la producción de conocimiento. En el capítulo siguiente se ofrece una revisión amplia de los hallazgos de investigaciones pasadas que sostienen, con la evidencia que reportan, las premisas de la teoría que vincula a la evaluación de la investigación con la producción de conocimiento. El segundo objetivo de este capítulo es plantear los elementos de juicio que permiten la integración –que se hace en esta tesis- de esta teoría con las contribuciones de Gibbons et al. (1994) sobre las formas diferenciales de producción de conocimiento. Este capítulo representa el segundo paso para avanzar hacia la construcción del marco teórico conceptual que justifica y provee los fundamentos teóricos sobre los que se construye el análisis que se realiza en esta tesis.

Para alcanzar estos objetivos, este capítulo se organiza en cinco secciones. Le sigue a esta introducción, la sección III.1 que describe cinco grandes consecuencias probables sobre la organización y la dirección de la producción de conocimiento, derivadas de la implementación de sistemas de evaluación de la investigación (Whitley, 2007). La sección III.2 plantea que las respuestas estratégicas de los investigadores para adaptarse a las presiones impuestas por el entorno, representan el mecanismo que conecta a la evaluación de la investigación con el contenido y características epistemológicas de la producción de conocimiento (Gläser y Laudel, 2007b). La sección III.3 argumenta la relevancia del sistema público de ciencia, del sistema de evaluación de la investigación y de los diversos campos científicos, ya que dependiendo de las características contextuales específicas que exhiban esas tres grandes dimensiones, también variará el impacto esperado de la evaluación sobre el contenido y la dirección de la producción de conocimiento (Whitley, 2007). La sección III.4 sostiene que los cambios sobre la dirección y el contenido de la producción de conocimiento, que probablemente se derivan de la implementación de sistemas de evaluación de la investigación, dependen en última instancia, de la gobernanza de la ciencia (Gläser, 2012). Es decir, de los balances –o desbalances- en las distribuciones de poder que se derivan de las –desiguales- relaciones de autoridad (Whitley, 2010) que en cada contexto específico tienen lugar. Finalmente, la sección III.5 concluye con una reflexión en torno a los elementos teóricos que permiten integrar las formas diferenciales de producción de conocimiento (Gibbons et al., 1994) con la teoría que vincula la evaluación de la investigación con la producción de conocimiento y se justifica y fundamenta teóricamente el análisis de las relaciones estudiadas en esta tesis. Esta sección finaliza con la figura 1 que representa una síntesis de las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento.

### III.1 Las consecuencias probables sobre la organización de la investigación y dirección de la producción de conocimiento

Whitley (2007) realizó un análisis comparativo de la institucionalización –desde mediados de la década de los ochenta- de diferentes sistemas de evaluación de los resultados de investigación<sup>6</sup> en

---

<sup>6</sup> El autor define un sistema de evaluación de la investigación como un conjunto organizado de procedimientos, implementado regularmente por el Estado o agencias estatales, para evaluar

distintos sistemas académicos de diversos países (Australia, Alemania, Países Bajos, España, Suecia, Japón, Estados Unidos y Reino Unido). Sobre esta base propone un marco conceptual para entender cómo y por qué ocurren ciertos cambios en la gobernanza de las ciencias y cuáles son sus consecuencias para la organización de la investigación y la dirección de la producción de conocimiento. En particular, identifica cinco grandes consecuencias probables sobre la producción de conocimiento, derivadas de la implementación de esos sistemas de evaluación.

De acuerdo con el autor y como consecuencia probable de la implementación de esos sistemas de evaluación, los evaluadores, en base a la revisión por pares de los méritos relativos de los productos de investigación de los investigadores, desarrollarán y aplicarán criterios estándar de calidad y relevancia para los campos, y con ello, centralizarán juicios y objetivos. Conforme pasa el tiempo y ese procedimiento se hace regular, dichos juicios, normas y objetivos se institucionalizan como dominantes en los diversos campos, tanto para los investigadores como para las universidades, centros de investigación y otras organizaciones productoras de conocimiento. Así es esperable que disminuya la incertidumbre sobre cuáles son los temas intelectualmente relevantes y sus objetivos de investigación, cuáles son las mejores estrategias de investigación para alcanzarlos y en definitiva, cuáles son los resultados significativos para el avance del conocimiento. (Whitley, 2007)

A la vez, siguiendo al autor, es esperable que los investigadores sean cada vez más conscientes de la competencia por el reconocimiento de las élites científicas. Por lo que es esperable que, para obtener ese reconocimiento, los investigadores coordinen –adapten– sus proyectos de investigación para contribuir a esos objetivos en sus respectivos campos. En ese sentido, es probable que la investigación en general esté más integrada (organizada y coordinada) en torno a dichos objetivos. (Whitley, 2007)

De ese modo y a medida que pasa el tiempo, la centralización y estandarización de los criterios de evaluación, a lo largo de los diversos campos disciplinares, supone una tendencia a la reducción de la diversidad, tanto de objetivos de investigación como de enfoques. Especialmente de aquellos objetivos y enfoques no conformistas, o que desafían a la corriente principal. A la vez, la evaluación se hace cada vez más importante para los investigadores y para las organizaciones que los emplean. En ese sentido, es esperable que aumenten por un lado, el costo de desarrollar estrategias de investigación que se desvían de la pauta establecida por el sistema de evaluación y, por otro lado, las presiones para demostrar cómo los resultados de investigación contribuyen a los objetivos disciplinarios dominantes. En particular, dichos efectos se agudizan entre los investigadores más jóvenes que buscan iniciar y sostener su carrera académica. (Whitley, 2007)

Simultáneamente, debido al reforzamiento de esos estándares y objetivos disciplinares, es esperable que el desarrollo de nuevos campos se vea inhibido. La creciente competencia por ganar reputación, ya que de ésta depende el acceso a recursos para la investigación, incrementa el riesgo de transitar hacia áreas de investigación nuevas, y/o de adoptar métodos, técnicas y enfoques de otras disciplinas. Así la originalidad y/o novedad de la producción de conocimiento tenderá a centrarse en las ciencias existentes y en sus preocupaciones. (Whitley, 2007)

---

retrospectivamente -o *ex-post*- los méritos de los resultados de la investigación realizada con financiamiento público (Whitley, 2007).

Finalmente, los sistemas de evaluación y sus listas ordenadas –*rankings*– sobre la calidad del desempeño de los investigadores, grupos de investigación, centros, institutos, universidades, etc., probablemente intensifiquen la estratificación de individuos, grupos y organizaciones, y con ello, induzca a una mayor concentración de los recursos para la investigación (Whitley, 2007). A este fenómeno, Merton (1988) lo llamó *Efecto Mateo*<sup>7</sup>.

A la vez, Whitley (2007) señala que esas consecuencias probables -o efectos esperados- sobre la organización de la investigación y la dirección de la producción de conocimiento, derivados de la implementación de los sistemas de evaluación, son más importantes o notorios si se trata de *sistemas de evaluación fuertes*.

Siguiendo al autor, un *sistema de evaluación fuerte* es aquel que cuenta con una evaluación de la calidad de los productos de investigación institucionalizada, cuyos resultados son públicos y que es realizada por las élites científicas en períodos regulares, de acuerdo con reglas y procedimientos altamente formalizados. Es decir, si el sistema de evaluación tiene alta frecuencia –es periódico y sistemático–, sus procedimientos están formalizados, sus criterios de evaluación están relativamente estandarizados y si es transparente -sus procedimientos, criterios y resultados se hacen públicos–, entonces se trata de un *sistema de evaluación fuerte*. A la vez y en términos generales, están organizados en torno a las disciplinas existentes, sus resultados afectan la asignación de recursos para la investigación y tienen impactos sobre la gestión de las universidades u otras organizaciones productoras de conocimiento. (Whitley, 2007)

Por lo dicho hasta aquí, se puede sostener que los sistemas de evaluación fuertes del desempeño de la investigación, normados y estructurados en torno a los intereses disciplinarios de las élites científicas dominantes, pueden poner en riesgo el avance del conocimiento y progreso de la ciencia. Por un lado, porque tienden a reducir la diversidad de la producción de conocimiento en términos de sus objetivos de investigación y enfoques –teóricos, metodológicos y empíricos–, en particular si dichos objetivos y enfoques cuestionan los fundamentos de la corriente principal –*mainstream*–. Por otra parte, porque tienden a inhibir o dificultar el desarrollo de nuevos campos científicos, en particular de los no disciplinares o que combinan conocimientos y enfoques de diversas disciplinas.

### III.2 Adaptación a las presiones del entorno: las respuestas estratégicas de los investigadores y el contenido de la producción de conocimiento

Los investigadores se anticipan y adaptan para hacer frente a las evaluaciones (van der Most, 2010). Es decir, cuando el acceso a recursos para la investigación depende del éxito en las evaluaciones, los investigadores adaptan su comportamiento porque saben *ex-ante* que la evaluación va a suceder. Además destaca, en coincidencia con Whitley (2007), que si el sistema de evaluación es cíclico o recurrente, es esperable que tenga mayores efectos sobre la producción de conocimiento que las evaluaciones aisladas o puntuales (van der Most, 2010).

---

<sup>7</sup> Con esta expresión el autor refiere al Evangelio según Mateo que dice: “*al que tiene se le dará más, pero al que no tiene, aún lo que tiene se le quitará*”. Así busca remarcar que aquellos que ostentan posiciones relativas superiores en el sistema de ciencia, tendrán mayores posibilidades de acceso a diversos recursos, lo que les permitirá continuar con la investigación. Sin embargo, quienes se encuentren en posiciones relativas más bajas, encontrarán diversas dificultades para investigar, debido a sus menores posibilidades de acceso a dichos recursos.

En sintonía, Gläser y Laudel (2007b) sostienen que los sistemas de evaluación son instituciones formales e informales que gobiernan la acción de evaluadores e investigadores (evaluados). De ese modo se fusionan las acciones de política con la conducta y el contenido de la investigación (Gläser y Laudel, 2007b). Es decir, estos autores plantean que el contenido de la investigación puede variar por la adaptación de las estrategias de investigación y de los enfoques para cumplir con la *expectativa de calidad* que sugiere el sistema de evaluación.

Además, dado que los criterios de evaluación miden solo aspectos parciales de la calidad de la producción de conocimiento, y que los aspectos de la calidad no medidos por el sistema de evaluación están asociados a ciertas características epistemológicas de la producción de conocimiento, entonces son esperables cambios –consecuencias- no deseados<sup>8</sup> sobre las características epistemológicas de la producción de conocimiento derivados de la implementación de *sistemas de evaluación fuertes* (Gläser y Laudel, 2007b). Es decir, los sistemas de evaluación fuertes pueden limitar ciertas características epistemológicas de los procesos de producción de conocimiento.

Estos autores identifican de la literatura la siguiente lista de características epistemológicas de la producción de conocimiento: i) tipo de investigación (teórica, metodológica, experimental, empírica) y orientación predominante de la investigación (básica, estratégica, aplicada); ii) relación de la investigación con la opinión mayoritaria de la comunidad (corriente principal –*mainstream-versus* no conformistas); iii) características temporales de los procesos de investigación (largo plazo *versus* corto plazo); iv) grado de heterogeneidad en la combinación de conocimientos (interdiscipliniedad); v) grado de riesgo intelectual asumido; y, vi) confiabilidad de los resultados. (Gläser y Laudel, 2007b).

En el contexto de esta tesis la definición conceptual de la dimensión o constructo que se ha denominado como *flexibilidad epistemológica del sistema de evaluación* se inspira –aunque no solamente- en las características epistemológicas identificadas por Gläser y Laudel (2007b). En el capítulo V –sobre marco teórico conceptual- se discuten en profundidad todos los aspectos vinculados a la definición conceptual de flexibilidad epistemológica del sistema de evaluación y de los restantes constructos relevantes para el análisis y los objetivos de investigación de esta tesis.

Siguiendo a Gläser y Laudel (2007b), la relación entre los efectos probables de la evaluación de los resultados de la investigación y las características epistemológicas de la producción de conocimiento, depende del conjunto de decisiones estratégicas que toman los investigadores, durante el proceso de producción de conocimiento, para *adaptarse* a sus condiciones de trabajo. De acuerdo con esos autores, dichas decisiones estratégicas se vinculan con: i) la selección de los temas y problemas a investigar; ii) los objetos y métodos de investigación; iii) el desarrollo de alianzas con otros investigadores –ya sean del ámbito nacional y/o internacional-, u otros actores extra-académicos –de la sociedad civil, empresas, gobierno, etc.-; y, iv) la selección de los medios de difusión de los resultados de investigación.

Importa marcar que esas decisiones estratégicas de los investigadores también están asociadas o afectan en última instancia los modos diferenciales de producción de conocimiento (Modo 1 y Modo

---

<sup>8</sup> En el capítulo siguiente, a partir del concepto *efecto constitutivo* introducido por Dahler-Larsen (2014), se refina esta idea de (Gläser y Laudel, 2007b) sobre los criterios de evaluación y sus *consecuencias no deseadas*.



2) en los que los investigadores trabajan predominantemente. En particular -o al menos- en lo relacionado con la selección de temas y problemas a investigar, el desarrollo de alianzas con otros actores (del ámbito académico –investigadores- o de otros ámbitos extra-académicos), y respecto a cuáles son los mejores medios para difundir los resultados de investigación. Además, los modos diferenciales de producción de conocimiento también están asociados a ciertas características epistemológicas de la producción de conocimiento. Específicamente porque los Modos 1 y 2 tienen prevalencias diferenciales respecto a su orientación predominante y la heterogeneidad en la combinación de conocimientos.

Dado que de acuerdo con la teoría, la evaluación del desempeño de la investigación afecta las decisiones estratégicas de los investigadores –quienes adaptan su comportamiento para cumplir con las exigencias que impone el sistema de evaluación- alterando con ello también las características epistemológicas de los procesos de producción de conocimiento que realizan; y si además, como se argumentó en el párrafo anterior, esas decisiones estratégicas y características epistemológicas también están asociadas a los modos diferenciales de producción de conocimiento; entonces es esperable que la evaluación del desempeño de la investigación también se relacione –afecte- los modos diferenciales de producción en que los investigadores predominantemente deciden trabajar.

A la vez y en la medida que dichas decisiones estratégicas son tomadas por los investigadores para adaptar su desempeño a las condiciones que le impone su entorno, se puede sostener que las mismas son dependientes del contexto. Adicionalmente, es posible señalar que dado que la naturaleza y la utilización de sistemas de evaluación de la investigación varían dependiendo de ciertas características específicas de los distintos contextos nacionales en que se implementen, también es razonable suponer variaciones en sus consecuencias probables sobre las características epistemológicas de la investigación y, sobre la dirección y el contenido de la producción de conocimiento, y en última instancia, sobre los modos diferenciales de producción de conocimiento. A continuación se desarrollan estas ideas.

### III.3 Los impactos esperados de los sistemas de evaluación fuertes dependen de las características específicas del contexto en que se implementen

Siguiendo a Whitley (2007), la institucionalización de distintos sistemas de evaluación de la investigación en los distintos sistemas académicos nacionales, es uno de los componentes importantes de los cambios en la gobernanza de la ciencia que se observan desde mediados de las década de los ochenta. En ese sentido, el autor destaca que no es posible comprender dichos cambios ni identificar sus consecuencias, si no se ubican en su contexto social específico.

De acuerdo con el autor, las consecuencias probables sobre la organización de la investigación y la dirección de la producción de conocimiento derivadas de la implementación de los sistemas de evaluación de la investigación se verán reforzados o amortiguados –variarán- dependiendo de las características específicas que asuman tres grandes dimensiones contextuales o ámbitos, a saber:

- i) el sistema público de Ciencia en términos de su *régimen de financiamiento*, es decir, del alcance y la periodicidad del financiamiento (presupuestal, proyectos, programas), la variedad de agencias para financiar la investigación y la incorporación -o no- de los objetivos de política pública en dichos financiamientos; de las *capacidades estratégicas*

- o autonomía de las organizaciones productoras de conocimiento; y, del grado de segmentación de esas organizaciones y del mercado de trabajo de los investigadores;*
- ii) *el sistema de evaluación, en particular la gobernanza y estructura del sistema, es decir, su frecuencia de evaluación, la formalización de sus procedimientos, el grado de estandarización de sus criterios de evaluación, su transparencia y unidades de evaluación (programas, proyectos, grupos, individuos); y sus implicancias para la asignación de recursos de investigación; y,*
  - iii) *los campos científicos, específicamente en relación a la diversidad de agencias de financiamiento a las que tienen acceso, a las audiencias que legitiman sus resultados, a la probabilidad que tienen los investigadores de los campos de acceder a fuentes de financiamiento externas, al alcance y variabilidad de los objetivos y objetos de investigación, marcos conceptuales y metodológicos de referencia, al grado de concentración en diversos –o pocos- problemas de investigación, a la centralidad y prestigio del campo en relación a los otros campos, y a la cohesión y prestigio de la élite científica. (Whitley, 2007).*

A continuación, se ponen en juego -a través de un escenario hipotético- las dimensiones contextuales específicas que median los efectos probables de los sistemas de evaluación fuertes sobre la organización de la investigación y la dirección de la producción de conocimiento.

Siguiendo a Whitley (2007), conforme: i) aumenta el financiamiento externo de la investigación basado en proyectos competitivos, ii) menor sea la variedad de agencias de financiamiento y sus objetivos, iii) mayor sea la frecuencia de la evaluación, iv) más fuerte sea el sistema de evaluación y sus implicancias en la asignación de recursos para la investigación, v) menor sea la diversidad de audiencias que legitiman la reputación de los investigadores, vi) menor sea la variabilidad de problemas, enfoques y objetivos de investigación en el campo, y vii) cuanto más fuerte y cohesiva sea la élite científica en la implementación de esos sistemas de evaluación; entonces es esperable que las consecuencias del sistema de evaluación fuerte se refuercen. A saber, que aumente la estratificación de individuos y de las organizaciones productoras de conocimiento, se incremente la concentración de los recursos para la investigación, se intensifique la competencia reputacional, y la investigación esté más integrada en torno a ciertos objetivos dominantes en el campo –mayor coordinación de los objetivos de investigación-, y con ello, aumente la importancia relativa de las disciplinas centrales, sus estándares y prioridades de investigación. Su contracara es el aumento de las restricciones para el desarrollo de nuevos campos científicos y su consecuente disminución de la diversidad intelectual y el pluralismo de estrategias y enfoques de investigación. (Whitley, 2007)

De acuerdo con el objetivo de investigación de esta tesis, estas ideas de Whitley (2007) son importantes para contextualizar y comprender las consecuencias de los cambios en la gobernanza de la ciencia que ocurrieron en México a inicios de la década de 1980 -cuando se institucionalizó un sistema de evaluación del desempeño de la investigación (el SNI) en el sistema público de CTI de ese país-, sobre la dirección de la producción de conocimiento, o más específicamente, sobre los modos diferenciales de producción de conocimiento (Modo 1 y Modo 2), en tanto representan direcciones distintas de la producción de conocimiento observable en ese contexto.

Influida por estas ideas y con el propósito de comprender la naturaleza de las relaciones inmersas entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de

conocimiento en el contexto mexicano, en esta tesis se retoman dos de las tres dimensiones propuestas por Whitley (2007) para alcanzar una comprensión contextualizada de la naturaleza de las relaciones entre un sistema de evaluación fuerte como el SNI y los Modo 1 y Modo 2 de producción de conocimiento de los investigadores categorizados en ese sistema.

En particular, en esta tesis se caracterizan el sistema público de CTI de México -en términos de su *régimen de financiamiento* y la *segmentación de las organizaciones productoras de conocimiento* y del *mercado de trabajo de los investigadores*-, y el sistema de evaluación -en términos de su *gobernanza y estructura* y sus *implicancias en la asignación de recursos para la investigación*-. Dado el alcance de la investigación de esta tesis y que una caracterización de los campos científicos -como la que propone Whitley (2007)- para el caso del SNI supone la realización múltiples esfuerzos de investigación en sí mismos, abordar esta dimensión contextual es materialmente imposible en el contexto de esta investigación. Sin embargo, debido a la importancia de los campos científicos en la relación entre la evaluación de la investigación y la producción de conocimiento, en esta tesis se controlan los resultados de las estimaciones realizadas a través de la introducción de un conjunto de variables independientes indicadoras del área de conocimiento. En el capítulo VI se caracterizan el sistema público de CTI de México y el SNI. A continuación, muy sintéticamente y a modo de adelanto, se enuncian parte de los resultados de ese trabajo.

Desde el punto de vista de la caracterización del sistema público de CTI de México, los resultados de la caracterización realizada muestran que: i) la investigación en México se financia con fondos externos, a través de proyectos de investigación altamente competitivos dada la gran escasez de recursos para la investigación y desarrollo (I+D), cuya duración va de uno a tres años; ii) una muy baja variedad de agencias y objetivos de financiamiento, muy concentrada en el Consejo Nacional de Ciencia y Tecnología (CONACYT) y en algunos de sus programas (SNI, Becas y Fondo de Ciencia Básica); y, iii) una alta concentración de las organizaciones productoras de conocimiento y del mercado de trabajo de los investigadores en muy pocas universidades públicas, donde se destaca la UNAM seguida de la UAM.

Desde el punto de vista de la caracterización del sistema evaluación, los resultados permiten sostener que el SNI es un *sistema de evaluación fuerte* del desempeño de la investigación realizada por los investigadores categorizados en ese sistema, con una relativamente alta frecuencia de la evaluación -entre tres y cuatro años en promedio aproximadamente- y, a la vez, con la exigencia de cierta cantidad de publicaciones internacionales anuales, y con implicancias importantes para el acceso a recursos externos para la investigación.

Dados esos resultados y siguiendo a la teoría, es razonable esperar consecuencias notables del SNI sobre el contenido y dirección de la producción de conocimiento mexicana, en particular en relación a los modos diferenciales de producción de conocimiento. De aquí también que si la teoría es cierta, entonces se deberían observar relaciones significativas entre la evaluación del desempeño de la investigación que se realiza en el SNI y los modos diferenciales de producción de conocimiento.

Adicionalmente, los resultados de la caracterización realizada tanto del SNI como del sistema público de CTI de México, permitió una comprensión analítica de ciertas tensiones que emergen de los debates públicos -entre políticos, gestores, autoridades de las universidades u otras organizaciones, evaluadores e investigadores- sobre el SNI y sus efectos percibidos. Muchos de esos efectos percibidos, se vinculan con las consecuencias indicadas por Whitley (2007) en el escenario

hipotético planteado anteriormente. En particular respecto a: i) la estratificación de los académicos –entre los que están categorizados en el SNI y por ello son reconocidos socialmente como investigadores y los que no están, aunque investiguen, no son reconocidos como tales-, de las organizaciones productoras de conocimiento, grupos y posgrados; ii) la concentración de los recursos para la investigación; iii) el aumento de la competencia entre los académicos tanto por el reconocimiento y prestigio que provee el SNI como por el acceso a los recursos para la investigación; y iv) la relativa disminución de la diversidad de la producción de conocimiento, en términos de sus temas y objetivos de investigación, enfoques y orientación predominante. En el capítulo VII se plantean en profundidad todos estos temas, en base a una revisión del debate nacional en torno al SNI y sus consecuencias.

#### III.4 La gobernanza de la ciencia: los balances de poder y su relación con la producción de conocimiento

Adicionalmente y según Whitley (2010), las consecuencias probables sobre la organización de la investigación y la dirección de la producción de conocimiento, derivadas de la implementación de los sistemas de evaluación fuertes, también estarán reforzados o amortiguados de acuerdo con las distribuciones relativas de poder que resultan de las relaciones de autoridad que tengan lugar en cada contexto específico.

Whitley (2010) coincide con Ziman (2000), en que los cambios en las formas en que se produce, difunde y evalúa el conocimiento científico en los distintos sistemas públicos de ciencia de la mayoría de las sociedades industrializadas, son observables desde el final de la Segunda Guerra Mundial. En sintonía también señala que, entre las transformaciones más notorias que dieron lugar a esos cambios en los distintos sistemas públicos de ciencia están: i) el *estancamiento en el crecimiento del financiamiento público de la investigación*, que fue acompañado por una *mayor competencia* para acceder a recursos públicos para la investigación; ii) el desarrollo de *políticas orientadas a direccionar la producción de conocimiento* hacia aquellas áreas que ofrecieran mayores retribuciones económicas del financiamiento público de la investigación, reorganizando así las relaciones entre ciencia y sector empresarial; y, iii) la *institucionalización y difusión de diversos procedimientos de evaluación* para ponderar tanto el desempeño de individuos y organizaciones productoras de conocimiento, como para auditar sus productos y resultados de investigación. (Whitley, 2010)

Simultáneamente, Whitley (2010) resalta que esas transformaciones han cambiado las relaciones de autoridad que gobiernan la selección de los objetivos de la ciencia y la evaluación de sus resultados en gran parte de los países de la OCDE. Sostiene que el cambio en las relaciones de autoridad es consecuencia de la participación activa de diversos actores que inciden en el sistema de ciencia, tales como: las agencias de financiamiento públicas, nacionales e internacionales, las élites y comunidades científicas nacionales e internacionales, el Estado, las organizaciones productoras de conocimiento, y otros actores empresariales o de la sociedad civil. A la vez, la incidencia de esos actores se expresa en la elección y definición de los problemas de investigación, en la selección de los enfoques teóricos y metodológicos para abordarlos, en la expansión a gran escala de las publicaciones científicas en revistas arbitradas, libros y otros medios de difusión, en la alta especialización e internacionalización de los campos de investigación y sus comunidades académicas, tanto en términos de la evaluación como en el reconocimiento de los méritos de los

resultados de investigación y su incorporación –o no- dentro del conjunto de conocimientos científicos formales y, en la asignación de recursos para la investigación. (Whitley, 2010)

De acuerdo con el autor, lo anterior ha transformado significativamente las condiciones o balances de poder que gobiernan la producción, coordinación y control del conocimiento científico formal. Es más, sostiene que muchos de los cambios señalados contribuyen a mejorar más las capacidades estratégicas y la influencia de las agencias de financiamiento, las élites científicas y las organizaciones empleadoras de los investigadores, que las de los investigadores, sus estrategias de investigación y las normas que rigen su funcionamiento. (Whitley, 2010)

En ese sentido, destaca que es esperable que las transformaciones en las relaciones de autoridad, que gobiernan las prioridades de investigación y los criterios de evaluación, también tengan un conjunto de consecuencias importantes sobre la intensidad de la competencia por prestigio disciplinar, el alcance nacional e internacional de la coordinación de las estrategias de investigación, el período de tiempo en que los investigadores esperan alcanzar resultados de investigación interesantes y publicables, y la facilidad con la cual pueden establecerse nuevos campos y líneas de investigación científica. (Whitley, 2010)

Gläser (2012) sugiere algunas ideas para el desarrollo de una teoría de rango medio para vincular las variaciones en la *governanza* de la ciencia con los cambios en el contenido de la investigación. Este autor, siguiendo el trabajo de Whitley (2010) recién referenciado y el de Rosenau (2004), define *governanza* como “*la construcción y ejercicio de autoridad en un espacio complejo y heterogéneo de actores interdependientes*” (Gläser, 2012: 3).

El autor, parte de las situaciones que hacen que los investigadores tomen ciertas decisiones estratégicas. En ese marco, sostiene que en realidad, tales situaciones son *influencias o palancas* para la *governanza* –“*levers*” for *governance*-. Su fundamento se basa en que dichas situaciones están moldeadas por actores externos que pueden influir indirectamente en las decisiones sobre el contenido de la producción de conocimiento. Es decir, los investigadores tomarán ciertas decisiones estratégicas en el proceso de producción de conocimiento, para poder cumplir con las expectativas que emergen de los actores de *governanza* (Gläser, 2012).

Es así que los cambios en la *governanza* de la ciencia han determinado una transformación general del contexto en el que los investigadores toman sus decisiones estratégicas sobre sus procesos de producción de conocimiento. Como se verá en el capítulo siguiente, en la actualidad, la *governanza* de la ciencia está caracterizada por la amplia difusión de indicadores bibliométricos como criterios centrales, sino únicos, de evaluación del desempeño de la investigación.

### III.5 Conclusión: las relaciones entre la evaluación y los modos diferenciales de producción de conocimiento: una justificación y el fundamento teórico de las relaciones estudiadas en esta tesis

A lo largo de este capítulo se han revisado las contribuciones de los autores que han vinculado teóricamente a la evaluación del desempeño de la investigación con la dirección, contenido y características epistemológicas de la producción de conocimiento. A continuación y en primer lugar se ofrece una síntesis de estas contribuciones. En segundo lugar, se argumenta sobre los elementos teóricos que permiten integrar estas contribuciones con las de Gibbons et al. (1994) y se justifica y fundamenta teóricamente el análisis de las relaciones estudiadas en esta tesis.

Tal como se planteó en las secciones III.1 y III.2 de este capítulo, la *evaluación del desempeño de la investigación* incide en la *dirección* (Whitley, 2007), *contenido y características epistemológicas de la producción de conocimiento* (Gläser y Laudel, 2007b). El mecanismo central que conecta a la evaluación de los resultados de la investigación con el contenido de la producción de conocimiento, es la respuesta estratégica de los investigadores que se expresa en las decisiones que toman sobre los procesos de investigación que emprenden, para adaptarse a las presiones que le impone su entorno (Gläser y Laudel, 2007b), y en particular, para cumplir con las exigencias que les impone el sistema de evaluación. Dicha adaptación será tanto más fuerte según los incentivos materiales y simbólicos que provea el sistema de evaluación, ya sea en términos del mejor acceso a recursos para la investigación así como en relación al reconocimiento social y prestigio que otorga.

De acuerdo con lo planteado en la sección III.3, la magnitud y alcance de la incidencia de la evaluación sobre la dirección, contenido y características epistemológicas de la producción de conocimiento, está mediada por las características contextuales específicas que asuman el sistema público de ciencia, el sistema de evaluación del desempeño de la investigación, en particular si se trata de un *sistema de evaluación fuerte* y los diversos campos científicos. (Whitley, 2007)

A la vez y de acuerdo con lo señalada en la sección III.4, dichos sistemas de evaluación inciden en la *gobernanza de la ciencia*. Los sistemas de evaluación fuertes inciden en la gobernanza de la ciencia en la medida que afectan las relaciones de autoridad y los (des)balances de poder entre los actores que gobiernan la selección de los objetivos de la ciencia y la evaluación de sus resultados (Whitley, 2010). De acuerdo con Gläser (2012), el contenido de la investigación cambia porque los investigadores tomarán ciertas decisiones estratégicas en el proceso de producción de conocimiento, para poder cumplir con las expectativas que emergen de los actores de gobernanza (Gläser, 2012).

El concepto de *gobernanza de la ciencia* permite integrar la teoría que vincula a la evaluación y la producción de conocimiento con las contribuciones de Gibbons et al. (1994) sobre los modos diferenciales de producción de conocimiento y su control de calidad.

Siguiendo a Gibbons et al. (1994) el control de calidad –evaluación– de la producción de conocimiento científico y tecnológico depende del *consenso provisional alcanzado en una comunidad de practicantes* (Gibbons et al., 1994), o por la *élite científica* (Whitley, 2007). En términos más amplios, ese *consenso* depende de los balances de poder derivados de las relaciones de autoridad que gobiernan las prioridades de investigación y los criterios de evaluación en cada contexto específico (Whitley, 2010). En ese sentido, “*los sistemas de evaluación son construcciones sociales en disputa*” (Bianco, Gras y Sutz, 2016: 400).

Además, tal como se presentó en la sección II.2 del capítulo anterior, Gibbons et al. (1994) caracterizan la *producción de conocimiento* de acuerdo a sus formas diferenciales de producción en *Modo 1* y *Modo 2*. De acuerdo con esos autores, los Modo 1 y Modo 2 de producción de conocimiento se vinculan con la *gobernanza de la ciencia*.

Por un lado y como se describió en las sub-secciones II.2.1 a II.2.4 del capítulo anterior, los Modo 1 y Modo 2 de producción de conocimiento se vinculan con la gobernanza de la ciencia porque dependiendo del *Modo de producción*, estará definido qué conocimiento se produce, cómo se produce, en qué contextos, cómo se organiza y difunde. De aquí es que en esta tesis se sostiene que

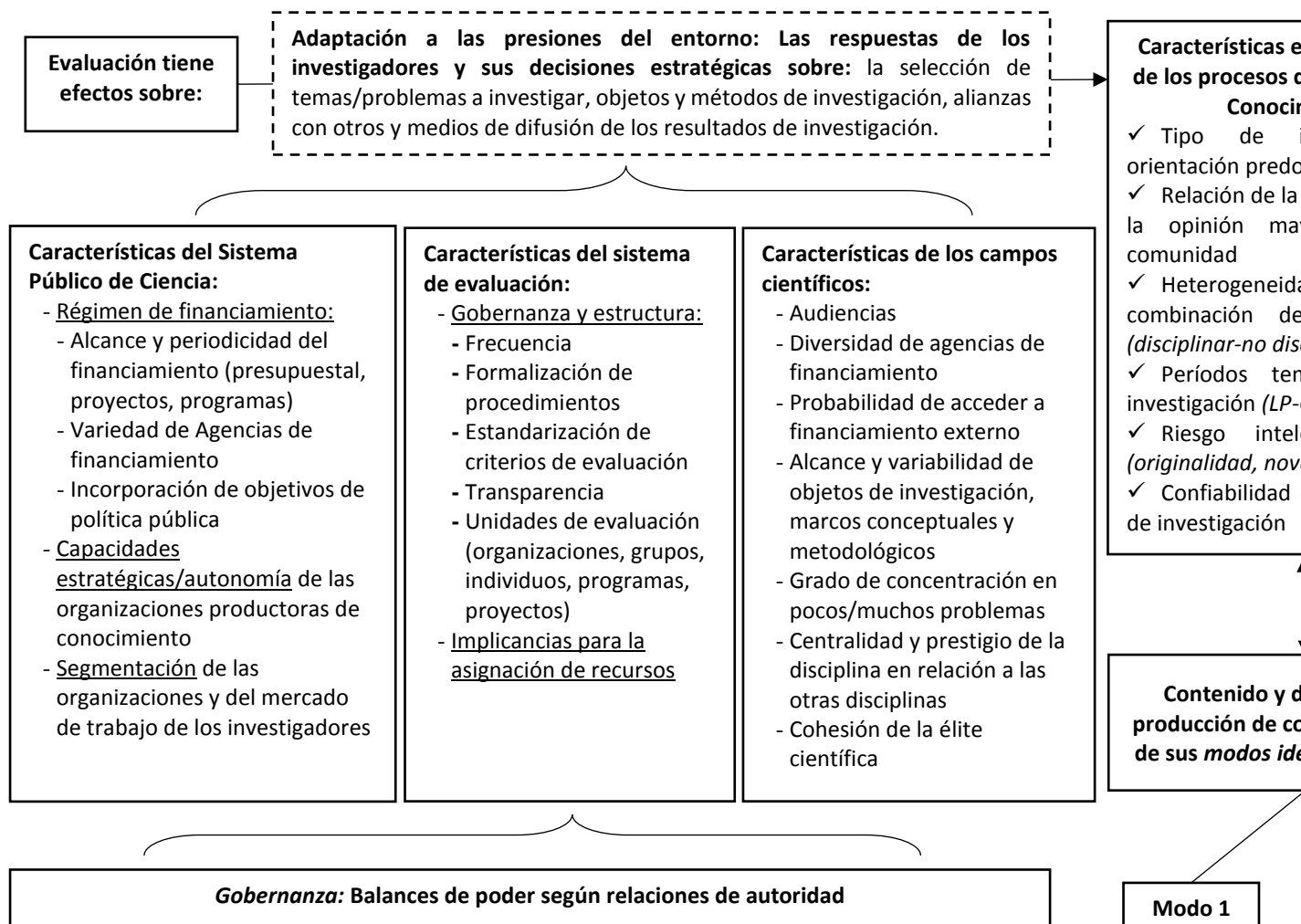
cada uno de esos modos de producción de conocimiento están asociados con *contenidos* y *direcciones* distintas de la producción de conocimiento. Además, los Modos 1 y 2 responden a decisiones estratégicas distintas de los investigadores sobre sus procesos de producción de conocimiento y exhiben prevalencias diferenciales respecto a su orientación predominante y la heterogeneidad en la combinación de conocimientos. De aquí es que en esta tesis se sostiene, que cada uno de esos modos de producción de conocimiento también están asociados con ciertas *características epistemológicas de la producción de conocimiento*. Hasta aquí los elementos teóricos que permiten integrar las contribuciones de Gibbons et al. (1994) con las de Whitley (2007) y Gläser y Laudel (2007b)

Por otro lado y con base en las contribuciones de Gibbons et al. (1994), de Jong et al. (2011), Regeer et al. (2009), Spaapen, Dijstelbloem y Wamelink (2007) y Hemlin y Barlebo Rasmussen (2006) descritas en la sub-sección II.2.5 y en la sección II.3 del capítulo anterior, los Modo 1 y Modo 2 de producción de conocimiento se vinculan con la gobernanza de la ciencia, también porque requieren de sistemas de evaluación del desempeño de la investigación flexibles y contextualizados para estimular integralmente la “buena ciencia”. Es decir, sistemas de evaluación que sean capaces de valorar la calidad, desempeño e impacto del conocimiento producido en cada uno de esos modos, recompensando y reconociendo los múltiples objetivos de investigación que se plantean, los diversos productos que generan, las características propias del contexto en el que se produce el conocimiento, sus usos y la valoración que hacen de los resultados de investigación las audiencias para las cuáles ese conocimiento fue creado.

Entonces, si la evaluación del desempeño de la investigación incide en el contenido, dirección y características epistemológicas de la producción de conocimiento y si además, esas características, contenido y dirección de la producción de conocimiento están asociadas con los modos diferenciales de producción de conocimiento, entonces es razonable esperar que dicha evaluación también se relacione –afecte– los Modos 1 y 2 de producción de conocimiento. De aquí el fundamento teórico y la justificación de que en esta tesis se espera encontrar relaciones significativas entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento.

A continuación se presenta la figura 1 que representa una síntesis de las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento.

Figura 1: La evaluación de la investigación y los modos diferenciales de producción de conocimiento



Fuente: Elaboración propia con base en Whitley (2007 y 2010), Gläser y Laudel (2007b), Gläser (2012) y Gibbons et al. (1994).

Notas: Las flechas: i) unidireccionales indican relaciones de dependencia; y, las ii) bidireccionales indican asociación. Las llaves representan mecanismos. El rectángulo punteado representa mecanismos.



## CAPITULO IV: Los *indicadores bibliométricos* y sus *efectos constitutivos* sobre el contenido, dirección y características epistemológicas de la producción de conocimiento

Este capítulo ofrece una revisión bibliográfica sobre diversos aspectos vinculados con la evaluación del desempeño de la investigación y la producción de conocimiento. Se revisa la evolución y amplia difusión de los indicadores bibliométricos en la gobernanza de la ciencia, la discusión sobre su uso como criterios centrales de evaluación y sus implicancias sobre la dirección, contenido y características epistemológicas de la producción de conocimiento de acuerdo con los hallazgos de investigaciones pasadas reportados en la literatura. Estos hallazgos de investigaciones pasadas sostienen, con la evidencia que reportan, las premisas de la teoría que vincula a la evaluación de la investigación con la producción de conocimiento revisada en el capítulo anterior. Además, también proveen los elementos teóricos que justifican y fundamentan la construcción que se hace en esta tesis de las dimensiones conceptuales relevantes para los objetivos de investigación de este trabajo, así como la dirección de las relaciones aquí estudiadas.

El capítulo se estructura en cuatro secciones. Le sigue a esta introducción, la sección IV.1 que se centra en las tensiones que genera, en el ámbito internacional, la amplia difusión de los indicadores bibliométricos como criterios centrales en la evaluación del desempeño de la investigación. Se revisan algunas de las razones que permiten entender la rápida difusión de estos indicadores en la gobernanza de la ciencia, cierta parte de sus limitaciones, las críticas y el cuestionamiento que diversos actores del ámbito internacional hacen sobre su uso en la evaluación y las recomendaciones que éstos realizan orientadas al desarrollo de buenas prácticas de evaluación. En la sección IV.2 se analiza el problema de la validez de los indicadores en general y de los bibliométricos en particular y se plantean sus *efectos constitutivos* teóricos. La sección IV.3 presenta una revisión de la literatura –internacional y nacional- que documenta múltiples efectos sobre la dirección, contenido y características epistemológicas de la producción de conocimiento, derivados de la implementación de sistemas de evaluación fuertes, cuyos criterios se basan centralmente en indicadores cuantitativos (bibliométricos). Finalmente, la sección IV.4 concluye con una síntesis de los hallazgos de investigaciones pasadas sobre la relación entre la evaluación cuantitativa del desempeño de la investigación y sus efectos constitutivos sobre la dirección, contenido y características epistemológicas de la producción de conocimiento que inspiran la construcción de las dimensiones relevantes para el análisis y que dan sentido a las hipótesis que se ponen a prueba en esta tesis. Las dimensiones de análisis relevantes para esta tesis y las hipótesis que se hacen sobre las relaciones entre ellas se fundamentan y justifican teóricamente en el capítulo siguiente, retomando ciertos conceptos y sus relaciones que emergen tanto de este capítulo como de los dos anteriores.

### IV.1 La difusión de los indicadores bibliométricos en la gobernanza de la ciencia y el cuestionamiento internacional

Si bien diversos autores coinciden en que distintas formas de producción de conocimiento requieren de enfoques diferenciales de evaluación de la investigación (Gibbons et al., 1994; de Jong et al., 2011; Regeer et al., 2009; Spaapen, Dijstelbloem y Wamelink, 2007; Hemlin y Barlebo Rasmussen, 2006, entre otros), la evaluación por pares -ha sido y sigue siendo- la forma tradicional e histórica

utilizada para evaluar los resultados de investigación y, a partir de mediados de la década de 1970, la evaluación centrada en indicadores bibliométricos comienza a ser reconocida en los ámbitos de la política científica y de la gestión de la ciencia (Gläser y Laudel, 2007a).

Durante la década de 1980, los métodos bibliométricos siguieron ganando espacio en la evaluación y además cobraron importancia en la toma de decisiones sobre la asignación de recursos para el financiamiento de la investigación. Entre los casos más conocidos en el ámbito internacional están los sistemas de evaluación del Reino Unido, Australia y Nueva Zelanda. (van der Most, 2010).

Durante la década de 1990 los *rankings* de universidades comenzaron a difundirse internacionalmente (van der Most, 2010) y, según Hicks et al. (2015), a las universidades de todas partes del mundo cada vez les *obsesiona* más obtener una posición relativa mejor en los *rankings* internacionales –por ejemplo en el de *Shanghai*, en el del *Times Higher Education*, etc-.

En la actualidad, Hicks et al. (2015) señalan que el uso de métricas que buscan proveer una medida de las actividades científicas es cada vez más importante en la *gobernanza de la ciencia*. Argumentan que la evaluación de la investigación se ha convertido en una actividad cada vez más rutinaria y centrada en ciertas métricas.

Estas métricas refieren fundamentalmente al cálculo de diversos indicadores bibliométricos tales como: número de publicaciones, número totales de citas, promedio de citas por artículo, por autor, por año, número de autores por artículo, factor de impacto de las revistas, *índice-h*, *índice-g* entre otros.

En 1955 Eugene Garfield en su artículo "*Citation Indexes for Science. A New Dimension in Documentation through Association of Ideas*", publicado en *Science*, propone un sistema bibliográfico para la literatura científica con el objetivo de construir índices de citación. Es en este momento que plantea sus ideas para el desarrollo de un índice nuevo de citas, que denominó *factor de impacto*<sup>9</sup> (Garfield, 1955).

No es casual que Eugene Garfield haya fundado en 1960 el *Institute for Scientific Information* (ISI), que hasta 2016 fue propiedad de Thomson Reuters (Martin, 2016), año en que Thomson Reuters vendió la propiedad intelectual de ese y otros productos a Clarivate Analytics<sup>10</sup>. Desde 1995 el interés en el factor de impacto de las revistas no ha dejado de aumentar (Hicks et al., 2015).

Sahel (2011) señala que los índices de citaciones han evolucionado, lo que se expresa a través del desarrollo de nuevos indicadores a partir de la existencia de las bases de datos bibliográficas, que contienen -para la mayoría de las disciplinas- una parte importante de su literatura científica y sus citas.

Hasta el año 2000, los análisis especializados realizados por los expertos en bibliometría se realizaban con base en el *Science Citation Index* del ISI en su versión de CD-ROM (Hicks et al., 2015). En 2002 Thomson Reuters pone a disposición de un público amplio la base de datos

---

<sup>9</sup> El cálculo del factor de impacto "[s]e basa en 2 elementos: el numerador, que es el número de citas en el año en curso que recibe cualquier artículo publicado en la revista en los 2 años anteriores, y el denominador, que es el número de artículos sustantivos publicados en los mismos 2 años." (Garfield, 1999: 979, citado en Martin, 2016:2).

<sup>10</sup> Ver <https://www.thomsonreuters.com/en.html>

bibliográfica *Web of Science* (Hicks et al., 2015), que incluye el *Science Citation Index* (SCI), el *Social Science Citation Index* (SSCI) y el *Art and Humanities Citation Index* (A&HCI), con los cuales se construye el *Journal Citation Report* (JCR) y los *Essential Science Indicators* (ESI). Después de ese momento, en 2004, nuevas bases de datos bibliográficas<sup>11</sup>, con sus respectivos índices de citas, aparecen para competir en el mercado dominado por *Web of Science*, entre ellas *Scopus* de Elsevier y *Google Scholar* (Hicks et al., 2015).

En 2005, Jorge Hirsch en su artículo “*An index to quantify an individual’s scientific research output*”, publicado en *PNAS*, propone el *índice-h* para caracterizar la producción científica de un investigador y aproximar una medida del *impacto científico de su trabajo* (Hirsch, 2005). Este índice conceptualiza impacto como: el número de artículos científicos de un investigador con al menos *h* citas en cada uno de ellos, porque en realidad es eso lo que mide. Por lo tanto, un investigador *x* con un *índice-h* igual a *n* –o impacto igual a *n*- implica que el investigador *x* ha publicado *n* artículos y que cada uno de ellos ha sido citado al menos *n* veces (Hirsch, 2005). Hicks et al. (2015) señalan que el *índice-h* “*popularizó el recuento de citas de investigadores individuales*” (Hicks et al., 2015:1).

Hirsch (2005) señala que el *índice-h* supera las limitaciones de otros indicadores bibliométricos individuales, entre ellos: i) la cantidad total de artículos publicados que no tiene en cuenta una medida de la relevancia de esa producción; y, ii) la cantidad total de citas, que puede proveer una medida de impacto sobrevalorada, debido a la existencia de algunos “*grandes éxitos*” no representativos y porque no pondera los artículos de revisión -generalmente mucho más citados que los artículos originales de investigación-.<sup>12</sup> En el ámbito internacional, algunas organizaciones piden el *índice-h* a los aspirantes a sus plazas, o hacen depender los ascensos de la superación de

---

<sup>11</sup> Importa señalar que las bases de datos bibliográficas varían significativamente de acuerdo a la cobertura de las disciplinas, tipos de literatura (artículos, libros, capítulos de libros, revisiones, reseñas, etc.) e idiomas. Además, los indicadores -que de ellas se derivan- no están normalizados, con lo cual sus valores también dependen significativamente de los tamaños de las comunidades científicas en cada una de las disciplinas o subdisciplinas.

<sup>12</sup> Para una descripción detallada de todas las limitaciones que presentan otros indicadores y que el *índice-h* supera, ver Hirsch (2005). La principal ventaja del *índice-h* es que provee simultáneamente una medida que integra la cantidad de artículos publicados por un investigador a lo largo del tiempo y el número de citas acumuladas de cada artículo, por ello es preferible a otras medidas simples (Sahel, 2011). Sin embargo, el *índice-h* tampoco está exento de limitaciones, a saber: i) sus resultados varían significativamente de acuerdo al campo científico (Hicks et al., 2015); ii) favorece a los investigadores consolidados porque es acumulativo y no decrece con la edad (Hicks et al., 2015) ni con la interrupción o discontinuidad de la investigación; iii) sus valores para un mismo autor en el mismo período de tiempo varían según la cobertura de la base de datos bibliográfica (Hicks et al., 2015) que se utilice para su cálculo; iv) no considera el contexto de la citas (si se trata de una crítica, rectificación de resultados, etc.); v) está fuertemente afectado por el número total de artículos, así subestima la producción científica realizada por investigadores con carreras cortas o que han publicado pocos artículos, aunque todos ellos sean de gran relevancia; y, vi) no distingue artículos muy influyentes y puede desconocer artículos muy citados. (Sahel, 2011). Este autor señala que el *índice-g* es similar al *índice-h* -aunque el valor del primero siempre es mayor que el valor del segundo- y reconoce a los artículos más citados, ya que su forma de cálculo les otorga una mayor ponderación. El *índice-g* fue introducido por Egghe (2006) –citado en Sahel (2011)- como una mejora del *índice-h* para alcanzar una medida más apropiada del rendimiento global en términos de las citas que recibe el conjunto de artículos publicados por un investigador.

ciertos valores umbrales de dicho índice y de la cantidad de artículos publicados en revistas de *alto impacto*. (Hicks et al., 2015)

A pesar de la amplia difusión de los métodos bibliométricos en las prácticas de evaluación, el debate internacional -entre investigadores, hacedores de políticas públicas, gestores de la ciencia y evaluadores- también está abierto. No existe consenso sobre cómo evaluar y los criterios a utilizar tanto para apreciar y valorar las formas diferenciales de producción de conocimiento como para ponderar, incluso en el ámbito de la ciencia académica, la relevancia científica, la pertinencia (social, productiva o de política), la originalidad, la calidad y la heterogeneidad en la combinación de conocimientos (MIT) de los resultados de los procesos de producción de conocimiento. A continuación se revisa este debate.

Existe una creciente inconformidad internacional sobre la difusión de los indicadores bibliométricos como criterios centrales de evaluación. A continuación se discuten algunas razones de su difusión, así como sus críticas y principales recomendaciones.

La evaluación del desempeño de los individuos es uno de los componentes esenciales de la evaluación de la investigación, cuyos resultados –simultáneamente- tienen un rol muy importante en las estrategias organizacionales (Sahel, 2011) de universidades, de agencias públicas de financiamiento u otras organizaciones (empresas, centros públicos de investigación y tecnología, institutos de investigación, etc.) del sistema de CTI (Bianco, Gras y Sutz, 2014). Sin embargo, no existen indicadores ni estándares internacionalmente aceptados a través de los cuales se pueda medir objetivamente el desempeño científico (Sahel, 2011).

El proceso de evaluación de la investigación, ya sea *ex ante* (por ejemplo de propuestas de investigación) o *ex post* de resultados (por ejemplo del desempeño de individuos, de grupos de investigación o de organizaciones) encuentra diversas dificultades debido a su complejidad y multidimensionalidad, y por lo mismo, no ha encontrado, hasta la actualidad, soluciones estándares –satisfactorias para todos los actores involucrados- para superarlas (Bianco, Gras y Sutz, 2014).

Expresiones claras que dan cuenta de la inconformidad o insatisfacción de la implementación de medidas bibliométricas como criterios centrales de evaluación –sino únicos- del desempeño de la investigación -o de la calidad de de las revistas científicas- en el ámbito internacional son -entre otras-:

- i) el informe crítico realizado por la Academia Francesa de Ciencias (AFC) publicado en Enero de 2011<sup>13</sup> (en adelante AFC, 2011);
- ii) la declaración titulada: “*Declaration on Research Assessment: Putting science into the assessment of research*”<sup>14</sup>, elaborada en Diciembre de 2012 (en adelante DORA, 2012) durante la reunión anual de la Sociedad Americana de Biología Celular en San Francisco, avalada y firmada<sup>15</sup> por la comunidad de biólogos celulares, otros investigadores de

---

<sup>13</sup> Disponible en [http://www.sauvonsluniversite.com/IMG/pdf/avis\\_170111.pdf](http://www.sauvonsluniversite.com/IMG/pdf/avis_170111.pdf), fecha de consulta actualizada 10/09/2017.

<sup>14</sup> El texto completo de la declaración y sus firmantes puede encontrarse en <http://www.ascb.org/dora/>, fecha de consulta actualizada 10/09/2017.

<sup>15</sup> Al 10/09/2017 DORA ha sido avalada por 12788 “*individuos firmantes*” y 871 “*organizaciones firmantes*” (<http://www.ascb.org/dora/>).

- diversos campos científicos, editores de revistas científicas, directores de academias de ciencias y universidades, y responsables de la política científica, de diversas partes del mundo;
- iii) la editorial de la revista *Science*, realizada por Bruce Alberts y publicada en su Vol. 340 en Mayo de 2013 (en adelante Alberts, 2013);
  - iv) el *Manifiesto de Leiden* titulado “*The Leiden Manifesto for research metrics*”, realizado por Diana Hicks, Paul Wouters, Ludo Waltman, Sarah De Rijcke e Ismael Rafols, y publicado en Abril de 2015 en la revista *Nature* en su Vol. 520<sup>16</sup> (referido como Hicks et al., 2015). El contenido del *Manifiesto* fue discutido en el marco de la Conferencia Internacional sobre indicadores de CTI que tuvo lugar en Septiembre de 2014 en Leiden; y,
  - v) la editorial de la revista *Research Policy*, realizada por Ben Martin y publicada en su Vol. 45 en Febrero de 2016 (referido como Martin, 2016).

El informe de la AFC (2011) es crítico respecto al uso de indicadores bibliométricos como métodos para evaluar el desempeño de los investigadores individuales. Simultáneamente, realiza un conjunto de recomendaciones para integrar en la evaluación criterios que permitan valorar y juzgar correctamente la calidad de las contribuciones intelectuales que realizan los investigadores. La recomendación más general y de fondo de la AFC es que los indicadores bibliométricos solo pueden utilizarse si son complementarios a la evaluación cualitativa realizada por pares -expertos-.

En coincidencia, la preocupación central de la declaración DORA está puesta en la necesidad de mejorar las prácticas de evaluación de los resultados de la investigación científica que implementan las agencias de financiamiento y las organizaciones académicas, y plantea un conjunto de recomendaciones para ello (DORA, 2012). En consonancia con la AFC, DORA (2012) destaca que los productos de la investigación científica son múltiples y variados, y señala la necesidad de evaluar con precisión e inteligencia la calidad y el impacto de la producción científica.

Alberts (2013) señala su coincidencia con la totalidad de los contenidos de la declaración (DORA, 2012) y enfatiza la necesidad de detener el uso del factor de impacto en la evaluación de la investigación. Esa necesidad, entre otras cosas, encuentra su fundamento en la falta de credibilidad y validez de dicho indicador como consecuencia de ciertas estrategias, cuestionables éticamente, que son implementadas por algunos editores de las revistas científicas, para incrementar artificialmente sus respectivos factores de impacto.

Martin (2016) explica cómo se calcula el factor de impacto de una revista, plantea todas sus limitaciones técnicas y describe las diversas estrategias empleadas por los editores para aumentarlo. Simultáneamente señala que mientras que algunas de esas estrategias son aceptables, otras como las “*citas coercitivas*” y los “*cárteles de citas cruzadas*”, claramente incumplen las convenciones sobre comportamiento académico, y otras están en algún lugar intermedio entre esas dos posiciones. Entre sus conclusiones señala que el factor de impacto como indicador para la evaluación ha perdido credibilidad.

---

<sup>16</sup> La traducción al español de este material está disponible en [www.ingenio.upv.es/manifiesto](http://www.ingenio.upv.es/manifiesto), fecha de consulta actualizada 10/09/2017.

El *Manifiesto de Leiden*, en coincidencia con la AFC (2011), DORA (2012) y Alberts (2013), plantea que el problema está en que la evaluación de la investigación se rutinizó, dejó de estar fuertemente basada en la valoración cualitativa de los *expertos*, y pasó a depender centralmente –o exclusivamente- de ciertas métricas. Sostiene que los indicadores utilizados en la evaluación si bien cuentan con amplia difusión internacional, y en general, están bien intencionados, no siempre están bien informados y, en la gran mayoría de los casos, son mal aplicados. Así y como consecuencia de la aplicación irreflexiva de ciertos indicadores y el desconocimiento de sus implicancias –por parte de las organizaciones encargadas de las evaluaciones-, los mismos instrumentos que buscaban fortalecer el sistema científico, terminan aumentando el riesgo de que éste sea dañado (Hicks et al., 2015).

Dicho manifiesto plantea diez principios que sintetizan las *buenas prácticas* de la evaluación basada en indicadores métricos. Busca proveer elementos de juicio tanto “*para que los investigadores puedan pedir cuentas a los evaluadores, [como] para que los evaluadores puedan pedir cuentas a los indicadores*”. (Hicks et al., 2015:2)

De acuerdo con Hicks et al. (2015), si bien los indicadores pueden proporcionar información sustantiva para la evaluación, no pueden ser la única fuente para informarla. Estos autores, en coincidencia con la AFC (2011) y DORA (2012), destacan que las mejores decisiones –en materia de evaluación- requieren que se tomen sobre la base de *evidencia cuantitativa y cualitativa*, integrando diversos indicadores robustos que sean sensibles a los objetivos y naturaleza de la investigación evaluada. Es decir, las “[d]ecisiones sobre la ciencia tienen que ser tomadas en base a procesos de alta calidad informados por datos de la mayor calidad.” (Hicks et al., 2015:5)

En la misma dirección, Sahel (2011) planteaba que la evaluación del desempeño de investigación de los individuos es una tarea lo suficientemente compleja como para que descansa exclusivamente en medidas bibliométricas que, por sí mismas, no son capaces de realizar (Sahel, 2011).

De acuerdo con Gläser y Laudel (2007a), la rápida difusión de los indicadores bibliométricos en las prácticas de evaluación fue estimulado por un paradigma de política científica que se construye sobre el supuesto de que la *competencia* representa la forma más *eficiente* de asignación de recursos públicos para la investigación, porque de ese modo, sólo los mejores *productores* –en este caso de conocimiento- sobrevivirán. Premisas tales como: *eficiencia, transparencia, rendición de cuentas –accountability-, calidad y competencia*, propias del enfoque de la nueva gestión pública –*new public management*- marcaron la gestión del sector público (Power, 1996) con consecuencias importantes sobre la *gobernanza de la ciencia*, en particular respecto a los sistemas de evaluación y la asignación de recursos públicos para la investigación (Whitley y Gläser, 2007; Gläser y Laudel, 2007a; Whitley, 2010).

Martin y Whitley (2010) en su estudio sobre el *Research Assessment Exercise* (RAE) del Reino Unido muestran la incidencia que tuvieron dichas premisas. Señalan que la reducción del gasto público –a partir de la decisión de Margaret Thatcher, quien asumió el poder en 1979- indujo a cambios en la gestión de las organizaciones financiadas con recursos públicos, en particular respecto a: i) la definición de sus objetivos; ii) la toma de decisiones; iii) la instalación de evaluaciones de desempeño; y, iv) la implementación de un sistema de asignación de recursos basado en resultados. Cambios similares se pueden encontrar en México a inicios de la década de 1980 (Bensusán et al., 2014) y en Suiza en la década de 1990 (Benninghoff y Braun, 2010), ambos casos caracterizados por

contextos marcados por crisis económica. Las universidades de Alemania (Meier y Schimank, 2010) y Australia (Gläser y Laudel, 2007b) son otros ejemplos del mismo fenómeno.

Gläser y Laudel (2007a) señalan tres percepciones sobre los indicadores bibliométricos que pueden ser útiles para entender su amplio predominio en la evaluación de los resultados de la investigación. Los indicadores bibliométricos se perciben como: i) *un método más económico de evaluación*; ii) *legítimos* porque se construyen con base en productos de investigación que han sido juzgados y legitimados por otros pares, y en ese sentido, se perciben como *medidas objetivas, válidas y más confiables* que la revisión por pares; y, iii) *accesibles* para políticos y gestores de la ciencia —es decir, tienen menor dependencia en la evaluación del juicio informado de expertos o especialistas -de los pares-.

La evaluación por pares enfrenta algunos desafíos. Por un lado, los problemas prácticos que se derivan de la revisión detallada de una producción de conocimiento creciente, amplificadas por la excesiva demanda de evaluaciones. En ese sentido, la evaluación por pares del desempeño de los investigadores insume múltiples recursos. Como señalan Bianco, Gras y Sutz (2014), los procesos de evaluación de la investigación altamente reflexivos son demandantes, no solamente en términos económicos, sino que también requieren de tiempo y energía. Por otro lado, está la *subjetividad* de la evaluación por pares, eventualmente agravada, por ciertos comportamientos inapropiados de algunos evaluadores, ya sea por conflictos de interés y/o por la existencia de grupos de poder que influyen la revisión de las solicitudes (AFC, 2011).

Siguiendo a Sahel (2011), la evaluación por pares puede sesgarse por: i) los favoritismos disciplinares, locales o institucionales; ii) la falta de conocimiento específico de los evaluadores en el área de investigación objeto de evaluación; y, iii) la presión que ejerce una creciente cantidad de solicitudes de evaluación sobre los evaluadores, lo que, en algunos casos, pueden generar revisiones superficiales de las contribuciones de los investigadores (Sahel, 2011).

Estos desafíos en la revisión por pares no son ajenos al caso del SNI en México como tampoco lo son para el caso del RAE del Reino Unido. La implementación y uso de ciertos indicadores bibliométricos en las evaluaciones, que se realizan en esos dos sistemas, buscaban —de alguna manera- proveer una alternativa eficiente y objetiva para enfrentar dichos desafíos.<sup>17</sup>

Más allá de los desafíos que impone la evaluación por pares, la AFC (2011) reivindica su valor insustituible para apreciar la originalidad, calidad, relevancia, pertinencia y alcance o impacto de la investigación y así valorar, cualitativamente, los productos que de ella se derivan.

De acuerdo con Gläser y Laudel (2007a) las percepciones sobre la economicidad, validez y accesibilidad de los indicadores bibliométricos para los procesos de evaluación se sustentan en lo que esos autores denominan como *modalidades omitidas*<sup>18</sup>. Es decir, esas percepciones sobre los indicadores bibliométricos son tales porque no consideran, desconocen u omiten las condiciones bajo las cuales dichas medidas pueden ser creadas, aplicadas e interpretadas para alcanzar resultados de evaluación válidos y confiables. Los propietarios de las bases de datos bibliográficas,

---

<sup>17</sup> Para una descripción detallada del caso del SNI ver Bensusán et al. (2014), y del caso del RAE a Martin y Whitley (2010).

<sup>18</sup> El concepto de modalidad refiere a las condiciones bajo las cuales cierta información —ej.: indicador- puede ser creada, aplicada e interpretada para producir resultados válidos y confiables (Gläser y Laudel, 2007a).

políticos, gestores de la ciencia, académicos y evaluadores crean indicadores y/o toman decisiones y asignan recursos con base en la aplicación de métodos bibliométricos sin conocimiento calificado para ello, o cuyas implicancias metodológicas desconocen u omiten. Es por esta razón que también es altamente cuestionada la validez y confiabilidad de las fuentes sobre las cuales basan sus decisiones.

Gläser y Laudel (2007a) indican que los métodos bibliométricos son fáciles de implementar solo *aparentemente* y pueden dar lugar a resultados engañosos –o que proveen una imagen distorsionada de la realidad que buscan medir/reflejar- si no tienen en cuenta múltiples dimensiones. Simultáneamente destacan que a diferencia de lo que se percibe o supone, lograr indicadores bibliométricos robustos estadísticamente, válidos y confiables es costoso, requiere de tiempo, conocimiento experto, experiencia y disciplina.

En el artículo titulado “*The four literatures of social science*”<sup>19</sup>, (Hicks, 2004) ya había advertido las dificultades metodológicas que impone el uso de indicadores bibliométricos provistos por el ISI para evaluar el desempeño de la investigación en las ciencias sociales y humanidades. En la medida que la producción de conocimiento de esos campos trasciende largamente la producción de artículos científicos indexados en WoS, los indicadores bibliométricos que de allí se deriven, pueden dar lugar a una imagen distorsionada de su avance y desempeño, porque no son capaces de capturar su especificidad y diversidad. La autora además destaca que los propósitos de cada uno de esos tipos de literaturas, para difundir los resultados de investigación, son múltiples y están orientados a audiencias distintas. Además, resalta que los objetivos de investigación en ciencias sociales y las humanidades tienen una vinculación mucho más estrecha con su contexto geográfico, social y político, en relación a otras ciencias. (Hicks, 2004).

Posteriormente, Hicks (2006) reflexiona sobre los méritos del enfoque bibliométrico, examinando la naturaleza de esos *cuatro tipos de literatura* en relación a la evaluación del desempeño de la investigación centrada en los indicadores bibliométricos del SSCI-ISI. Concluye que sí es posible evaluar el desempeño de la investigación en ciencias sociales y humanidades a través de medidas bibliométricas. Sin embargo, sostiene que para hacerlo correctamente, se deben incluir todos los diferentes tipos de literatura que en ellos se producen, no solamente lo que se produce y publica en las revistas indexadas en WoS. En sintonía con Gläser y Laudel (2007a), remarca que eso supone un proceso complejo, desordenado y costoso (Hicks, 2006).

En esa dirección, Hicks et al. (2015) en el Manifiesto de Leiden, sostienen que la recolección de datos de alta calidad, precisos y rigurosos lleva tiempo y es costoso, en ese sentido recomiendan que los

---

<sup>19</sup> Los cuatro tipos de literatura son: i) artículos científicos en revistas internacionales; ii) libros; iii) artículos científicos en revistas nacionales; iv) publicaciones no académicas. En las ciencias sociales y en las humanidades además de producir artículos, se producen –con propósitos diversos y dirigidos audiencias específicas- libros, capítulos de libros, monografías, artículos para revistas académicas y no académicas nacionales, artículos de prensa, reportes para informar a la toma de decisiones de política pública, entre otros. Por ejemplo, mientras que publicaciones en las revistas académicas nacionales buscan difundir los resultados de investigación para enriquecer la discusión teórica, conceptual, metodológica y/o empírica con la comunidad académica local, las publicaciones en revistas no académicas buscan poner los resultados de investigación en interacción con los contextos de aplicación. (Hicks, 2004)



reponsables de esos datos deben asignar los recursos económicos suficientes para poder cumplir con los requerimientos de calidad de la información.

Una de las modalidades omitidas más frecuentemente es que ninguno de los indicadores basados en citas miden calidad, en todo caso aproximan un “*aspecto importante de la calidad*” su difusión, impacto o influencia. Más allá de que eso sea un parámetro importante para valorar el rendimiento de la investigación, los indicadores basados en citas son débiles y por ello es indispensable que sus resultados sean interpretados por los expertos en el campo. (Gläser y Laudel, 2007a)

Entre las limitaciones de los indicadores basados en citas, están: i) el número de citas depende de la calidad de la base de datos que se utilice; ii) puede suceder que los artículos sean muy citados por razones ajenas a su calidad (Kostoff, 1997); iii) los artículos de revisión en promedio reciben mayor cantidad de citas que los artículos originales de investigación; iv) los artículos publicados en revistas de alto impacto son mejor valorados que los de igual calidad pero publicados en revistas de menor visibilidad; v) los artículos escritos en inglés son en promedio más citados, lo que se ve reforzado por la cobertura de las bases de datos bibliográficas más utilizadas en la evaluación (ej.: *WoS*, *SCOPUS*), así artículos de calidad pero escritos en otros idiomas son invisibilizados; etc. (Sahel, 2011; AFC, 2011).

La AFC (2011) también advierte sobre el uso inadecuado del factor de impacto de las revistas. El factor de impacto de una revista no da cuenta de la calidad del artículo realizado por un investigador, sin embargo, y frecuentemente es utilizado para evaluar su calidad. Según la AFC esta práctica es ampliamente utilizada pero es peligrosa porque algunas revistas con alto factor de impacto contienen un porcentaje significativo de artículos de calidad media.

La Declaración de San Francisco es radical y sostiene que en ningún caso y bajo ninguna circunstancia debería utilizarse el factor de impacto como indicador aproximado de la calidad de una contribución científica (DORA, 2012). El factor de impacto no mide el número de citas de un artículo específico, ni refleja el interés particular que pueda tener algún artículo individualmente (Pendlebury; 2009; Moed 2009).

Entre las limitaciones más importantes del factor de impacto de una revista para evaluar la calidad de un artículo, están: i) la alta variabilidad en los factores de impacto entre los campos científicos, las revistas de ciencias básicas tienen factores de impacto en promedio más altos que las revistas en áreas muy especializadas o aplicadas; ii) las auto-citas no están corregidas en el cálculo del factor de impacto; iii) la información utilizada para el cálculo de factores de impacto no es transparente y no está disponible para el público; y, iv) los factores de impacto de las revistas pueden ser manipulados por políticas editoriales -además de que dependen de ellas- (Sahel, 2011; AFC, 2011; DORA, 2012). La editorial de Martin (2016) en *Research Policy* es especialmente ilustrativa sobre ciertas estrategias editoriales de algunas revistas orientadas a incrementar artificialmente su factor de impacto, lo que determina la pérdida de credibilidad de este indicador (Martin, 2016).

El Manifiesto de Leiden en sus cuarto y quinto principios destacan que los procesos de construcción de bases de datos para la evaluación y su análisis deben ser abiertos, transparentes, simples y tienen que estar disponibles para que los evaluados puedan verificarlos (Hicks et al., 2015). En la misma dirección la AFC (2011) y DORA (2012) sostienen que el proceso de evaluación tiene que ser transparente y replicable, con metas claras y ajustadas al contexto.

La AFC (2011) recomienda que, si se tienen en cuenta un conjunto importante de precauciones, la evaluación bibliométrica podría ser un insumo adicional y complementario, que contribuya a la evaluación cualitativa realizada por pares y a la toma de decisiones, pero en ningún caso debería ser el único insumo de evaluación. En esta misma dirección apuntan las recomendaciones de DORA (2012) cuando señala que la evaluación tiene que centrarse en los méritos intrínsecos de esa investigación, más que en la jerarquía o prestigio de la revista en la que se publican sus resultados (DORA, 2012). Asimismo, el Manifiesto de Leiden en su séptimo principio resalta que la evaluación del desempeño individual de los investigadores debe realizarse con base en la *valoración cualitativa de su portafolio de investigación*, y recomienda la adopción de un enfoque de evaluación que considere información diversa sobre cada investigador –incluidos sus conocimientos, experiencia, actividades e influencia-. (Hicks et al 2015)

A la vez, el primer principio del Manifiesto de Leiden, justamente destaca que la evaluación cuantitativa tiene que estar al servicio de –o apoyar– la valoración cualitativa realizada por expertos, y el segundo recomienda que *“la excelencia en investigación de relevancia local debe ser protegida”* (Hicks et al., 2015:3). Esta recomendación apunta directamente a evitar lo que sucede en muchos sistemas de evaluación que asimilan excelencia únicamente con publicaciones en revistas de alto impacto. Esa asimilación unívoca produce múltiples sesgos en los resultados de la evaluación y, dado que dichos factores de impacto se calculan para revistas indexadas en WoS, induce a que la investigación se centre en temas que no necesariamente tienen relevancia y pertinencia local y/o regional. (Hicks et al., 2015)

Otra modalidad frecuentemente omitida es que los indicadores basados en citas son válidos si son *estadísticamente representativos* del total de publicaciones. Los resultados que proveen los indicadores bibliométricos son válidos si se calculan sobre grandes cantidades de publicaciones. Si existe toda esa información y está disponible, entonces es posible medir un aspecto de la calidad de la investigación: su impacto internacional o rendimiento de la investigación. (Gläser y Laudel, 2007a)

A la vez y por lo anterior, es importante llamar la atención sobre la cobertura de las bases de datos bibliográficas que se utilizan para calcular dichos indicadores bibliométricos. Tal como señalan Gläser y Laudel (2007a), para que los resultados sobre el rendimiento de la investigación sean válidos, es preciso que los análisis bibliométricos de citas consideren todas las publicaciones de la unidad evaluada, de lo contrario se crean errores estadísticos significativos, especialmente si se trata de la evaluación de grupos de investigación o más aún, de individuos. Representando otra modalidad generalmente omitida. (Gläser y Laudel, 2007a)

Simultáneamente, las citas que representan influencia de los artículos publicados y, a partir de ello, eventualmente, un aspecto de su calidad, tienen lugar con al menos tres años de retardo. En este sentido, la validez de esos indicadores depende del período seleccionado para el análisis. Esta es otra modalidad habitualmente no considerada. (Gläser y Laudel, 2007a)

Similarmente, la AFC (2011) indica que para garantizar indicadores bibliométricos correctamente calculados es necesario acceder a múltiples bases de datos en su totalidad. Agrega que si bien las bases de datos existentes tienen calidad y crecen continuamente, su cobertura varía según las disciplinas, siendo las Ciencias Sociales y las Humanidades las peor representadas en ellas. Asimismo sostiene que no existe indicador o conjunto de indicadores capaces de resumir en soledad la calidad

de la producción científica de un investigador, no tienen ningún valor intrínseco por sí mismos, y deben ser analizados por expertos. (AFC, 2011)

Distintos autores han destacado que distintos modos de producción de conocimiento requieren distintos enfoques de evaluación (de Jong et al., 2011; Regeer et al., 2009; Spaapen, Dijstelbloem y Wamelink, 2007; Hemlin y Barlebo Rasmussen, 2006; Gibbons et al., 1994) para poder estimular la diversidad y la calidad en todas las áreas de conocimiento, atendiendo a sus diferencias, reconociendo y valorando todos sus resultados y productos de acuerdo a sus objetivos de investigación (Hicks et al., 2015, Bianco, Gras y Sutz, 2014; Hicks, 2013; de Jong et al., 2011).

En ese sentido, el Manifiesto de Leiden es claro cuando señala que los indicadores que se utilicen para medir el desempeño de la investigación tienen que estar estrechamente vinculados con los objetivos de esa investigación. Destaca que tales indicadores tienen que considerar los contextos en que se realiza la producción de conocimiento y sus propósitos, y que no hay un modelo de evaluación pasible de ser aplicado a todos los contextos. Así, el segundo principio del Manifiesto de Leiden dice: “[e]l desempeño debe ser medido de acuerdo con las misiones de investigación de la institución, grupo o investigador” (Hicks et al., 2015:2). Sin embargo, el estudio realizado por Bensusán et al. (2014), sobre el SNI de México, señala que la implementación de indicadores en la evaluación del desempeño de investigación de los individuos que se realiza en ese sistema, no atiende a las características contextuales en las que los investigadores desarrollan sus actividades de investigación.

La AFC (2011), DORA (2012) y Hicks et al. (2015) recomiendan considerar en la evaluación los diversos productos de la investigación. Mientras que DORA (2012) destaca la importancia de valorar el impacto de todos los productos de investigación no solo de las publicaciones científicas, a efectos de ponderar la influencia que tiene la producción de conocimiento en la política y en la práctica (DORA, 2012). La AFC (2011) es especialmente insistente con la importancia de considerar diversos indicadores para apreciar los impactos de la investigación orientada a la resolución de problemas complejos, ya que los mismos no siempre pueden ser valorados satisfactoriamente a través de los métodos bibliométricos. La importancia creciente de la bibliometría como método de evaluación en algunas disciplinas incluso puede estimular a algunos investigadores a re-orientar sus agendas de investigación y en consecuencia su producción de conocimiento hacia los temas preferidos o agendas de conocimiento de las revistas de alto factor de impacto (AFC, 2011).

De acuerdo con Gläser y Laudel (2007a), otra modalidad comúnmente omitida tiene que ver con la especificidad de la producción de conocimiento en los diferentes campos científicos. Como las prácticas de producción de conocimiento, publicación y citación varían entre los campos, los resultados derivados de los indicadores bibliométricos también varían según las diversas prácticas que se registren en los distintos campos de conocimiento. En ese sentido, sería incorrecto agregar resultados o realizar comparaciones entre campos, si los datos no son normalizados según los valores de referencia de cada campo específico (Gläser y Laudel, 2007a). Los resultados del cálculo de indicadores bibliométricos varían significativamente dependiendo de la base de datos bibliográfica que se utilice, y si su cálculo no es realizado por profesionales sino por “amateurs” (Gläser y Laudel, 2007a), puede dar lugar a una visión distorsionada y errónea de la calidad, impacto y desempeño de los investigadores.

Es por esto que la AFC (2011) es enfática en la importancia que tiene la validación y normalización de los datos que surgen de tales indicadores para realizar comparaciones entre investigadores, disciplinas y sub-disciplinas. La AFC (2011) señala las diferencias entre los tamaños de las comunidades científicas en cada disciplina y sub-disciplina, y destaca la necesidad de normalizar los indicadores de acuerdo con esos parámetros para que las comparaciones sean apropiadas.

En la misma dirección, el sexto principio del Manifiesto de Leiden resalta la importancia de considerar las diferencias en las prácticas de publicación y citación de los diversos campos científicos. Para la evaluación de los historiadores y los científicos sociales indica que se deben incluir y valorar la producción de libros y la literatura en el idioma local; en el caso de la evaluación de investigadores en informática hay que valorar también sus contribuciones en conferencias. Respecto a los indicadores basados en citas, también resalta la necesidad de indicadores normalizados por campos, las revistas más citadas en matemáticas tienen factores de impacto en el entorno de 3, mientras que en biología celular ese valor asciende a 30. (Hicks et al., 2015)

De acuerdo con Gläser y Laudel (2007a) los resultados de evaluación que se deriven de estas fuentes de información contienen múltiples errores que cuestionan su validez. Para los usuarios no expertos, la información proporcionada por las bases de datos bibliográficas y citas parece simple, accesible y directa porque omiten el tratamiento metodológico que requieren para asegurar la obtención de resultados válidos y confiables. El uso de indicadores bibliométricos con fines de evaluación proporciona resultados válidos y confiables en ciertos campos científicos y bajo el cumplimiento de ciertas condiciones metodológicas. Se perciben como accesibles y fáciles de entender porque se presentan como números, pero en realidad requieren de un tratamiento metodológico especializado sustantivo porque de lo contrario proveen resultados de evaluación que carecen de validez y pueden conducir a una visión distorsionada y errónea del desempeño de la investigación de universidades, campos científicos e investigadores. (Gläser y Laudel, 2007a)

Los últimos tres principios sobre buenas prácticas de evaluación de la investigación del Manifiesto de Leiden sostienen que: i) deben utilizarse múltiples indicadores –cualitativos y cuantitativos- para construir una imagen robusta y fiel del desempeño de la investigación. Sahel (2011) coincide con este principio. Si se trata de indicadores cuantitativos, sus valores deberían estar acompañados de sus respectivos errores de medición y predicción; ii) los indicadores deben revisarse y actualizarse periódicamente; y, iii) deben anticiparse los efectos sistémicos de la evaluación y los indicadores. Diversos indicadores son siempre preferibles, porque la implementación sistemática de uno o dos (por ejemplo: cantidad de publicaciones y de citas) además de todo lo dicho, puede inducir a comportamientos estratégicos que conduzcan a sustituir objetivos (por ejemplo: cuando alcanzar determinado valor en esos indicadores se vuelve un fin en sí mismo) y con ello alterar el contenido de la investigación, deteriorar la calidad, originalidad y pertinencia de la producción de conocimiento. (Hicks et al., 2015)

En la siguiente sección se revisa un tema central, que fue advertido en esta sección, el problema de la validez de los indicadores en general y de los bibliométricos en particular, y sus *consecuencias constitutivas* sobre la producción de conocimiento cuando se usan en la evaluación del desempeño de la investigación.

## IV.2 El problema de la validez de los indicadores y sus consecuencias constitutivas

Barré (2010) en su trabajo destaca que el proceso de construcción e interpretación de indicadores de ciencia y tecnología –incluidos los indicadores bibliométricos- es normativo y por lo tanto debatible. La construcción de indicadores refiere a la cuantificación. A la vez, dicha cuantificación no es otra cosa que un proceso de *naturaleza política* (Barré, 2010) porque se basa en una forma, modelo o configuración específica de comprensión de cómo las cosas funcionan o deberían funcionar. Es decir, los indicadores no son *verdad revelada o realidad técnica y científica* indiscutible. A la vez, su proceso de interpretación representa “*un ejercicio cargado de valores*”. Así, la construcción de indicadores y su interpretación no es más que la expresión de una representación sobre muchas otras representaciones posibles. (Barré, 2010)

De acuerdo a la sistematización realizada en la sección anterior, las deficiencias, tanto en el uso como en la interpretación de tales indicadores, son múltiples y pueden generar –debido a su importancia en la gobernanza de la ciencia- “*impactos desproporcionados e irreversibles, no sólo sobre grupos e instituciones particulares, sino también sobre las trayectorias cognitivas*” (Barré, 2010: 227-228) de la producción de conocimiento. Para contribuir a una comprensión analítica de estos impactos, Dahler-Larsen (2014) introduce el concepto *efectos constitutivos* de los indicadores.

Dahler-Larsen (2014), en coincidencia con Barré (2010), destaca que los indicadores responden a definiciones y representaciones particulares del fenómeno que buscan medir. Son una manera específica de *creación de sentido asistida* porque ofrecen claves para la interpretación, definen discursos y orientan acciones en direcciones específicas.

De ese modo, los indicadores son instrumentos de política y sus efectos también son políticos, en la medida en que definen categorías que son colectivamente significantes en una sociedad (Dahler-Larsen, 2014). Es decir, los indicadores son construcciones sociales (Gläser y Laudel, 2007a; Barré, 2010; Dahler-Larsen, 2014) y, también, construyen realidad (Dahler-Larsen, 2014).

De acuerdo con Dahler-Larsen (2014) los indicadores contribuyen a la definición operacional de una versión específica de un concepto, por ejemplo del concepto *calidad de la investigación*. Así el indicador por un lado, ordena lo que se entiende por *calidad de la investigación* y, por el otro, en la medida en que responde a una conceptualización específica, también puede ser controvertido y discutido (Dahler-Larsen, 2014).

Los indicadores producen una forma más o menos ordenada –estilizada- de la realidad (cuantifican, seleccionan, retienen, etc.). A la vez, proveen un lenguaje, y junto con la cuantificación, una categorización o taxonomización simplificada de una realidad o fenómeno mucho más complejo y dinámico. En dicha categorización se determina qué se va a considerar -actividades, insumos, procesos, resultados o productos- y para qué. Lo que se va a medir de un fenómeno no está dado, por el contrario, depende de un conjunto de operaciones –conceptuales, metodológicas y técnicas- para garantizar la validez y confiabilidad de esa medición. Al realizar esas operaciones, fenómenos ricos en significados se reducen más o menos dependiendo de la complejidad de la medición o, dicho de otro modo, de las diversas dimensiones del fenómeno que son retenidas en la medición. (Dahler-Larsen, 2014).

Conforme se retienen menor cantidad de dimensiones del fenómeno complejo que se busca medir, disminuye la validez de la medición. En el extremo, cuando sólo se retiene una única dimensión del

mismo, la relación entre la medición y ese fenómeno es baja. En ese caso, la medida no tiene validez porque mide otra cosa distinta de lo que buscaba medir: únicamente mide una de las múltiples dimensiones que caracterizan a ese fenómeno.

Cuando los indicadores que buscan medir desempeño -que vinculados a incentivos orientan la acción de individuos, grupos u organizaciones- carecen de validez, generan múltiples “*consecuencias no intencionadas*”. Más precisamente, como tales indicadores miden aspectos puntuales de un todo más amplio, producen *consecuencias constitutivas* (Dahler-Larsen, 2014). A medida que pasa el tiempo, y si se usan mediciones regularmente para categorizar individuos, grupos u organizaciones y/u otorgar reconocimiento, mayor prestigio o mejor acceso a recursos, que se basan en indicadores con escasa validez, se generan las condiciones para *la paradoja del rendimiento* (Dahler-Larsen, 2014). Así, el indicador se distorsiona -se vuelve un fin en sí mismo- y pierde significado.

Como ya se indicó, ningún indicador basado en citas mide calidad, en todo caso aproxima una de las diversas dimensiones importantes para determinar la calidad de la producción de conocimiento: su difusión o influencia internacional. Así, los indicadores bibliométricos basados en citas se relacionan débilmente con la calidad y, en ese sentido, carecen de validez. La *paradoja del rendimiento* en este caso -por ejemplo- se da cuando los investigadores en vez de preocuparse por la calidad sustantiva de su producción de conocimiento, se ocupan centralmente de maximizar el número de citas que sus artículos potencialmente pueden alcanzar.

El trabajo de Mårtensson et al. (2016) es ilustrativo de la multidimensionalidad y complejidad que encierran tanto la definición conceptual y operacional de la calidad de la investigación como también la construcción de una medida válida capaz de aproximarla. A la vez, el efecto performativo de esos indicadores probablemente sea un aumento en las citas, sin embargo, sus efectos constitutivos probablemente sean sobre las otras dimensiones de la calidad que tales indicadores no consideran. Siguiendo la conceptualización de calidad que proponen Mårtensson et al. (2016), entre esas dimensiones están: la credibilidad y alcance de los resultados de investigación, la relevancia conceptual, pertinencia y originalidad de la contribución, entre otras. Sobre estas dimensiones -probablemente- puedan observarse diversos efectos constitutivos de los indicadores bibliométricos basados en citas que carecen de validez para medir *calidad*.

Dahler-Larsen (2014) plantea que los indicadores en general y los bibliométricos en particular inciden en distintos aspectos en los que pueden observarse sus diversos efectos constitutivos. Siguiendo a ese autor y con el fin de ilustrarlos, a continuación se plantean -en *cursiva*- los distintos aspectos que son influenciados por los indicadores bibliométricos cuando son utilizados en la evaluación de la investigación, y se subrayan sus efectos constitutivos.

Los indicadores *definen marcos interpretativos y visiones del mundo* y cuando son utilizados en la evaluación del desempeño de la investigación, reducen la actividad de investigación a la producción de ciertas publicaciones (Dahler-Larsen, 2014) -preferentemente en revistas de alto impacto, de amplia circulación internacional, indexadas en determinadas bases de datos y en inglés-.

Los indicadores bibliométricos en la evaluación también *determinan contenidos y establecen períodos temporales*; el énfasis puesto en la obtención de cierta cantidad de artículos científicos presiona a los investigadores, quienes tenderán a dividir su trabajo en piezas publicables más pequeñas, focalizarlo en preguntas de investigación no riesgosas -para obtener resultados seguros

y rápidamente publicables- y orientarlo hacia temas de investigación de alto interés para las audiencias internacionales. (Dahler-Larsen, 2014)

En definitiva la productividad de la investigación y la calidad de la producción de conocimiento son conceptualizadas y medidas regularmente -aunque no sin controversia- en relación a la cantidad de artículos científicos publicados y su difusión y transcendencia internacional. Esto simultáneamente incide –directa o indirectamente, por acción u omisión- en la elección de los formatos y medios de difusión, el idioma y los temas a investigar. En última instancia, esos indicadores y su uso en el ámbito de los sistemas de evaluación del desempeño de la investigación, están definiendo qué ciencia se hace y para quiénes.

El uso de los indicadores bibliométricos en la evaluación del desempeño de los investigadores también *define las relaciones sociales y las identidades*, generando tensión y competencia entre los académicos. Los indicadores y las estadísticas apoyan la definición de poblaciones y sus identidades colectivas (Dahler-Larsen, 2014).

Así, al tiempo que estos indicadores jerarquizan cierta investigación y a los investigadores que cumplen con los parámetros de medición, marginan a la enseñanza y los docentes, y también a la investigación y los investigadores que no siguen esos parámetros de rendimiento. Es decir, son los indicadores implementados en los sistemas de evaluación los que establecen demarcaciones entre la población académica, entre quienes son los investigadores y quienes no lo son, o entre quienes merecen ser reconocidos como investigadores y quienes no, aunque todos produzcan conocimiento científico.

Finalmente, la magnitud y alcance de los diversos efectos performativos y constitutivos dependen del contexto específico en que se implementen y de las respuestas estratégicas de los diversos actores involucrados. (Dahler-Larsen, 2014).

A continuación y en base a una revisión bibliográfica, se presenta la evidencia reportada en la literatura –internacional y nacional- respecto a las consecuencias probablemente derivadas de los sistemas de evaluación de la investigación, fundamentalmente centrados en medidas bibliométricas, sobre la dirección, contenido y características epistemológicas de la producción de conocimiento. Se pone énfasis en los efectos constitutivos más que en los performativos de esos sistemas de evaluación.

### IV.3 Las relaciones entre la evaluación del desempeño de la investigación y la producción de conocimiento: los hallazgos de investigaciones pasadas

A pesar de su importancia para la ciencia y su futuro, los efectos constitutivos sobre la producción de conocimiento, que pueden derivarse de la implementación de sistemas de evaluación del desempeño de la investigación, han sido poco explorados empíricamente. Tal como señalan De Rijcke et al. (2016), la agenda de investigación que explora esas relaciones es relativamente nueva y si bien se conocen algunas de las implicancias de las evaluaciones cuantitativas, el conocimiento que se tiene en la actualidad todavía es escaso.

El análisis y la evidencia empírica que aporta esta tesis, busca contribuir a aumentar el conocimiento sobre las implicancias de la evaluación -predominantemente cuantitativa- del desempeño individual de investigación, en relación a dos formas diferenciales de producción de conocimiento. En

particular, los resultados de esta tesis pueden aportar evidencia nueva que sugiera efectos constitutivos de la evaluación del SNI sobre lo que Gibbons et al. (1994) conceptualizaron como Modo 1 y Modo 2 de producción de conocimiento.

En base a una revisión de la literatura que documenta diversas experiencias caracterizadas por la implementación y el uso sistemático de indicadores bibliométricos en la evaluación de la investigación, a continuación se plantean algunos de sus efectos constitutivos sobre la producción de conocimiento. En particular sobre ciertas características epistemológicas, tales como: la calidad, la originalidad, la relevancia conceptual, la pertinencia social y/o productiva, la multi/inter/transdisciplinariedad y los períodos temporales de la producción de conocimiento.

Bianco, Gras y Sutz (2014), en base a una revisión de la literatura y la reflexión sobre los procesos de evaluación académica, implementados fundamentalmente en la Universidad de la República de Uruguay, alertan sobre las posibles *consecuencias no anticipadas* de las políticas de estímulo a la investigación y sus instrumentos. En particular, las autoras llaman la atención sobre las señales “anti” -*anti-interdisciplina, anti-novedad, anti-pertinencia, anti-relevancia y anti-calidad*- para la producción de conocimiento, que pueden derivarse de los juicios emitidos en modelos de evaluación académica poco reflexivos y centrados en ciertas medidas bibliométricas.

En base a tres proyectos de investigación cualitativa realizados en Holanda y Austria, Müller y De Rijcke (2017) estudian los impactos de los indicadores cuantitativos, utilizados para evaluar el desempeño de la investigación, sobre las características epistemológicas de la producción de conocimiento en el campo de las *ciencias de la vida*. El material empírico del caso holandés fue construido con base en observación etnográfica y entrevistas semiestructuradas a investigadores (con diferentes grados de avance en su trayectorias académicas y nacionalidades), técnicos, líderes de grupos, directores de departamento y administradores –managers-, pertenecientes a tres grupos de investigación universitarios y dos centros médicos holandeses, que trabajan en los campos de biología celular molecular, oncología quirúrgica y estadística médica. El material empírico del caso austríaco también se basa en entrevistas semiestructuradas a estudiantes de doctorado y posdoctorado, investigadores jóvenes y *senior*, y a líderes de grupos de investigación. Todos ellos trabajan en el campo de las ciencias de la vida y desempeñan sus actividades en las principales universidades y organizaciones de investigación de Austria. (Müller y De Rijcke, 2017).

Sus resultados proveen evidencia que demuestra que la relevancia de la investigación que se realiza en el campo de las *ciencias de la vida*, es evaluada y definida según su rendimiento potencial en términos cuantitativos (por ejemplo: cantidad de publicaciones de alto impacto). A la vez, al tiempo que las métricas de rendimiento de la investigación son integradas a las prácticas de producción de conocimiento, las consideraciones sobre la calidad, originalidad, progreso científico a largo plazo, relevancia y responsabilidad social de la investigación reciben menos atención en la producción de conocimiento, en sus procesos de creación y en la evaluación. Las autoras resaltan que esos resultados obtenidos ponen en evidencia la tensión que tienen con los objetivos de la política pública que, buscan fomentar la investigación de calidad, conceptualmente relevante, original y socialmente pertinente. (Müller y De Rijcke, 2017)

De Rijcke et al. (2016) aportan una revisión bibliográfica actualizada de las principales tendencias empíricas sobre los posibles efectos de la evaluación de la investigación, de los indicadores y las respuestas estratégicas de los investigadores -*comunidades científicas*- y otros actores interesados.



De acuerdo con estos autores, la evaluación altera el comportamiento de los investigadores. Por un lado, induce a lo que De Rijcke et al. (2016) denominan como *desplazamiento de la meta*. Es decir, alcanzar altas puntuaciones en los criterios de evaluación se transforma en el objetivo y con ello, dichos criterios pierden su significado, ya no son un medio para evaluar el desempeño de la investigación, sino que se convierten en un fin en sí mismo. Por otro lado, ese desplazamiento de la meta induce a ciertas respuestas o *comportamientos estratégicos* de los investigadores, quienes evitarán la elección de temas de investigación riesgosos y las consideraciones que hacen sobre la calidad de la investigación, pierden importancia relativa cuando esos criterios están centrados en la obtención de ciertos niveles de productividad –aproximada por cantidad de publicaciones-. Los casos de los sistemas de evaluación de Australia, Dinamarca, Noruega y el Reino Unido aportan evidencia empírica de estos efectos. (De Rijcke et al, 2016)

Estos autores también reportan una tendencia empírica orientada a la *reducción de tareas*. En particular, destacan que los investigadores tienden a focalizar su actividad hacia cierto tipo de investigación y formato para la difusión de sus resultados. La centralidad que ha adquirido la producción de artículos científicos en inglés en la evaluación, ha inducido, por un lado, a prácticas imitativas de ese estilo de difusión por parte de las disciplinas que gozan de menor prestigio en el sistema de ciencia –los casos de las ciencias sociales y las humanidades son paradigmáticos-, y, por otro lado, a una estandarización de los tipos y productos de la investigación. (De Rijcke et al, 2016)

La evaluación del desempeño de la investigación al tiempo que ha jerarquizado la publicación de artículos, ha devaluado otros tipos de investigación, actividades y productos, a saber: i) la investigación orientada a la obtención de resultados y/o productos pertinentes para la política pública; ii) la investigación aplicada; iii) la investigación original o disruptiva con los enfoques de la corriente principal; iv) el desarrollo de proyectos de investigación “excepcionales” orientados a responder grandes preguntas; v) el trabajo de investigación colaborativo; y vi) las actividades de enseñanza, entre otras actividades académicas. Distintos estudios sobre los sistemas de evaluación de Australia y el Reino Unido proveen evidencia de estos efectos constitutivos. (De Rijcke et al, 2016)

El estudio de Butler (2003) muestra la *reacción* de los académicos australianos ante la implementación y uso -en 1995- de una medida de desempeño de la calidad de la investigación -centrada en la cantidad de publicaciones- para asignar recursos públicos a las universidades en Australia. A la vez, las universidades refuerzan ese mecanismo re-distribuyendo –entre sus departamentos y/o individuos- esos recursos con base en el mismo criterio. A partir de un análisis empírico de los datos del SCI, en el período comprendido entre 1985 y 2000, la autora comprueba, a partir del momento en que se implementa dicha medida de desempeño, un aumento significativo del número de publicaciones de los investigadores australianos. Sin embargo, y a pesar de este notable efecto performativo en todos los campos científicos, la autora sugiere que la calidad de la investigación fue sacrificada, porque esas publicaciones se concentraron en las revistas de menor impacto. (Butler, 2003)

Gläser y Laudel (2007b), al igual que Butler (2003), estudian el sistema de evaluación de la investigación con el que se asignan recursos para las universidades en Australia. Sin embargo, su objetivo es diferente. Buscan identificar los efectos epistémicos, derivados de la implementación de un sistema de evaluación fuerte como el australiano, sobre la dirección de la investigación. Para ello, diseñan un estudio de caso múltiple y realizan un análisis cualitativo comparado de tres tipos de

universidades australianas (las más antiguas, las que emergen con la reforma del sistema de evaluación en 1995 y las universidades tecnológicas) en seis campos científicos: i) matemáticas, bioquímica, física y geología (ciencias exactas y naturales); ii) ciencia política (ciencias sociales); e, iii) historia (humanidades). Los datos utilizados en el estudio combinan: i) análisis documental; ii) análisis bibliométrico; y iii) entrevistas semiestructuradas a los administradores –managers- de las universidades y los académicos.

Sus resultados muestran que el sistema de evaluación incide indirectamente sobre la dirección de la investigación, con diversos efectos y adaptaciones –organizacionales e individuales- como respuestas estratégicas a las presiones que impone el contexto en el que las universidades y los académicos desarrollan sus actividades. A continuación se presentan muy sintéticamente solo algunos de los resultados obtenidos por Gläser y Laudel (2007b). En particular los relacionados con la incidencia del sistema de evaluación fuerte en las respuestas estratégicas de los académicos, las características epistemológicas de la investigación y su contenido.

Los autores entre sus resultados reportan: i) aumentos en el número de publicaciones, en el número de publicaciones de autor único y en su impacto internacional; iii) abandono de líneas de investigación por su escaso éxito en los parámetros de medición de desempeño; iv) mayor orientación de la investigación hacia los temas de “moda” o hacia los temas que tienen mejores oportunidades de financiamiento; v) una eventual reducción de la relevancia científica de la producción de conocimiento debido a que se desarrollan más proyectos según el financiamiento disponible más que por su contribución científica; vi) reducción de los períodos temporales de la investigación, de los problemas abordados y de los objetivos de investigación, porque los recursos se otorgan a través de proyectos de corto plazo; vii) una eventual reducción de la calidad de la investigación debido a que la selección de los métodos de investigación y objetos empíricos son guiados más por su economicidad que por su adecuación para alcanzar los objetivos de investigación planteados, por lo que los resultados de la investigación serían menos rigurosos y confiables; y viii) un desplazamiento de las actividades docentes y recursos, desde la enseñanza hacia las actividades de investigación. (Gläser y Laudel, 2007b)

Wang, Veugelers y Stephan (2017) muestran como los métodos de evaluación de la investigación basados en indicadores que enfatizan las citas y el factor de impacto de las revistas, pueden dar lugar a resultados sesgados contra la investigación de alto riesgo y original.

Utilizando datos de citas y sus trayectorias temporales de todos los artículos publicados en 2001 e indexados en WoS, en todas las disciplinas cubiertas por esa base de datos bibliográfica, y los respectivos factores de impacto de las revistas, Wang, Veugelers y Stephan (2017) exploran la relación entre *novedad*, *riesgo* de la investigación e *impacto de citación*. Para ello, primero definen conceptualmente novedad de la investigación como una nueva combinación de conocimientos y la operacionalizan a través de la construcción de una medida *ex ante*, que aproxima empíricamente la novedad de la investigación como una combinatoria de las citas referenciadas en cada artículo. Una vez calculada la *novedad de la investigación* contenida en cada uno de esos artículos, los autores analizan y comparan los perfiles de citación de la *investigación novedosa* en relación a la *no novedosa* y los factores de impacto de las revistas en que se publica.

Sus resultados, por un lado, confirman la asociación entre la investigación más novedosa y la de mayor riesgo intelectual de los artículos (medido según la dispersión de las citas). Por otro lado, al

comparar los perfiles de citación e impacto de los artículos que contienen resultados de investigación novedosa *versus* no novedosa, muestran que: i) la investigación más novedosa requiere de períodos temporales más amplios para ser reconocida, si bien en el largo plazo la investigación novedosa tiene más probabilidades de ser más citada, en el corto plazo es la menos citada; ii) la investigación más novedosa es más citada –reconocida- en campos alejados o distintos del propio o del campo principal donde se originó; y, iii) la investigación más novedosa tiene menor probabilidad de ser publicada en las revistas con alto factor de impacto (Wang, Veugelers y Stephan, 2017).

Bianco, Gras y Sutz (2016) reflexionan teóricamente sobre la centralidad que ha ganado un *tipo universal de indicadores* para medir la productividad de la producción de conocimiento y sus efectos en el contenido de la investigación. En particular, subrayan que los sistemas de evaluación centrados en ese tipo de indicadores, envían señales negativas a los investigadores interesados en desarrollar investigaciones sobre agendas contextualizadas. Las autoras argumentan, con base en los estudios teóricos y empíricos documentados en la literatura y su experiencia en el desarrollo de políticas de estímulo a la investigación en la Universidad de la República de Uruguay, que la consolidación de esas prácticas de evaluación deterioran y desestimulan un tipo de producción de conocimiento –la que se ocupa de atender los problemas complejos (fundamentales o prácticos) de su contexto-, cuya relevancia y pertinencia social y económica es sustantiva y de gran importancia en el contexto de los países en desarrollo.

En base a una encuesta de alcance mundial a demógrafos de países desarrollados y en desarrollo<sup>20</sup> van Dalen y Henkens (2012) analizan la percepción que tienen los investigadores respecto a la *presión por publicar*, el grado en que dicha presión afecta el prestigio académico del que gozan y la orientación de la investigación que realizan. Sus resultados muestran que: i) la mayoría de los demógrafos de todos los países analizados, perciben una presión por publicar alta, pero en Estados Unidos, Australia, Canadá y Reino Unido la proporción de demógrafos que perciben una alta presión es significativamente mayor; ii) tienen mayor prestigio aquellos que publican más; iii) una creciente y cada vez más importante cantidad de publicaciones que no se citan, los autores también señalan que importa más cuan seguido se escribe y con quién, que lo que se escribe y para quién; vi) la presión por publicar conduce a los investigadores a estrategias que denominan como “*salami tactics*” –subdivisión de los temas de investigación en piezas más pequeñas y publicables rápidamente-; y, v) la presión por publicar se asocia negativamente con la orientación de la investigación hacia temas de la política pública nacional o temas de relevancia local y con la publicación en revistas de circulación nacional (excepto en el caso de Estados Unidos como es esperable).

El estudio realizado por Chavarro, Tang y Rafols (2017) es especialmente relevante para comprender el rol que desempeñan las revistas científicas, que denominan como *no principales* –“non-

---

<sup>20</sup> Los autores seleccionaron cuatro tipos de países y regiones relevantes en demografía: i) Estados Unidos (como líder científico en el campo); ii) Australia, Canadá, Reino Unido (como países anglosajones competidores en el campo); iii) Europa Occidental (excluido el Reino Unido); y iv) Africa, Asia, América Latina y Europa del Este (como competidores en economías emergentes y países en desarrollo). (van Dalen y Henkes, 2012)

*mainstream journals*<sup>21</sup>- en la comunicación científica. En particular, estos autores estudian las funciones de dichas revistas en tres campos científicos: ciencias agrícolas, negocios y administración, y química en Colombia. Para ello, en primer lugar exploran –bibliométricamente- las prácticas de publicación de los investigadores de esos campos de ese país en las revistas *principales* –aludiendo a que están indexadas en *WoS* y *SCOPUS* y por ello son reconocidas en los sistemas de evaluación dominantes-, y en las *no principales* –aludiendo a que están indexadas en Scielo y RedALyC<sup>22</sup> y por ello no son tan bien reconocidas en los sistemas de evaluación dominantes-. En segundo lugar, con la información obtenida sobre los patrones de publicación y validada por la revisión de *Currículum Vitae*, seleccionan un conjunto de investigadores en esos campos, a quienes les realizan entrevistas semiestructuradas. El objetivo de las entrevistas es profundizar sobre sus motivaciones para publicar.

Sus resultados muestran que contribuir al conocimiento y avanzar profesionalmente son las principales razones para publicar. El avance profesional se asocia con publicar en *revistas “top” –top journals-*. Respecto a las motivaciones para publicar en *revistas no principales*, sus resultados muestran que los investigadores le asignan tres grandes funciones a esas revistas. (Chavarro, Tang y Rafols, 2017).

Las *revistas no principales* son vistas por algunos investigadores como un espacio de aprendizaje y formación de capacidades. Es decir, es el paso previo para poder dar el “salto cualitativo” y publicar en las *revistas principales* y “avanzar profesionalmente”. Esta percepción se basa en la idea –expresada por algunos investigadores- de que “calidad” es sinónimo de *WoS*. Sin embargo, no todos los entrevistados convergen en esta visión o función. (Chavarro, Tang y Rafols, 2017)

Para otros, las *revistas no principales* permiten el acceso y la lectura de resultados de investigación a comunidades que no acceden a las *revistas principales*. Facilitan el acceso a materiales para la enseñanza –de grado y posgrado-, permiten el acceso libre a artículos de investigación, difunden el conocimiento escrito en inglés a otros idiomas –en este caso, a lectores de español y/o portugués-, y habilitan la introducción de conceptos, temas y métodos en la comunidad científica local. (Chavarro, Tang y Rafols, 2017)

Finalmente, la amplia mayoría coincide en que las *revistas no principales* son fundamentales para publicar sobre temas originales de investigación que no son cubiertos por las *revistas principales* o sobre temas cuyos abordajes teóricos, metodológicos y/o empíricos desafían a la corriente principal. A la vez, y cuando esos temas tienen pertinencia local porque atienden problemas específicos dependientes del contexto, la preferencia por publicar en *revistas no principales* es mayor. Esto se debe no solamente a que esos temas no tienen buena cobertura en las *revistas principales*, sino también a que permiten una más rápida y mejor difusión de los resultados de investigación hacia las audiencias locales relevantes. Además, algunos investigadores señalan que para mantenerse publicando en *revistas principales* tuvieron que transformar el contenido de su agenda de investigación, pasando de una agenda de investigación regional-local –contextualizada en términos de Bianco, Gras y Sutz (2016)- hacia una más internacionalizada (Chavarro, Tang y Rafols, 2017). Los

---

<sup>21</sup> Los autores cuando refieren a revistas científicas no principales aluden a aquellas revistas que no son bien ponderadas en los sistemas de evaluación por no estar indexadas en bases de datos internacionales tales como *WoS* o *SCOPUS*. Es decir, son revistas no principales para los sistemas de evaluación dominantes.

<sup>22</sup> Dos bases de datos bibliográficas líderes en América Latina (Chavarro, Tang y Rafols, 2017).

autores concluyen que la publicación en *revistas no principales* cumplen diversas funciones relevantes que las alejan de la pre-concepción que sostiene que allí se publican – fundamentalmente- trabajos de baja calidad científica. En ese sentido sostienen que la evaluación de la investigación debería otorgar mayor valor a estas publicaciones (Chavarro, Tang y Rafols, 2017).

Vessuri, Guédon y Cetto (2014) cuestionan que la competencia por prestigio y reputación, gestionada a través de la publicación en *revistas centrales* –referidas en esta tesis como revistas indexadas en bases bibliográficas internacionales tales como WoS o SCOPUS-, sea la única forma de demostrar la excelencia científica –individual, grupal u organizacional- y la calidad de la producción de conocimiento. Los autores argumentan que esa forma generalizada de competencia para alcanzar reconocimiento y prestigio, no contribuye mucho a mejorar la calidad general de la investigación en América Latina y otras regiones periféricas, ya que más que contribuir a enriquecer la investigación preocupada por atender los temas que afectan a su desarrollo social y económico, los niega y desplaza –re-orienta- la producción de conocimiento hacia temas que son de interés para los países centrales, sus revistas y la comunidad científica internacional. En definitiva, ese mecanismo termina estimulando un estilo único -o dominante- de producción de conocimiento internacionalizado y descontextualizado de su espacio regional, social, político, disciplinar e idiomático (Vessuri, Guédon y Cetto, 2014).

El objetivo del estudio de Rafols et al. (2012) es ilustrar como el uso de *rankings* de revistas en la evaluación del desempeño de la investigación puede poner en desventaja a la investigación interdisciplinaria. A través de datos de publicaciones y citas de WoS, y diversas métricas y técnicas de mapeo, el estudio compara el grado de interdisciplinariedad y desempeño de investigación de seis unidades organizacionales, tres de ellas en el área de los estudios de innovación y las restantes en administración y negocios, todas del Reino Unido. Sus resultados muestran que: i) la investigación en el área de los estudios de innovación es consistentemente más interdisciplinaria que en el área de la administración y negocios; ii) las revistas que ocupan las posiciones superiores en el *ranking* de la *Association of Business Schools* contienen un conjunto menos diverso de disciplinas, que las revistas ubicadas en las posiciones más bajas de ese ordenamiento; conduciendo iii) a una evaluación más favorable del desempeño de la investigación en el área de la administración y negocios, ya que la investigación en esa área es más disciplinar (Rafols et al., 2012).

Los autores concluyen con un conjunto de consecuencias probables sobre la producción de conocimiento. En primer lugar destacan que los resultados de la evaluación del desempeño de la investigación, que utiliza estos *rankings*, exhiben un sesgo sistemático a favor de la producción de conocimiento monodisciplinar. Así, la producción de conocimiento interdisciplinaria queda peor posicionada para obtener recursos para su investigación, y con el tiempo, esto puede traducirse en que los investigadores perciban menos incentivos a trabajar interdisciplinariamente, y por lo tanto, también menores posibilidades de trabajar en la búsqueda de soluciones a problemas complejos (científicos, productivos, sociales, ambientales, energéticos, etc.). Finalmente y en coincidencia con Whitley (2007), entre otros autores, señalan que es esperable que la producción de conocimiento se integre en torno a los temas de investigación disciplinar dominantes y se reduzca la diversidad cognitiva (Rafols et al., 2012).

El estudio de Hicks (2012) compara, a través de la identificación de todos los sistemas de asignación de recursos para la investigación basados en la evaluación de su desempeño -*Performance-based research funding systems (PRFS)*-, sus características específicas de operación, racionalidades, diseños y relación con los sistemas de financiamiento. Con ello y en base a una extensa revisión bibliográfica, busca extraer lecciones de la experiencia que se deriva de la implementación de estos sistemas de asignación de recursos, que puedan contribuir a enriquecer la comprensión sobre cómo evolucionan las políticas de investigación y los sistemas de innovación. Los sistemas de asignación de recursos para la investigación basados en el desempeño, son sistemas nacionales de evaluación de los resultados de la investigación utilizados para distribuir recursos para la investigación a las universidades<sup>23</sup>. La amplia difusión de estos sistemas representa un cambio importante en las políticas nacionales de investigación y también en los sistemas nacionales de innovación, debido a la centralidad que en estos últimos ocupan la investigación y las universidades. (Hicks, 2012)

La racionalidad que justifica la implementación de esos sistemas se vincula –aunque con énfasis diferentes en los distintos países- con las ideas proveniente del *new public management* para maximizar la eficiencia de la asignación de los recursos –escasos- para la investigación. Se basa en el supuesto de que la competencia por esos recursos conduce a la excelencia académica a través de estímulos orientados a incrementar la publicación internacional (en inglés). (Hicks, 2012)

Desde el punto de vista del diseño, las unidades de evaluación de esos sistemas también varían, pueden ser universidades, departamentos e individuos; y con ellos varía el énfasis en los distintos métodos de evaluación seleccionados. La evaluación a nivel de universidades se centra exclusivamente en indicadores bibliométricos (cantidad de publicaciones y/o citas) y su frecuencia es anual, excepto en los casos de Polonia y República Eslovaca que es cada cinco y tres años respectivamente. La evaluación a nivel de departamento teóricamente es realizada *por pares* en base a *indicadores bibliométricos*, y su frecuencia varía dependiendo del sistema. El sistema de Australia (ERA) es el más frecuente -cada dos años- y el RAE del Reino Unido el menos frecuente y se realiza cada siete años –antes se realizaba cada tres, cuatro y cinco años-. La evaluación individual se realiza en los sistemas de España y Nueva Zelanda, teóricamente a través de revisión por pares y cada seis años (Hicks, 2012).

---

<sup>23</sup> En base a un conjunto de criterios Hicks (2012) identifica catorce sistemas que han sido adoptados de forma generalizada desde 1986, cuando se implementó el *Research Assessment Exercise (RAE)* en el Reino Unido – en la actualidad, conocido como *Research Excellence Framework (REF)*-. Desde ese momento y hasta el 2010, identifica: i) *sexenio* desde 1989 en España; ii) *evaluación paramétrica* desde 1991 en Polonia; iii) sistema de evaluación desde 1992 en República de Eslovaca; iv) *RAE* desde 1993 en Hong Kong, China; v) *Research Quality Framework (RQF)/Excellence in Research for Australia (ERA)* desde 1995/2010; vi) *Research Unit Evaluation* desde 1996 en Portugal; vii) *Funding formula for allocation of university resources* desde 1998 en Finlandia; viii) *Valutazione Triennale della Ricerca (VTR)/Valutazione Quinquennale della Ricerca (VQR)* desde 2001 en Italia; ix) *Performance-based research funding (PBRF)* desde 2003 en Nueva Zelanda; x) *BOF-key* desde 2003 en Bélgica (Comunidad Flamenca); xi) *Norwegian model* desde 2006 en Noruega; xii) *New model for allocation of resources* desde 2009 en Suecia; y, xiii) *Norwegian model* desde 2012 en Dinamarca. (Hicks, 2012). Señala que todos estos sistemas pueden ser caracterizados como *sistemas de evaluación fuertes* de acuerdo a Whitley (2007).

Sin embargo, y siguiendo a Hicks (2012), en la práctica todas esas evaluaciones se hacen casi que exclusivamente a través de *indicadores bibliométricos* provistos por WoS o SCOPUS, lo que se constituye como el principal foco de los cuestionamientos que reciben estos sistemas de evaluación. Entre otras cosas, porque no consideran las especificidades de los productos y resultados de investigación de los distintos campos científicos. Para Hicks (2012), al igual que para otros autores, la contabilización de artículos científicos publicados en revistas indexadas en WoS o SCOPUS, puede ser un indicador que refleje adecuadamente el desempeño de las ciencias naturales pero no el de las ciencias sociales y de las humanidades.

Finalmente, Hicks (2012) concluye que los sistemas de evaluación basados en el desempeño, como instrumentos de política, tienen por objetivo promover la excelencia académica, sin embargo, lo logran sacrificando *valores públicos* importantes de las universidades, tales como la diversidad y la equidad. A la vez, estos sistemas, debido a sus diseños y las métricas que emplean, comprometen los objetivos de las políticas orientadas a estimular la innovación. Es decir, concentran los recursos para la investigación y a la vez, la producción de conocimiento se estandariza, adquiriendo el formato de artículos científicos en inglés. Sus objetivos y temas de investigación se integran en torno al interés de las audiencias internacionales. La contribución de las universidades a la comprensión de la realidad nacional y la identidad cultural probablemente se pierda por su devaluación ante los parámetros de estos sistemas. Por último, dichos sistemas tampoco contribuyen a fomentar la interacción de las universidades con la industria, ni promueven tipos de investigación aplicada. (Hicks, 2012)

En términos más generales, Nowotny, Scott y Gibbons (2001) destacan que si la evaluación de la investigación no considera las diversas especificidades de la producción de conocimiento, no la reconoce y, en ese sentido, puede castigar *formas de producción de conocimiento* que se caracterizan por su orientación hacia la atención de problemas complejos, transdisciplinarios, que involucran la participación de actores diversos y que ocurre en los *contextos de aplicación*.

Hicks (2013) destaca que no considerar la especificidad de la producción de conocimiento en ciencias sociales y humanidades en la evaluación, puede inducir en el largo plazo, a dos fenómenos. Por un lado, la presión por publicar internacionalmente y en inglés obligaría a los académicos de las ciencias sociales y las humanidades a adoptar las perspectivas de los académicos estadounidenses, líderes y dominantes en WoS. Además, agrega “[s]i este fuera el caso, aquellos que piensen sobre el futuro de una sociedad lo harán pensando en términos Americanos” (Hicks, 2013:86). Por otro lado, se reduciría el impacto social de la producción de conocimiento en ciencias sociales y humanidades. En particular, en términos de su difusión hacia y pertinencia para los actores no académicos (hacedores de políticas públicas o público en general) (Hicks, 2013).

El estudio de Das et al. (2009) aporta evidencia cuantitativa sólida en el campo de la economía que, de alguna manera, da cuenta del primer fenómeno advertido por Hicks (2013). A través de la construcción de una base de datos que contiene información de 76046 artículos empíricos en economía, publicados entre 1985 y 2004 en las revistas consideradas más importantes de ese campo (202). Los resultados del trabajo muestran –entre otras cosas- que los artículos sobre Estados Unidos tienen una probabilidad significativamente más alta -de 2.6 puntos porcentuales- de ser

publicados en las 5 revistas de economía<sup>24</sup> mejor posicionadas, que aquellos enfocados en el estudio de otros países. Importa señalar que este resultado se mantiene incluso después de haber controlado por la calidad de los artículos<sup>25</sup>. Más aún, los autores destacan que ese es un efecto muy grande e importante, porque únicamente el 1.5% de todos los artículos que estudian aspectos económicos de la realidad de otros países -distintos de Estados Unidos-, se publican en esas cinco revistas “*top-five journal*”.

Hammarfelt y De Rijcke (2014) muestran una tendencia de cambio sobre las estrategias de publicación en el campo de las humanidades en la Universidad de Uppsala en Suecia. Los autores asocian esa tendencia como una consecuencia probablemente derivada de la presión externa que impuso la introducción, en el año 2009 y a nivel nacional, de un *nuevo sistema de asignación de recursos basado en la evaluación del desempeño de investigación*. El estudio se realiza con base en la combinación de datos bibliométricos y, datos cuantitativos y cualitativos. Los datos bibliométricos fueron extraídos para el período 2006-2013 de la base de datos bibliográfica DiVA -*Digital Archive Online*-, que contiene todos los tipos de publicaciones y, con ello, evitar los sesgos introducidos por WoS o SCOPUS –es decir, sobre-representación de artículos escritos en inglés-. Los datos cuantitativos y cualitativos fueron recogidos a través de un cuestionario, que fue enviado a todos los académicos (incluidos estudiantes de doctorado, profesores e investigadores) en el campo de las humanidades de la Facultad de Artes de esa universidad. Debido a su escasa tasa de respuesta (26%), el análisis estadístico es representativo a nivel de los departamentos pero no a nivel disciplinar.

Sus resultados muestran que a partir de la introducción del nuevo sistema de evaluación: i) aumentan las publicaciones en inglés y disminuyen las publicaciones en sueco; ii) aumentan las publicaciones arbitradas; y, iii) disminuyen las *monografías*, tanto escritas en inglés como en sueco. A pesar de que todos los académicos en el campo destacan que este tipo de producto es fundamental para las humanidades, sostienen que no hay incentivos suficientes para seguir generándolas porque son muy poco valoradas y ponderadas en la evaluación. Desde la perspectiva cualitativa, los autores resaltan que los académicos cada vez están más preocupados por las medidas bibliométricas y buscan publicar en revistas internacionales mejor posicionadas en los *rankings*, adaptando sus estrategias de publicación para imitar las prácticas de otras disciplinas ajenas al campo de las humanidades. De ese modo los académicos esperan mejorar sus posibilidades de competencia y mantener sus posiciones dentro de su campo. Finalmente, los autores dan cuenta del comportamiento de los más jóvenes, quienes muestran una tendencia más inclinada a seguir las presiones externas, focalizando su producción de conocimiento hacia la publicación en revistas internacionales. (Hammarfelt y De Rijcke, 2014)

El estudio realizado por Leisyte y Westerheijden (2014) explora y compara las implicancias de dos sistemas de evaluación de la investigación sobre el “*corazón*” del trabajo académico, más específicamente, sobre las prácticas de investigación. Para ello analizan cualitativamente, con base

---

<sup>24</sup> En las cinco primeras posiciones del *ranking* utilizado por estos autores están: i) *American Economic Review*; ii) *Econometrica*; iii) *Journal of Political Economy*; iv) *Quarterly Journal of Economics*; y v) *Review of Economic Studies* (Das et al., 2009).

<sup>25</sup> Ver Das et al. (2009) para una discusión detallada del enfoque metodológico y métodos empleados para controlar los resultados según la calidad de los artículos.



en la revisión de la literatura, el RAE del Reino Unido y el modelo de evaluación de la investigación de Holanda.

Antes de describir tales implicancias, conviene describir brevemente –siguiendo a estos autores– algunas de las características del sistema de evaluación holandés. Algunas de las características del RAE ya han sido descritas a lo largo de este capítulo y en esta sección. De todas maneras, para quienes quieran profundizar más sobre ellas, pueden consultar –entre otros estudios– el trabajo de Martin y Whitley (2010) que, además de describirlo detalladamente, presentan su evolución a lo largo del tiempo desde su creación.

El modelo de evaluación holandés se implementa desde 1983, tres años antes que el RAE del Reino Unido, pero a diferencia de éste, no tiene consecuencias directas sobre la asignación de recursos públicos para la investigación. Sin embargo, sí tiene consecuencias indirectas. En particular, porque los encargados de redistribuir los recursos para la investigación al interior de las universidades en Holanda, utilizan los resultados de esas evaluaciones para hacerlo. Más adelante se describe otro mecanismo, también indirecto, que da cuenta de cómo el sistema de evaluación holandés afecta la asignación de recursos –en ese caso, externos– para la investigación. (Leisyte y Westerheijden, 2014)

El objetivo de política que persigue el modelo de evaluación holandés es estimular la “buena ciencia”. Desde el punto de vista de su diseño: i) se organiza en *comités de pares*, cada comité se asocia con grandes áreas de conocimiento; ii) estos comités juzgan las actividades de investigación de *grupos*, la unidad de evaluación son los grupos de investigación; iii) los juicios emitidos por dichos comités se realizan con base en diversas fuentes (información provista por el grupo, incluyendo sus principales productos, entrevistas a los líderes de grupos y otros integrantes, e indicadores bibliométricos); y, iv) la evaluación de los grupos se realiza cada seis años. Los resultados de la evaluación son públicos. Si bien no tiene el énfasis en el *new public management* que sí tiene el RAE (ver Martin y Whitley, 2010), sí cumple la función de *transparencia y rendición de cuentas*. Finalmente, los juicios sobre los grupos se elaboran con base a escalas de cinco puntos sobre cuatro dimensiones, a saber: i) *productividad* (cantidad de publicaciones y títulos de doctorado); ii) *calidad* (entendida como “reconocimiento internacional, potencial innovador y entusiasmo”, pero está fundamentalmente asociada al impacto de las publicaciones en términos de las citas que reciben y factores de impacto de los órganos de publicación); iii) *relevancia* (entendida como “impacto científico y socio-económico”, que es la dimensión más subjetiva, aunque han desarrollado algunos procedimientos e indicadores (van der Meulen, 1995; citado en Leisyte y Westerheijden, 2014); y, iv) *viabilidad de largo plazo*, que se refiere a una estimación prospectiva sobre si el grupo seguirá siendo productivo, con alto nivel de calidad y relevancia. (Leisyte y Westerheijden, 2014)

Leisyte y Westerheijden (2014) también analizan el RAE del Reino Unido, entre las consecuencias que reportan del RAE están: i) aumento de la competencia entre los académicos por el acceso a recursos, de la estratificación de organizaciones e investigadores, y de la concentración de recursos –el ya aludido *Efecto Mateo*–, donde hay “*ganadores y perdedores*” –los *ganadores* son los que sí pueden afrontar la investigación –los investigadores– y los *perdedores* son los que dan clases –los docentes–; ii) aumento de los estados emocionales de ansiedad de los investigadores y grupos de investigación ante cada edición de evaluación del RAE; iii) aumento de la presión sobre la realización de *investigaciones publicables internacionalmente en revistas de alto factor de impacto*; iv) aumento de la carga, ritmos de trabajo y presión por conseguir recursos externos para la investigación, lo que

ha inducido al desarrollo de estrategias para afrontar estos desafíos, basadas en la tendencia a la colaboración entre los investigadores; v) reforzamiento de los estándares de excelencia disciplinares; vi) aumento de la productividad –aproximada por cantidad de artículos científicos publicados–; vii) transformación en las formas en que los académicos reflexionan sobre su trabajo, especialmente en las humanidades y ciencias sociales; viii) reducción de los períodos temporales de la investigación, gestionada a través de proyectos y ajustada a la obtención de resultados publicables para cumplir con los parámetros de evaluación; ix) reducción de la producción de libros debido a que requieren períodos de tiempo más extensos de trabajo; x) aumento de la presión sobre los estudiantes de doctorado para que realicen “*investigaciones naturalmente publicables*” (National Postgraduate Committee, 1996 citado en Leisyte y Westerheijden, 2014: 13); xi) la presión por publicar condujo a prácticas “*salami publishing*” –división de los temas de investigación en unidades más pequeñas para maximizar la cantidad de publicaciones- o a publicaciones prematuras (Martin y Whitley, 2010); xii) mayor homogeneización –estandarización- de la investigación que se expresa por un lado, a través de la presión sobre todos los temas y campos de investigación para que adopten –imiten- el modelo de difusión de los resultados de investigación de las ciencias naturales y exactas, y por otro lado, en la concentración en pocos y específicos temas de investigación estrechamente definidos y que habiliten la obtención rápida de resultados, internacionalmente publicables en las revistas con mayor factor de impacto; y xiii) aumento del rendimiento de la investigación universitaria en términos de su incidencia internacional. (Leisyte y Westerheijden, 2014)

El trabajo de Martin y Whitley (2010) también muestra estos efectos y agrega –entre otras cosas- que el contenido de la producción de conocimiento es menos original y se orienta más por la investigación académica que por la investigación aplicada. Además, destacan que los procesos de investigación se caracterizan por ser menos riesgosos y ambiciosos, más disciplinares que interdisciplinares, y están más alineados con los temas, enfoques y métodos de la corriente principal (Martin y Whitley, 2010). De acuerdo con estos autores, el avance del conocimiento es más incremental y se concentra o integra entorno a los parámetros de la *ciencia normal* (Kuhn, 2012 [1962]). Tal como señalan Leisyte y Westerheijden (2014), algunas de las consecuencias reportadas para el caso del RAE son coincidentes con las que reportan para el caso del modelo de evaluación holandés.

Entre las implicancias para el caso del modelo de evaluación holandés Leisyte y Westerheijden (2014) destacan que los investigadores publican más estratégicamente, y el objetivo de tales estrategias es maximizar el número de publicaciones y la cantidad de citas para posicionarse mejor en los índices bibliométricos. Al igual que en el caso del RAE y de los demógrafos (van Dalen y Henkens, 2012) ya referenciado, aplican “*salami publishing*”.

En términos del desempeño, al igual que en todos los otros sistemas de evaluación, el modelo de evaluación holandés también contribuyó al aumento de la cantidad de artículos científicos publicados internacionalmente (Leisyte y Westerheijden, 2014). Se producen más artículos de forma cada vez más acelerada. En términos constitutivos, se acortan los períodos temporales de la investigación. Las reglas orientadoras son: i) *dos publicaciones por año*; ii) *elija temas de investigación rápidamente publicables internacionalmente*; y, iii) *primero publique artículos en revistas con alto factor de impacto, luego –si le da el tiempo- publique libros en editoriales de prestigio*.

En el caso de las *ciencias de la vida* el objetivo es publicar en revistas con alto factor de impacto y los indicadores bibliométricos basados en citas son muy centralmente tenidos en cuenta. En este campo, esta es y ha sido históricamente su práctica habitual de producción y difusión de conocimiento. Sin embargo, en el campo de la *historia*, los historiadores buscan un balance entre sus prácticas tradicionales de producción y difusión del conocimiento (libros) y el formato exigido (artículos). A la vez, los autores reportan que entre ellos se observan altos niveles de ansiedad debido al incremento de la competencia por publicar, por la obtención de los “puntos” que eso genera en la evaluación de desempeño holandesa, y por la presión que genera el dominio – creciente- de los formatos de difusión y producción de conocimiento de las “*ciencias duras*”. También y al igual que en el caso del RAE y de la evaluación de desempeño de Australia (Gläser y Laudel, 2007b) –por dar algunos ejemplos-, los investigadores de todos los campos en Holanda perciben, pero los historiadores más, que disponen de períodos cada vez más acotados para desarrollar sus investigaciones (Leisyte y Westerheijden, 2014).

La presión que perciben los investigadores sobre la cantidad de artículos que deben producir, más que inducirlos a realizar más investigaciones conceptualmente relevantes, socialmente pertinentes y/o más originales, los conduce a lo que Leisyte y Westerheijden (2014) elegantemente denominan como “*repackaging of ideas*” pero que refiere a un fenómeno que, en el *argot* académico mexicano, se conoce como “*refrito*”.

Finalmente, la ambición profesional de los investigadores converge en alcanzar grados más altos en la evaluación holandesa del desempeño de la investigación, porque además de generar prestigio eso afecta –indirectamente- el acceso a recursos. Leisyte y Westerheijden (2014) también reportan un fenómeno que lo asocian con los “*efectos halo*”, para referirse a cuando los resultados de la evaluación del desempeño influyen la obtención de plazas en el mercado de trabajo de los académicos, o el acceso a recursos externos para la investigación (por ejemplo a través de fondos competitivos por proyectos). Este es el otro mecanismo indirecto -mencionado más arriba- a través del cual el modelo de evaluación de desempeño holandés afecta la asignación de recursos externos para la investigación.

Muchas de las consecuencias constitutivas hasta aquí revisadas, probablemente derivadas de la implementación de sistemas de evaluación fuertes, sobre el contenido, la dirección y características epistemológicas de la producción de conocimiento, también han sido reportadas por diversos estudios en relación al caso del SNI de México. Estos estudios, desde diversos enfoques, han analizado distintos aspectos del SNI y su incidencia –probable- en las actividades académicas.

Paulatinamente, el SNI se ha convertido en un instrumento de política pública de CTI de México muy relevante, o el más importante dentro de los programas públicos de investigación (Reyes y Suriñach, 2015). Es percibido como un importante dispositivo de prestigio en el ámbito académico mexicano y ha desempeñado un papel central como “vector de internacionalización” de sus élites científicas (Didou y Gérard, 2011 y 2010), aumentando el reconocimiento, prestigio y profesionalización de la actividad de investigación (Bensusán et al., 2014; Valenti et al., 2013; Esteinou, 2012).

Diversos estudios describen la contribución del SNI al aumento de los recursos humanos altamente calificados dedicados a la actividad de investigación y al incremento de la participación de la producción de conocimiento nacional en el ámbito internacional (Bensusán et al., 2014; Valenti et

al., 2013, entre otros). Específicamente, a través de una cantidad creciente de artículos científicos especializados en revistas científicas internacionales de alto impacto (Esteinou, 2012).

Estudios anteriores, como por ejemplo los de Gonzalez-Brambila y Veloso (2007) y Gonzalez-Brambila (2005), ya habían analizado y destacado la incidencia del SNI sobre la productividad y el impacto internacional de la producción de los investigadores categorizados en ese sistema, en todas las áreas de conocimiento. Sin embargo, Gonzalez-Brambila (2005) demostró la existencia de diferencias significativas en la productividad según las diversas áreas.

La autora destaca que esos resultados sugieren que los investigadores de las distintas áreas tienen incentivos diferenciales para publicar en las revistas internacionales indexadas en WoS. En particular, señala que mientras que los investigadores de Ciencias Exactas, Biología, Química y Ciencias de la Salud publican más, por más años y desde que son más jóvenes; los de Ciencias Sociales, Humanidades, Ciencias Agrícolas, Biotecnología e Ingeniería eligen temas de investigación con intereses más locales, cuyos resultados tienden a difundirse en otros medios, menos globales y más asociados a las audiencias locales/regionales. La autora reconoce que los artículos científicos no son el único producto de investigación relevante, y que para lograr resultados más completos sobre la productividad de los investigadores, también hay que analizar otros productos, tales como: libros y capítulos, reportes para informar la toma de decisiones -por ejemplo en el ámbito de la política pública-, desarrollos tecnológicos, entre otros. (Gonzalez-Brambila, 2005).

Consistentemente, el estudio de Russell y Liberman (2002), con base en datos bibliométricos para el período 1997-2000, obtenidos de bases de datos bibliográficas, una internacional (WoS) y dos regionales (Periódica y Clase), de los trabajos publicados por investigadores de la UNAM en cinco grandes áreas de conocimiento, muestran que: i) para el caso de las ciencias naturales y exactas, el medio de difusión de mayor relevancia para sus resultados de investigación es a través de artículos científicos publicados en revistas internacionales; ii) en el caso de las ciencias sociales y las humanidades, el principal medio de difusión de sus resultados de investigación son los libros y los capítulos de libro, seguidos de la publicación de artículos en revistas de circulación regional; y que, iii) las ciencias aplicadas -ingenierías y biotecnología- se encuentran en un punto intermedio entre esos dos extremos. Los autores reflexionan sobre los modelos diferenciales de comunicación de la producción científica de las diversas áreas y concluyen que esas diferencias reflejan sus patrones de investigación también variantes.

Más allá de todos estos importantes efectos sobre el desempeño *-performativos-* de los investigadores y sus investigaciones, probablemente como consecuencia del SNI, los estudios también revelan un conjunto de consecuencias “perversas” *-efectos constitutivos-* e incluso contradictorias con los objetivos de política pública.

Izquierdo (2006) destaca que el SNI ha tenido gran incidencia en el cambio de los roles y valores académicos, generando diversas exclusiones. En particular, y entre otras cosas, la autora destaca: *“Es clara la intervención de las políticas federales en materia de financiamiento a la investigación y el acoplamiento que los académicos tienen que hacer para su ingreso y permanencia en el sistema (...) [produciendo] efectos contradictorios [con] el resultado [esperado por el] (...) gobierno federal para el desarrollo de la ciencia en México. (...) [P]or un lado se exige calidad pero se evalúa cantidad. Se pide investigación ambiciosa y de frontera pero se premia a la investigación modesta y de corto plazo. (...) [P]or otro lado, se requiere la formación de recursos humanos para el desarrollo de la*

*ciencia pero, se evalúa principalmente la cantidad de tesis que el investigador reporta ante el SNI, importando muy poco la calidad y el compromiso con el proceso de formación de los estudiantes.”* (Izquierdo, 2006:26, énfasis añadido)

Similarmente, el estudio de Ocampo y Rueda (2015) reporta cambios en las prácticas de los académicos de la Universidad Veracruzana tanto para ingresar y mantenerse en el SNI, como para promoverse –avanzar en su trayectoria académica- dentro de la universidad. En la misma dirección, el estudio de Álvarez (1999) documenta las respuestas de los académicos en ciencias sociales en contextos institucionales de precariedad, ante las oportunidades y limitaciones que se derivan de los cambios en las políticas de educación superior e investigación del gobierno federal.

El estudio de Heras (2005), entre otras cosas, y con base en varios autores -Cicero (2002), Drucker (2004), Díaz-Barriga (1996) y García Salord (2001)-, destaca diversos “efectos perversos” del SNI. Entre esos efectos, señala la reluctancia de los académicos a impartir clases y otras actividades de enseñanza en el nivel de la licenciatura, y una tendencia a concentrarse solamente en aquellas actividades que son reconocidas y premiadas por el sistema de evaluación, dejando de lado también las actividades de construcción institucional.

El desplazamiento desde las actividades de enseñanza y docencia hacia las actividades de investigación como consecuencia probable del SNI, ha sido un fenómeno muy estudiado y reportado en múltiples trabajos (Galaz Fontes et al., 2012; Padilla, 2010; Galaz Fontes, 2008; Galaz Fontes et al., 2008; Cordero, Galaz Fontes y Sevilla, 2003; Gil Antón, 2002; Ibarra, 1998 y 1999; Grediaga Kuri, 1998).

Algunos de esos estudios también destacan que el SNI genera múltiples dificultades para que *el académico mexicano* se oriente hacia actividades de vinculación con actores externos al ámbito de las universidades (Galaz Fontes y Gil Antón, 2009; Galaz Fontes et al., 2008; Grediaga Kuri, 1998) o hacia el desarrollo de investigaciones aplicadas (Valenti et al., 2013; Valenti, 2011) o las basadas en colaboraciones universidad-empresa (Amaro, Corona y Soria, 2008, Becerra, 2014). En la misma línea, el trabajo de Padilla (2010) también muestra el desplazamiento de las actividades orientadas al desarrollo tecnológico.

Por su parte, el estudio de Didou (2002) llama la atención sobre el escaso acceso que tienen los institutos tecnológicos federales al SNI y a otros programas públicos de estímulo a la investigación. El trabajo de Reyes y Suriñach (2012), a través de un análisis de la evolución de los criterios de evaluación del SNI y sus resultados, sugiere una baja coherencia entre ellos cuando se trata del desarrollo tecnológico.

El estudio de Bensusán et al. (2014), plantea diversas dificultades que enfrentan los investigadores del SNI para orientar su producción de conocimiento hacia la atención de problemas complejos dependientes del contexto, ya sean sociales, productivos o de política pública. Esteinou (2012) destaca que el SNI si bien reconoce la investigación internacionalizada, desconoce la producción de conocimiento orientada a atender los principales desafíos que impone el desarrollo social y económico de México.

El estudio de Rivera et al. (2011) señala que la cantidad de artículos publicados en revistas científicas como medida única para reflejar la productividad de la investigación, es un indicador muy pobre basado en una conceptualización muy estrecha de lo que realmente es la producción de

conocimiento y la diversidad de productos que genera. En particular, Rivera et al (2011) destacan para el caso de las ciencias agrícolas, que los investigadores generan variedad de productos, algunos de ellos orientados a mover la “frontera del conocimiento científico” y otros más orientados por la búsqueda de soluciones a problemas productivos, tales como la generación de nuevas recomendaciones y nuevas técnicas. Sin embargo, señalan que el SNI estimula la producción de artículos científicos pero no así otros productos técnicos. De acuerdo con los hallazgos reportados por estos autores, ese escaso o nulo reconocimiento de ese tipo de productos, es un desincentivo importante para la interacción entre la *academia y el sector productivo* nacional.

El estudio de Heras (2005) señala que entre los efectos del SNI también están: i) el escaso o nulo interés de los académicos mexicanos de publicar en revistas nacionales; ii) un cierto aumento en sus estados emocionales de ansiedad y de frustración “*si no entran al sistema*”; iii) cierto grado de malestar generalizado por la “*abierta competencia entre colegas*”; y, iv) una tendencia a las *estrategias de simulación* sobre los avances académicos -*repackaging of ideas y/o salami publishing* para ponerlo en los términos ya aludidos- (Heras, 2005:211). El trabajo de Bensusán et al. (2014), entre sus resultados, también documenta estos efectos.

Desde la perspectiva de diversos autores, el SNI ha tenido gran incidencia en las condiciones laborales de los académicos en México (Tamez y Pérez, 2009), en la mejora de los ingresos extrasalariales de los investigadores categorizados en ese sistema (Heras, 2005), en sus decisiones de jubilación –postergada- (Bensusán, Ahumada e Inclán, 2013) y en el estímulo al trabajo individual, basado en un esquema meritocrático desvinculado del contexto institucional (Díaz-Barriga, 2005).

Finalmente, el estudio de Bensusán et al. (2014), entre sus resultados, también reporta que el SNI impone restricciones para el desarrollo de estrategias de investigación innovadoras (originales) y de largo plazo. A la vez, señala que el SNI estimula la producción de conocimiento especializada, internacionalizada, acelerada e individualizada, y que tiene escasa apertura para valorar y reconocer la producción de conocimiento multi, inter o transdisciplinaria. Respecto a los criterios de evaluación del SNI, sus resultados sugieren: i) un escaso reconocimiento de los distintos perfiles de investigación en México; y, ii) una gran tensión debido al escaso balance entre los criterios de evaluación cuantitativos y cualitativos, siendo los primeros los dominantes, y para el caso de algunas áreas, por la imposición de estándares de calidad propios de las ciencias exactas y naturales y que le son ajenos a sus prácticas de investigación. Importa resaltar que el estudio de Bensusán et al. (2014) es el antecedente de investigación más directo de esta tesis, no solamente porque quien escribe participó activamente en el equipo de investigación, particularmente en el diseño y análisis de la encuesta a investigadores realizada en ese estudio, sino también porque esta investigación utiliza los datos contenidos en dicha encuesta.

#### IV.4 Conclusión: la relación entre la evaluación del desempeño de la investigación basada en indicadores y la dirección, contenido y características epistemológicas de la producción de conocimiento

De acuerdo con la revisión de los hallazgos de investigaciones pasadas reportados en la literatura y el debate internacional, en la actualidad existe una amplia difusión de los indicadores bibliométricos en la gobernanza de la ciencia, es decir, como criterios centrales de evaluación, con implicancias importantes tanto sobre la dirección y contenido de la producción de conocimiento, como también sobre las características epistemológicas de sus procesos de creación. Así, estos hallazgos sostienen,

con la evidencia que reportan, las premisas de la teoría que vincula a la evaluación del desempeño de la investigación con la producción de conocimiento revisada en el capítulo anterior.

El debate internacional -revisado en la sección IV.1 de este capítulo-, en torno a la amplia difusión de los indicadores bibliométricos en la evaluación del desempeño de la investigación, da cuenta de la inconformidad con esa forma de evaluación. Dicha inconformidad se explica por el hecho de que tales indicadores no son capaces de apreciar la diversidad de los productos de la investigación, ni la calidad ni el impacto de la producción de conocimiento y con ello, dan lugar a una imagen distorsionada de la productividad, calidad e impacto real de los resultados de la investigación y por lo tanto, del desempeño de los investigadores, grupos u organizaciones dedicadas a la producción de conocimiento.

A la vez, los actores que participan de ese debate proponen un conjunto de recomendaciones orientadas a mejorar las prácticas de evaluación. La recomendación de fondo es que las decisiones sobre la evaluación tienen que estar informadas por la valoración cualitativa de todas las contribuciones intelectuales realizadas por los investigadores y que los indicadores bibliométricos solo pueden utilizarse como insumos adicionales de información complementarios, siempre y cuando se consideren el conjunto de prescripciones metodológicas y/o sus limitaciones. De acuerdo con la opinión de los actores que participan de ese debate, solamente así se podrá valorar con inteligencia y precisión el rendimiento de la investigación -los productos que genera, su calidad e impacto- y el desempeño de investigadores, grupos y organizaciones productoras de conocimiento.

Tal como postula la teoría (capítulo anterior), los hallazgos de investigaciones pasadas -revisados en las secciones IV.2 y IV.3 de este capítulo- sugieren que los sistemas de evaluación fuertes, en tanto proveen incentivos materiales y simbólicos, contribuyen en la orientación de los investigadores, quienes tenderán a adaptar su comportamiento en materia de producción de conocimiento para cumplir con las exigencias que les impone el sistema de evaluación. Además, si esos sistemas de evaluación están informados por indicadores bibliométricos que carecen de validez, generan efectos constitutivos sobre la dirección, contenido y características epistemológicas de la producción de conocimiento.

Dicho de otro modo, y de acuerdo con los hallazgos de investigaciones pasadas sistematizados en la sección IV.3, la evidencia sugiere que el uso de indicadores bibliométricos en la evaluación de desempeño introduce un conjunto de sesgos en contra de la diversidad de: i) los tipos de productos derivados de las distintas prácticas de producción de conocimiento de los campos científicos; ii) el contenido y la dirección de esa producción; y iii) las características epistemológicas de los procesos de producción de conocimiento.

Respecto a los productos derivados de la producción de conocimiento, conforme esos sistemas de evaluación jerarquizan reconocen y valoran la producción de artículos científicos en inglés y su publicación en revistas internacionales indexadas en WoS y/o SCOPUS, desvalorizan la producción de: i) libros; ii) capítulos de libros; iii) artículos publicados en revistas científicas de circulación local y/o regional; iv) los reportes para informar la toma de decisiones de política pública; v) los reportes técnicos, el desarrollo de instrumentos y nuevas técnicas; vi) los desarrollos tecnológicos, entre otros.

Respecto al contenido y dirección de la producción de conocimiento, el uso de indicadores bibliométricos en la evaluación contribuye -debido a las estrategias adaptativas de los investigadores para cumplir con las exigencias que impone el sistema de evaluación, ya que de ello depende su prestigio y acceso a los recursos para la investigación- a que la producción de conocimiento esté orientada por el interés de las audiencias internacionales. Es decir, por el interés o las agendas de investigación de las revistas internacionales y sus editores, de las élites científicas dominantes o de la corriente principal y de las comunidades científicas internacionales. Así, la producción de conocimiento que se aparte de esas agendas de investigación internacionales es invisibilizada y devaluada ante los parámetros de rendimiento de esos sistemas de evaluación, en particular la producción de conocimiento contextualizada y preocupada por atender los desafíos que le impone su contexto social, político, económico y cultural.

Respecto a las características epistemológicas, la evaluación cuantitativa centrada en indicadores bibliométricos introduce sesgos sistemáticos en contra de la investigación orientada por la búsqueda de soluciones a problemas específicos dependientes del contexto, de la investigación no disciplinar o multi/inter/transdisciplinaria, de los procesos de producción de conocimiento de largo plazo, originales y/o disruptivos con la corriente principal y la producción de conocimiento colaborativa, especialmente la que supone la interacción con actores no académicos. Además, la presión que ejerce el sistema de evaluación respecto a la cantidad de publicaciones también parece inducir a una reducción de la calidad de la investigación, de la confiabilidad de sus resultados y relevancia científica.

Todo esto no solamente podría limitar el progreso de la ciencia, sino que además también parece restringir el impacto social y económico del conocimiento científico y tecnológico, y con ello, contradecir los objetivos de política pública orientados a estimular la producción de conocimiento de calidad y los orientados a poner al conocimiento científico y tecnológico al servicio del desarrollo social y económico de los países. Lo que resulta especialmente problemático para los países en desarrollo como México.



## CAPITULO V: El Marco Teórico-Conceptual: las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento

El objetivo de este capítulo es construir el marco teórico conceptual en que se sustenta el trabajo de investigación de esta tesis. Para ello, a continuación se retoman –de los capítulos II, III y IV- los conceptos y las relaciones que postulan la teoría y los hallazgos de investigaciones pasadas documentados en la literatura especializada sobre la relación entre la evaluación del desempeño de la investigación y la producción de conocimiento. La teoría y los hallazgos de investigaciones pasadas proveen los fundamentos que justifican y sobre los que se elabora tanto las definiciones conceptuales de las dimensiones relevantes para el análisis que se hace en esta investigación, así como las relaciones que se exploran entre ellas.

El capítulo se estructura en dos secciones. Le sigue a esta introducción la sección V.1 que, retomando ciertos conceptos provenientes de la teoría y los hallazgos de investigaciones pasadas, se enfoca en la definición conceptual de las dimensiones relevantes para el análisis que se hace en esta tesis. La sección V.2 se centra en el planteo de las relaciones teóricas hipotetizadas entre las dimensiones conceptuales relevantes para el análisis, que se ponen a prueba en esta tesis. El capítulo termina con la figura 2 que representa gráficamente el marco teórico conceptual que se construye en esta tesis, sintetizando las dimensiones conceptuales y las hipótesis que las vinculan. Es decir, la figura 2 representa sintéticamente la teoría estructural sobre las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento, que se propone y que se pone a prueba en esta investigación.

### V.1 Las definiciones conceptuales de las dimensiones relevantes: los constructos

La integración realizada en el capítulo III de la teoría que vincula a la evaluación de los resultados de la investigación y la producción de conocimiento con las contribuciones de Gibbons et al. (1994) – en adelante la teoría-, representa una síntesis importante porque justifica y provee los fundamentos teóricos sobre los que se elabora el análisis de las relaciones estudiadas en esta tesis.

Si la evaluación del desempeño de la investigación se relaciona con en el contenido, dirección y características epistemológicas de la producción de conocimiento y si además, esas características, contenido y dirección de la producción de conocimiento están asociadas con los modos diferenciales de producción de conocimiento, entonces es razonable esperar que dicha evaluación también se relacione con los Modos 1 y 2 de producción de conocimiento.

De ahí el fundamento teórico de que en esta tesis se espera encontrar relaciones significativas entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento. En particular entre tres dimensiones conceptuales –constructos- asociadas a la evaluación del desempeño de la investigación (*flexibilidad epistemológica del sistema de evaluación, satisfacción con el proceso de evaluación y dirección de la reforma del sistema de evaluación*) y los modos diferenciales de producción de conocimiento (*propensión modo 1 y propensión modo 2*).

Como se planteó, tanto la definición conceptual de los cinco constructos como las hipótesis que los vinculan, encuentran su justificación y fundamento en la teoría y los hallazgos de investigaciones pasadas. A continuación se presentan y se muestra qué conceptos provenientes de la teoría y de las

investigaciones pasadas se retomaron para justificar y fundamentar la definición conceptual de cada uno de los cinco constructos.

### V.1.1 Propensión modo 1

El primer constructo es *propensión modo 1* y se define conceptualmente en esta tesis como: *el grado de acuerdo de los investigadores categorizados en el sistema de evaluación con aquella producción de conocimiento que se caracteriza por desarrollar procesos de investigación orientados y organizados disciplinariamente, involucrar a actores homogéneos (investigadores) y difundir el conocimiento producido por canales formales (publicación de artículos, libros y capítulos de libro), lo que representa su principal fuente de reconocimiento científico.*

Así la definición conceptual del constructo *propensión modo 1* encuentra su justificación y fundamento teórico en la caracterización realizada por Gibbons et al. (1994) del fenómeno que denominó como Modo 1 de producción de conocimiento. En particular, porque la definición que se da de *propensión modo 1* retoma cinco de los seis atributos propuestos por dichos autores, cuatro son retomados textualmente y uno es adaptado de acuerdo a la información disponible y su consecuente posibilidad de operacionalización.

Los cuatro atributos que se retoman textualmente en la definición de *propensión modo 1* son: i) producción de conocimiento orientado por la disciplina; ii) producción de conocimiento disciplinar –organización-; iii) actores involucrados en la producción de conocimiento homogéneos; y iv) difusión del conocimiento a través de canales formales. El atributo que se adapta es el que Gibbons et al. (1994) vincula con que el *control de la calidad y legitimidad del conocimiento producido depende del juicio de los pares*. En la definición de *propensión modo 1* se adapta ese atributo cuando se enuncia que la principal fuente de reconocimiento científico son las publicaciones.

La racionalidad de la decisión tomada respecto a la adaptación de dicho atributo, se vincula con las posibilidades de operacionalizar el constructo *propensión modo 1* dada la información disponible en la encuesta. Es decir, antes de perder en la conceptualización y operacionalización de *propensión modo 1* el atributo teórico textual propuesto por Gibbons et al. (1994), se optó realizar dicha adaptación, ya que respeta el sentido conceptual del atributo teórico original.

El único atributo de los propuestos por Gibbons et al. (1994) en la caracterización del Modo 1 que no se retoma en la definición de *propensión modo 1*, es el vinculado con el contexto en que se produce el conocimiento, que en este caso y de acuerdo con los autores, corresponde al *contexto de la ciencia académica*. Si bien la encuesta contiene una variable con datos sobre este atributo, la inclusión de dicha variable observada en el modelo deterioraba la validez de constructo de *propensión modo 1*. Es por esta razón que se decidió no retener este atributo en la conceptualización de *propensión modo 1*, ni la variable observada en su medición.

Por lo argumentado hasta aquí, el constructo *propensión modo 1* es un concepto multidimensional y representa una aproximación válida y confiable (ver resultados en el capítulo X) para observar o medir, de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento de los investigadores categorizados en el SNI, el fenómeno Modo 1 de producción de conocimiento conceptualizado por Gibbons et al. (1994). En otros términos, el constructo *propensión modo 1* permite inferir en relación al Modo 1 como un tipo ideal de producción de conocimiento.

### V.1.2 Propensión modo 2

El segundo constructo es *propensión modo 2* y se define conceptualmente en esta tesis como: *el grado de acuerdo de los investigadores categorizados en el sistema de evaluación con aquella producción de conocimiento que se caracteriza por desarrollar procesos de investigación orientados por la búsqueda de soluciones a problemas dependientes del contexto y organizados inter/multi/transdisciplinariamente, involucrar a actores heterogéneos (investigadores y contrapartes interesadas en la solución de problemas) y difundir el conocimiento producido por canales informales (a través de las soluciones encontradas).*

Así, la definición conceptual del constructo *propensión modo 2* encuentra su justificación y fundamento teórico en la caracterización realizada por Gibbons et al. (1994) del fenómeno que denominó como Modo 2 de producción de conocimiento. En particular, porque la definición que se da de *propensión modo 2* retoma cuatro de los seis atributos propuestos por dichos autores, tres son retomados textualmente y uno es adaptado de acuerdo a la información disponible.

Los tres atributos que se retoman textualmente en la definición de *propensión modo 2* son: i) producción de conocimiento orientado por la búsqueda de soluciones a problemas dependientes del contexto; ii) actores involucrados en la producción de conocimiento heterogéneos; y iii) difusión del conocimiento a través de canales informales. El atributo que se adapta es el que Gibbons et al. (1994) vincula con *producción de conocimiento transdisciplinario*. En la definición de *propensión modo 2* se adapta ese atributo, agregando interdisciplinario y multidisciplinario. La racionalidad de la decisión tomada respecto a la adaptación de este atributo, se asocia con el hecho de que la variable observada disponible en la encuesta para operacionalizar este atributo es única e incluye indistintamente producción de conocimiento inter/multi/transdisciplinario. Antes de perder en la conceptualización y operacionalización de *propensión modo 2* el atributo teórico textual propuesto por Gibbons et al. (1994), se optó –prefirió– realizar dicha adaptación, ya que el autor utiliza el concepto de transdisciplinariedad para destacar que las normas que rigen la producción Modo 2 de conocimiento son no disciplinarias.

Los dos atributos de los propuestos por Gibbons et al. (1994) en la caracterización del Modo 2 que no se retoman en la definición de *propensión modo 2* son: i) el vinculado con el contexto en que se produce el conocimiento, que en este caso y de acuerdo con los autores, corresponde al *contexto de aplicación* y ii) el vinculado con que el control de la calidad y legitimidad del conocimiento producido depende del juicio de los pares y de los impares sobre el desempeño de las soluciones encontradas. Estos dos atributos no se retoman en la definición conceptual de *propensión modo 2* porque no hay posibilidades de operacionalización, ya que la encuesta no contiene ninguna variable que permita observar estos dos atributos. A pesar de esta limitación, *propensión modo 2* mantiene su validez de contenido, ya que retiene cuatro de los seis atributos teóricos y con ello su contenido refleja de buena manera el concepto Modo 2.

Por lo argumentado, el constructo *propensión modo 2* es un concepto multidimensional y representa una aproximación válida y confiable (ver resultados en el capítulo X) para observar o medir, de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento de los investigadores categorizados en el SNI, el fenómeno Modo 2 de producción de conocimiento conceptualizado por Gibbons et al. (1994). En otros términos, el constructo

*propensión modo 2* permite inferir en relación al Modo 2 como un tipo ideal de producción de conocimiento.

Los constructos *propensión modo 1* y *propensión modo 2* permiten medir e inferir en relación a los modos diferenciales de producción de conocimiento conceptualizados por Gibbons et al. (1994). De aquí su relevancia conceptual y empírica, hasta donde se sabe, en la literatura especializada - nacional e internacional- no existe ningún estudio que haya aproximado empíricamente esos conceptos con características similares a como se hace en este trabajo. Además, la revisión de la literatura realizada por Hessels y van Lente (2008), resalta la necesidad de realizar mayores esfuerzos orientados a aumentar la validez empírica de la conceptualización Modo 2 de Gibbons et al. (1994). Además, estos constructos son pertinentes para la política, ya que proveen información importante para identificar los modos diferenciales de producción de conocimiento y con ello, para el diseño de instrumentos de política científica específicos orientados a estimularlos.

### V.1.3 Flexibilidad epistemológica del sistema de evaluación

El tercer constructo es *flexibilidad epistemológica del sistema de evaluación* y se define conceptualmente en esta tesis como: *en qué medida los investigadores categorizados en el sistema de evaluación perciben que dicho sistema de evaluación permite diversidad de perfiles y de trayectorias de investigación, procesos de investigación de largo plazo y colectivos, el desarrollo de estrategias de investigación originales y novedosas, y la obtención de resultados de investigación suficientes –confiables- y científicamente relevantes.*

La definición conceptual del constructo *flexibilidad epistemológica del sistema de evaluación* encuentra su justificación y fundamento teórico en: i) la relación teórica postulada por la teoría; ii) las características epistemológicas de los procesos de producción de conocimiento identificadas por Gläser y Laudel (2007b); y iii) la contribución de Ziman (2000) sobre el concepto de *colectivización* para destacar que los procesos de investigación son cada vez más colectivos debido a los desafíos que impone el avance de la ciencia.

De acuerdo con la teoría, los *sistemas de evaluación fuertes* del desempeño de la investigación pueden limitar ciertas características epistemológicas de los procesos de producción de conocimiento. Teniendo presente esta relación y el tipo de información disponible para el análisis que se hace en esta tesis, es razonable esperar que exista un constructo - *flexibilidad epistemológica del sistema de evaluación*- capaz de explicar en qué medida los investigadores perciben que el sistema de evaluación permite ciertas características epistemológicas de los procesos de producción de conocimiento que desarrollan.

La definición conceptual del constructo *flexibilidad epistemológica del sistema de evaluación* también encuentra su justificación y fundamento teórico en las características epistemológicas de los procesos de producción de conocimiento identificadas por Gläser y Laudel (2007b) porque se inspira y adapta cuatro de las seis características epistemológicas planteadas por estos autores.

La adaptación que se hace de la conceptualización de Gläser y Laudel (2007b) tiene que ver con el hecho de que estos autores plantean las características epistemológicas de los procesos de producción de conocimiento. Sin embargo, en el contexto de esta tesis y dada la información disponible, interesa definir conceptualmente una dimensión que sirva para entender y observar en

qué medida los investigadores perciben que el sistema de evaluación permite ciertas características epistemológicas de los procesos de producción de conocimiento que desarrollan.

La primera característica epistemológica planteada por Gläser y Laudel (2007b) que se adapta es la vinculada con el *tipo de investigación y orientación predominante*. Con estos conceptos los autores refieren a los diversos tipos de investigación -teórica, metodológica, experimental y empírica-, y sus distintas orientaciones predominantes -básica, estratégica o aplicada-. En la definición de *flexibilidad epistemológica del sistema de evaluación* se adapta esta característica cuando se enuncia en qué medida, de acuerdo con la percepción de los investigadores categorizados en ese sistema, el sistema de evaluación permite diversidad de perfiles de investigación y de trayectorias de investigación. Como se adelantó, la racionalidad de la decisión tomada respecto a la adaptación de esta característica, se vincula con las posibilidades de operacionalizar el constructo *flexibilidad epistemológica del sistema de evaluación* dada la información disponible en la encuesta. Conceptualmente, la adaptación que se hace respeta la noción de diversidad derivada de los distintos tipos y orientaciones de la investigación que resaltan Gläser y Laudel (2007b) con esta característica epistemológica.

La segunda característica epistemológica planteada por Gläser y Laudel (2007b) que se adapta es la vinculada con las *características temporales de los procesos de investigación*. Con este concepto los autores caracterizan a los procesos de investigación en términos de su duración (investigación de corto plazo *versus* investigación de largo plazo). En la definición de *flexibilidad epistemológica del sistema de evaluación* se adapta esta característica cuando se enuncia en qué medida, de acuerdo con la percepción de los investigadores categorizados en ese sistema, el sistema de evaluación permite procesos de investigación de largo plazo. Nuevamente, la racionalidad de la decisión tomada respecto a la adaptación de esta característica, se vincula con las posibilidades de operacionalizar el constructo *flexibilidad epistemológica del sistema de evaluación* dada la información disponible en la encuesta. Conceptualmente, la adaptación que se hace respeta la noción asociada a las características temporales de los procesos de investigación que resaltan Gläser y Laudel (2007b).

La tercera característica epistemológica planteada por Gläser y Laudel (2007b) que se adapta es la vinculada con el *grado de riesgo intelectual asumido*. Con este concepto los autores caracterizan a los procesos de investigación en términos del riesgo intelectual que asumen. En la definición de *flexibilidad epistemológica del sistema de evaluación* se adapta esta característica cuando se enuncia en qué medida, de acuerdo con la percepción de los investigadores categorizados en ese sistema, el sistema de evaluación permite el desarrollo de estrategias de investigación originales y novedosas. Otra vez aquí, la racionalidad de la decisión tomada respecto a la adaptación de esta característica, se vincula con las posibilidades de operacionalizar el constructo *flexibilidad epistemológica del sistema de evaluación* dada la información disponible en la encuesta. Conceptualmente, la adaptación que se hace respeta la noción asociada al *riesgo intelectual asumido* que resaltan Gläser y Laudel (2007b), ya que los hallazgos de investigación de Wang, Veugelers y Stephan (2017) confirman la asociación positiva entre riesgo intelectual asumido y novedad.

La cuarta característica epistemológica planteada por Gläser y Laudel (2007b) que se adapta es la vinculada con la *confiabilidad de los resultados*. Con este concepto los autores refieren a la

confiabilidad de los resultados de investigación. En la definición de *flexibilidad epistemológica del sistema de evaluación* se adapta esta característica cuando se enuncia en qué medida, de acuerdo con la percepción de los investigadores categorizados en ese sistema, el sistema de evaluación permite obtener resultados suficientes y científicamente relevantes. Otra vez aquí, la racionalidad de la decisión tomada respecto a la adaptación de esta característica, se vincula con las posibilidades de operacionalizar el constructo *flexibilidad epistemológica del sistema de evaluación* dada la información disponible en la encuesta. Conceptualmente, la adaptación que se hace respeta y amplía la noción asociada a la confiabilidad de los resultados de investigación que resaltan Gläser y Laudel (2007b). Por un lado, la respeta porque en el contexto del formulario –a partir del cual se recolectaron los datos que se utilizan en esta investigación– cuando utiliza la expresión “resultados suficientes”, aunque no es preciso, puede interpretarse que dicha suficiencia tiene que ver con la obtención de resultados científicamente confiables. Por otro lado, la amplía porque introduce también la noción de relevancia científica.

Las dos características epistemológicas de los procesos de producción de conocimiento planteadas por Gläser y Laudel (2007b) que no se retoman en la definición de *flexibilidad epistemológica del sistema de evaluación* son: i) la relación de la investigación con la opinión mayoritaria de la comunidad (corriente principal –*mainstream*– versus no conformistas), y ii) el grado de heterogeneidad en la combinación de conocimientos (interdisciplinariedad). Estas dos características no se retoman en la definición conceptual de *flexibilidad epistemológica del sistema de evaluación* porque no hay posibilidades de operacionalización, ya que la encuesta no contiene ninguna variable que permita observarlas. Hasta aquí, *flexibilidad epistemológica del sistema de evaluación* mantiene su validez de contenido, ya que retiene cuatro de las seis características epistemológicas propuestas por Gläser y Laudel (2007b) y con ello, su contenido refleja de buena manera la relación de la evaluación con dichas características.

De todas maneras y a efectos de mejorar aún más la validez de contenido de *flexibilidad epistemológica del sistema de evaluación*, en su definición conceptual se introduce una característica epistemológica adicional a las identificadas por Gläser y Laudel (2007b), cuando se enuncia en qué medida, de acuerdo con la percepción de los investigadores categorizados en ese sistema, el sistema de evaluación permite procesos de investigación colectivos. Conceptualmente, se considera que los procesos de investigación colectivos representan una característica epistemológica de los procesos de producción de conocimiento ya que suponen una manera específica de conocer y desarrollar conocimiento, distinta de los procesos de investigación individuales. Tal como plantea Ziman (2000), los procesos de investigación son cada vez más colectivos debido a los desafíos que impone el avance de la ciencia. Dicho de otro modo, la ciencia ha avanzado tanto que sus principales problemas o desafíos tienen que ser estudiados colectivamente.

Por lo argumentado, la *flexibilidad epistemológica del sistema de evaluación* es un concepto multidimensional útil para entender y aproximar empíricamente en qué medida el sistema de evaluación permite ciertas características epistemológicas de los procesos de producción de conocimiento. Además es un constructo conceptualmente muy relevante y pertinente para la política y la ciencia, ya que vincula a la evaluación de la investigación con ciertas características epistemológicas que están asociadas a la calidad de la producción de conocimiento de acuerdo con la conceptualización de calidad que provee Mårtensson et al. (2016) –ver capítulo anterior-. En

resumen, la *flexibilidad epistemológica del sistema de evaluación* provee una aproximación válida y confiable (ver resultados en el capítulo X) para saber en qué medida el sistema de evaluación permite procesos de producción de conocimiento de calidad, originales, de largo plazo, colectivos y con diversidad de perfiles y trayectorias de investigación. Finalmente, el constructo *flexibilidad epistemológica del sistema de evaluación* permite medir e inferir sobre la relación entre la evaluación y las características epistemológicas de la producción de conocimiento.

#### V.1.4 Satisfacción con el proceso de evaluación

El cuarto constructo es *satisfacción con el proceso de evaluación* y se define conceptualmente en esta tesis como: *el nivel de satisfacción que tienen los investigadores categorizados en el sistema de evaluación respecto a si el proceso de evaluación es capaz de reconocer el desempeño, es decir, si cuenta con la información necesaria para elaborar juicios académicos apropiados sobre el desempeño de los investigadores e incluye todas sus contribuciones importantes, si los criterios y mecanismos de selección de los evaluadores garantizan evaluaciones académicas apropiadas en las diversas áreas de conocimiento y si la composición de las comisiones dictaminadoras respetan los equilibrios regionales e institucionales y dan garantías respecto a la objetividad de la evaluación.*

De acuerdo con la revisión del debate internacional -presentado en el capítulo IV-, en la actualidad no existen estándares internacionalmente aceptados para evaluar objetivamente el desempeño de la investigación. Los procesos de evaluación del desempeño de la investigación enfrentan diversos desafíos debido a su dificultad y multidimensionalidad y por lo mismo, todavía no se han podido encontrar procesos de evaluación que sean satisfactorios para todos los actores involucrados en ellos. Teniendo presente este debate y el tipo de información disponible para el análisis que se hace en esta tesis, es razonable esperar que exista un constructo - *satisfacción con el proceso de evaluación*- capaz de explicar el nivel de satisfacción que tienen los investigadores con el proceso de evaluación del desempeño de la investigación al que están sometidos.

La definición conceptual del constructo *satisfacción con el proceso de evaluación* encuentra su justificación y fundamento teórico en dicho debate. De acuerdo con ese debate, se puede sostener que las fuentes de insatisfacción con los procesos de evaluación del desempeño de la investigación se asocian con dos aspectos vinculados entre sí, a saber: i) los criterios de evaluación utilizados y su relación con los productos considerados en el proceso de evaluación; y ii) los evaluadores encargados de juzgar las contribuciones intelectuales -productos- de los investigadores para elaborar sus juicios de evaluación. El vínculo más general entre esas dos fuentes de insatisfacción está en que los evaluadores aplicarán ciertos criterios de evaluación para juzgar las contribuciones intelectuales de los investigadores y su desempeño. Dependiendo de cuáles sean esos criterios, serán también los productos que se consideran en la evaluación y cómo éstos se consideran o valoran, es decir, qué datos o parámetros se utilizan para informar el juicio de los evaluadores sobre el desempeño de la investigación que realizan los investigadores.

El constructo *satisfacción con el proceso de evaluación* retoma en su definición esas dos fuentes de insatisfacción y de ahí su justificación y fundamento teórico. A continuación se explica cómo son retomadas dichas fuentes.

En el debate internacional, diversas declaraciones y distintos estudios han planteado el problema derivado de la amplia difusión de indicadores bibliométricos en la gobernanza de la ciencia (Hicks et al., 2015) es decir, de su utilización como criterios de evaluación centrales, sino únicos, para

informar el juicio de los evaluadores sobre el desempeño de la investigación realizada por individuos, grupos u organizaciones productoras de conocimiento.

El sistema de evaluación del desempeño de la investigación que se estudia en esta tesis –el SNI- no es la excepción a esa tendencia mundial respecto a la difusión de los indicadores bibliométricos en la gobernanza de la ciencia. De acuerdo con la descripción realizada de las características específicas de dicho sistema (ver capítulo VI, sección VI.2), se puede argumentar que sus criterios de evaluación se basan prioritariamente en indicadores bibliométricos<sup>26</sup>.

El problema o fuente de insatisfacción con los procesos de evaluación centrados en el uso casi exclusivo de indicadores bibliométricos como medidas para aproximar productividad, calidad e impacto de la producción de conocimiento, tiene que ver con el hecho de que dichos indicadores no son capaces de valorar por sí solos el desempeño de la investigación (Sahel, 2011) porque no son sensibles a sus objetivos y naturaleza (Hicks et al., 2015) y por lo tanto no capturan la especificidad y diversidad de la producción de conocimiento (Hicks, 2004, DORA, 2012). Tal como discute Hicks (2012), los criterios de evaluación informados por indicadores bibliométricos no consideran las especificidades de los productos y los resultados de investigación de los distintos campos científicos.

Es decir, el indicador de productividad al tiempo que reconoce un tipo de producto -los artículos científicos publicados-, desconoce -o valora menos- los restantes productos también derivados de la producción de conocimiento. Así, la relación entre ese criterio de evaluación y los diversos productos que genera la investigación es baja, y de ese modo pueden conducir a una visión distorsionada de la productividad de los investigadores.

Adicionalmente, los indicadores basados en citas también son cuestionados porque no miden calidad, aproximan un aspecto de la calidad: su difusión internacional (Gläser y Laudel, 2007a). Los factores de impacto tampoco miden la calidad de los artículos publicados (DORA, 2012) ni reflejan el interés de la comunidad científica en ellos (Pendlebury; 2009; Moed 2009) y, además, tampoco son confiables debido a ciertas estrategias implementadas por algunos editores de las revistas para incrementarlos artificialmente (Martin, 2016). Simultáneamente, los indicadores bibliométricos en general y los factores de impacto en particular, no solamente varían significativamente según la cobertura de la base bibliográfica que se utilice para su cálculo, sino que también presentan alta variabilidad entre los campos científicos –en las ciencias básicas los factores de impacto son más altos que en áreas muy especializadas o aplicadas- (Sahel, 2011; AFC, 2011; DORA, 2012).

Así, la relación entre esos criterios de evaluación (cantidad de citas y factores de impacto) y la calidad e impacto de la producción de conocimiento es baja, y con ello pueden proveer una imagen distorsionada justamente de la calidad e impacto que en realidad tiene la producción de conocimiento (Gläser y Laudel, 2007a; Hicks, 2004).

En suma, los procesos de evaluación cuyos criterios están prioritariamente centrados en indicadores bibliométricos que se relacionan débilmente con la productividad, calidad e impacto de la producción de conocimiento realizada por los investigadores, generan insatisfacción entre ellos porque pueden dar lugar a una imagen distorsionada de su desempeño real.

---

<sup>26</sup> Tales como: i) cantidad de artículos publicados en revistas indexadas –preferentemente- en WoS; ii) cantidad de citas; y iii) los factores de impacto de las revistas.



La segunda fuente de insatisfacción con los procesos de evaluación se asocia con la subjetividad de los evaluadores que influencia la revisión y el juicio sobre las solicitudes (AFC, 2011), ya sea por favoritismos disciplinares, locales y/o institucionales o por falta de conocimiento específico de los evaluadores en el área de investigación objeto de evaluación (Sahel, 2011).

Estas dos fuentes de insatisfacción se retoman en la definición conceptual de *satisfacción con el proceso de evaluación* usada en esta tesis cuando se enuncia: qué nivel de satisfacción tienen los investigadores categorizados en el sistema de evaluación –SNI- respecto a la capacidad que tiene el proceso de evaluación de dicho sistema de reconocer el desempeño, es decir, si cuenta con la información necesaria para elaborar juicios académicos apropiados del desempeño de los investigadores e incluye todas sus contribuciones importantes, si los criterios y mecanismos de selección de los evaluadores garantizan evaluaciones académicas apropiadas en las diversas áreas de conocimiento y si la composición de las comisiones dictaminadoras respetan los equilibrios regionales e institucionales y dan garantías respecto a la objetividad de la evaluación.

Por lo argumentado, la *satisfacción con el proceso de evaluación* es un concepto multidimensional, válido y confiable (ver resultados en el capítulo X), que es útil para entender y aproximar empíricamente el nivel de satisfacción que tienen los investigadores categorizados en el sistema con el proceso de evaluación al que están sometidos, en particular respecto a si dicho proceso de evaluación es capaz de reconocer integralmente y objetivamente su desempeño. Además es un concepto pertinente para la política ya que aporta información relevante sobre si el proceso de evaluación es capaz de reconocer integralmente y objetivamente el desempeño de los investigadores. Finalmente, el constructo *satisfacción con el proceso de evaluación* permite medir e inferir sobre la relación entre la percepción que tienen los investigadores sobre el proceso de evaluación y su desempeño de investigación.

#### V.1.5 Dirección de la reforma del sistema de evaluación

El quinto constructo es *dirección de la reforma del sistema de evaluación* y se define conceptualmente en esta tesis como: *el interés que tienen los investigadores categorizados en el sistema de evaluación de reformarlo en una dirección que cambia la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, con mejor consideración de la investigación multi/inter/transdisciplinaria y que valore más los resultados de la investigación aplicada u orientada a la resolución de problemas.*

A pesar de la amplia difusión de los métodos bibliométricos en las prácticas de evaluación, el debate internacional está abierto. Como se dijo, no existe acuerdo sobre cómo evaluar y los criterios a utilizar. El debate internacional además de evidenciar las fuentes de insatisfacción con los procesos de evaluación del desempeño de la investigación, provee un conjunto de recomendaciones orientadas a mejorar o cambiar las prácticas de evaluación de tal forma que la misma sea capaz de apreciar y valorar integralmente el desempeño de la investigación y el rendimiento de los productos que genera, y evitar con ello imágenes distorsionadas tanto de la producción de conocimiento así como de su calidad e impacto. Teniendo presente estas recomendaciones y el tipo de información disponible para el análisis que se hace en esta tesis, es razonable esperar que exista un constructo *-dirección de la reforma del sistema de evaluación-* capaz de explicar el interés que tienen los investigadores de cambiar –reformar- el sistema de evaluación al que están sometidos.

La definición conceptual del constructo *dirección de la reforma del sistema de evaluación* encuentra su justificación y fundamento teórico en dicho debate y en los hallazgos de investigaciones pasadas reportados en la literatura especializada. En particular porque dicha definición se inspira en las recomendaciones que surgen de ese debate y la evidencia que reporta la literatura, cuya conjunción provee los elementos de juicio que habría que considerar para reformar el sistema de evaluación y definir en qué dirección.

A continuación se sintetizan, por un lado, las recomendaciones del debate internacional orientadas a mejorar las prácticas de evaluación de tal forma que la misma sea capaz de apreciar y valorar integralmente el desempeño de la investigación y el rendimiento de los productos que genera, y evitar con ello imágenes distorsionadas tanto de la productividad de la producción de conocimiento así como de su calidad e impacto. Por otro lado, se sintetizan los hallazgos de investigaciones pasadas que dan cuenta de porqué la evaluación del desempeño de la investigación centrada en indicadores bibliométricos proveen una imagen distorsionada del desempeño de la investigación. Estas dos síntesis, conjuntamente aportan los elementos a tener en cuenta para pensar la dirección que debiera asumir una reforma al sistema de evaluación y que sintéticamente están contenidas como elementos relevantes en la definición conceptual del constructo *dirección de la reforma del sistema de evaluación*.

La recomendación más general y de fondo es que los indicadores bibliométricos solo pueden utilizarse como criterios de evaluación si son complementarios a la evaluación cualitativa realizada por pares -expertos-, y en ningún caso, dichos indicadores deberían ser el único insumo de evaluación (AFC, 2011; Sahel, 2011; DORA, 2012; Hicks et al., 2015). Hicks et al. (2015), en coincidencia con la AFC (2011), DORA (2012) y Alberts (2013), plantean que el problema está en que la evaluación de la investigación dejó de estar fuertemente basada en la valoración cualitativa de los *expertos*, y pasó a depender centralmente –o exclusivamente- de ciertas métricas. Sin embargo, estas métricas no pueden ser la única fuente para informar el juicio de los pares (Hicks et al., 2015).

De acuerdo con la AFC (2011), la evaluación cualitativa de los resultados de la investigación, para informar el juicio que realizan los pares, tiene un valor insustituible para apreciar la originalidad, calidad, relevancia, pertinencia e impacto de la producción de conocimiento. En esta misma dirección apuntan las recomendaciones de DORA (2012) cuando señala que la evaluación tiene que centrarse en los méritos intrínsecos de la investigación, más que en la jerarquía o prestigio de la revista en la que se publican sus resultados (DORA, 2012). El Manifiesto de Leiden resalta que la evaluación del desempeño individual de los investigadores debe realizarse con base en la *valoración cualitativa de su portafolio de investigación*, y recomienda la adopción de un enfoque de evaluación que considere información diversa sobre cada investigador y que también valore la investigación de relevancia y pertinencia local (Hicks et al 2015).

Los productos de la investigación científica son múltiples y variados, por lo tanto los criterios de evaluación que se utilicen y sus indicadores, para que sean capaces de evaluar con precisión e inteligencia la calidad y el impacto de la producción de conocimiento (AFC, 2011; DORA, 2012), tienen que ser sensibles tanto a los objetivos y naturaleza de la investigación evaluada, como también a los contextos en los que se produce el conocimiento (Hicks et al., 2015). Sin embargo, el estudio realizado por Bensusán et al. (2014), sobre el SNI de México, señala que la implementación

de indicadores en la evaluación que se realiza en ese sistema, no atiende a las características contextuales en las que los investigadores desarrollan sus actividades de investigación.

La AFC (2011), DORA (2012) y Hicks et al. (2015) recomiendan considerar en la evaluación los diversos productos de la investigación y el impacto de todos ellos -no solamente de las publicaciones científicas- para ponderar la influencia que tiene la producción de conocimiento en la política y en la práctica (DORA, 2012) o en la resolución de problemas (AFC, 2011).

Estas recomendaciones apuntan directamente a evitar lo que sucede en muchos sistemas de evaluación, incluido el SNI, que asimilan –debido a los indicadores que utilizan en la evaluación- excelencia académica únicamente con publicaciones en revistas con alto factor de impacto. Esa asimilación unívoca produce múltiples sesgos en los resultados de la evaluación a favor de ciertos estilos de producción y difusión del conocimiento y sus productos; y con ello desestimula aquella producción de conocimiento, cuyos estilos de producción, difusión y los productos que genera no exhiben tan buen rendimiento en los parámetros que utiliza el sistema de evaluación para reconocer el desempeño.

Distintos estudios coinciden en que la amplia difusión de indicadores bibliométricos en la gobernanza de la ciencia como criterios centrales de evaluación ha inducido a una estandarización u homogeneización de los tipos de investigación y sus productos (De Rijcke et al., 2016), donde la producción de artículos científicos en inglés y su publicación en revistas –preferentemente- con alto factor de impacto e indexadas a bases de datos bibliográficas internacionales (WoS o SCOPUS) ha sido jerarquizada y mejor valorada en la evaluación del desempeño (Leisyte y Westerheijden, 2014; De Ibarrola, 2012).

Simultáneamente, diversos estudios también coinciden en que el uso sistemático de esos criterios de evaluación deterioran y desestimulan la producción de conocimiento contextualizada, es decir, la que se ocupa de atender los problemas complejos dependientes del contexto, cuya pertinencia social y económica es sustantiva, en particular para los países en desarrollo (Bianco, Gras y Sutz, 2016). Ese fenómeno se puede explicar por el hecho de que los temas que tienen gran pertinencia local no son bien cubiertos por las revistas internacionales indexadas a WoS o SCOPUS, por lo que los investigadores para cumplir con los requerimientos de los sistemas de evaluación (publicar en dichas revistas) cambian el contenido de su agenda de investigación, desde una con mayor pertinencia regional-local (contextualizada) hacia una más internacionalizada (Chavarro, Tang y Rafols, 2017). Los hallazgos del estudio de van Dalen y Henkes (2012) y del realizado por Müller y De Rijcke (2017) son consistentes con esta explicación, ya que el primero resalta que la presión por publicar internacionalmente se asocia negativamente con la orientación de la investigación hacia temas con pertinencia local, y el segundo destaca que la evaluación centrada en indicadores de rendimiento –bibliométricos- induce a que los investigadores pongan menos atención a la responsabilidad social de la investigación en sus prácticas de producción de conocimiento.

De acuerdo con Vessuri, Guédon y Cetto (2014), esos criterios han estimulado un estilo único o dominante de producción de conocimiento internacionalizado pero descontextualizado de su espacio regional, social, político, disciplinar e idiomático; negando y desplazando la investigación orientada a atender los problemas que afectan al desarrollo social y económico de América Latina y otras regiones periféricas. Tal como discuten Hicks et al. (2015), dichos indicadores inducen a que la investigación se centre en temas de interés internacional pero que no necesariamente tienen

relevancia y pertinencia local y/o regional, y en temas más disciplinares que inter/multi/transdisciplinares.

El estudio de Martin y Whitley (2010) muestra que la evaluación cuantitativa realizada en el RAE del Reino Unido ha inducido a que la investigación sea más disciplinar que interdisciplinar y el estudio realizado por Leisyte y Westerheijden (2014) consistentemente muestra que la evaluación del desempeño cuantitativa ha reforzado los estándares de excelencia disciplinares. Este fenómeno se puede explicar por el hecho de que las evaluaciones de desempeño que utilizan *rankings* de revistas exhiben sesgos sistemáticos en contra de la investigación interdisciplinar, lo que puede traducirse en que los investigadores perciban menores incentivos para trabajar interdisciplinariamente y por lo tanto, también para trabajar en la búsqueda de soluciones a problemas complejos (básicos-fundamentales y/o dependientes del contexto ya sean sociales, productivos, de política pública, ambientales, energéticos, etc.) (Rafols et al., 2012).

Consistentemente, en el caso de México, diversos estudios coinciden en que el SNI al tiempo que estimula la producción de conocimiento especializada disciplinarmente, acelerada, individualizada e internacionalizada, también impone diversas dificultades que limitan la orientación de la producción de conocimiento hacia la atención de problemas complejos dependientes del contexto (Bensusán et al., 2014), ya que sus criterios de evaluación no reconocen la producción de conocimiento orientada a atender los principales desafíos que impone el desarrollo social y económico de México (Esteinou, 2012) y tienen escasa apertura para reconocer y valorar la producción de conocimiento multi /inter/transdisciplinaria (Bensusán et al., 2014).

Otros estudios también han destacado que el SNI limita la investigación colaborativa, en particular la que se caracteriza por su interacción con actores externos al ámbito de la academia (Becerra, 2014; Galaz Fontes y Gil Antón, 2009; Amaro, Corona y Soria, 2008; Galaz Fontes et al., 2008; Grediaga Kuri, 1998). Adicionalmente, el SNI debido a los criterios que utiliza en la evaluación del desempeño de la investigación, también dificulta o devalúa la investigación aplicada (Valenti et al., 2013; Valenti, 2011) y desplaza las actividades orientadas al desarrollo tecnológico (Padilla, 2010). Tal como muestran los hallazgos de Rivera et al. (2011), los investigadores en el área de las ciencias agrícolas generan variedad de productos, algunos orientados a mover la frontera del conocimiento y otros más orientados por la búsqueda de soluciones a problemas productivos, sin embargo, el SNI estimula la producción de artículos científicos publicados en revistas internacionales pero no así otros productos técnicos. El escaso o nulo reconocimiento de ese tipo de productos determina un desincentivo importante para la interacción entre la academia y el sector productivo nacional.

Además de lo ya señalado, otros hallazgos de investigaciones pasadas reportados en la literatura internacional también son consistentes con lo observado para el caso del SNI de México. Los sistemas de evaluación centrados en indicadores bibliométricos devalúan los tipos de investigación aplicada (De Rijcke et al., 2016; Martin y Whitley, 2010) y comprometen los objetivos de política pública orientados a estimular la innovación y la interacción de la academia con el sector productivo (Hicks, 2012). En particular, porque dichos sistemas de evaluación devalúan la investigación orientada a la obtención de resultados y/o productos pertinentes para la política pública (De Rijcke et al., 2016; van Dalen y Henkes, 2012) y la que se caracteriza por desarrollarse a través del trabajo de investigación colaborativo (De Rijcke et al., 2016). El estudio de Gläser y Laudel (2007b) muestra que la evaluación cuantitativa de los resultados de investigación ha inducido a un aumento de la

producción de artículos elaborados por un único autor, es decir, artículos científicos que son resultado de procesos de producción de conocimiento individuales.

En síntesis, de acuerdo con el debate internacional, la principal recomendación para mejorar las prácticas de evaluación es que la evaluación del desempeño de la investigación tiene que estar centrada en la valoración y revisión cualitativa de todas las contribuciones intelectuales que realizan los investigadores en los diversos campos científicos, atendiendo a sus diferencias, reconociendo y valorando todos sus resultados y productos de acuerdo a sus objetivos de investigación. En este modelo de evaluación cualitativa, los indicadores bibliométricos pueden ser utilizados solo como insumos de información adicionales y complementarios, y nunca ser los únicos ni más importantes criterios para informar los juicios de los pares sobre el desempeño de la investigación. Con esta recomendación, los actores que participan en el debate internacional buscan evitar los múltiples sesgos, entre ellos, en contra de la investigación colectiva, no disciplinar y la orientada a la resolución de problemas dependientes del contexto (de política pública, sociales, productivos, etc.) que parece introducir la evaluación cuantitativa -centrada en indicadores bibliométricos- de los resultados de la investigación sobre la organización, contenido y dirección de la producción de conocimiento. Como fue señalado, estos sesgos han sido ampliamente documentados –en general a través de estudios de caso- entre los hallazgos de investigaciones pasadas y reportados en la literatura especializada internacional y nacional.

La definición conceptual del constructo *dirección de la reforma del sistema de evaluación* se inspira y conjuga esta recomendación y algunos de los diversos sesgos a los que contribuye la evaluación centrada en indicadores bibliométricos reportados en la literatura, cuando enuncia el interés que tienen los investigadores de cambiar la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, mejor consideración de la investigación multi/inter/transdisciplinaria y que valore más los resultados de la investigación aplicada u orientada a la resolución de problemas.

Así, el constructo *dirección de la reforma del sistema de evaluación* reúne algunos de los elementos de juicio que habría que considerar para reformar el sistema de evaluación y en qué dirección, de tal forma que la evaluación del desempeño de la investigación que se realiza en ese sistema sea capaz de apreciar y valorar integralmente el desempeño de la investigación y el rendimiento de los productos que genera, y evitar con ello imágenes distorsionadas tanto de la productividad de la producción de conocimiento así como de su calidad e impacto.

Por lo argumentado, la *dirección de la reforma del sistema de evaluación* es un concepto multidimensional válido y confiable (ver resultados en el capítulo X), que es útil para entender y aproximar empíricamente el interés que tienen los investigadores categorizados en el sistema de reformarlo en una dirección, cuyo proceso de evaluación sea más cualitativo, que respete las diferencias disciplinares en sus prácticas de producción de conocimiento y que valore y reconozca la investigación colectiva, multi/inter/transdisciplinaria y los resultados de la investigación aplicada u orientada a la resolución de problemas.

De acuerdo con lo argumentado hasta aquí, las definiciones conceptuales que se dan en esta tesis de los constructos *propensión modo 1*, *propensión modo 2*, *flexibilidad epistemológica del sistema de evaluación*, *satisfacción con el proceso de evaluación* y *dirección de la reforma del sistema de evaluación* se justifican y encuentran su fundamento en la teoría, el debate internacional y los

hallazgos de investigaciones pasadas. En el capítulo X se operacionalizan todos los constructos y se presentan los resultados asociados a su validez y confiabilidad.

A continuación, se argumenta que las relaciones teóricas hipotetizadas entre estos cinco constructos y que se ponen a prueba en esta tesis, también encuentran su justificación y fundamento en la teoría, el debate internacional y los hallazgos de investigaciones pasadas.

## V.2 La teoría estructural sobre las relaciones entre la evaluación y los modos diferenciales de producción de conocimiento: hipótesis sobre las relaciones entre los constructos

Distintos estudios coinciden en que distintos tipos de producción de conocimiento requieren formas diferenciales de evaluación (de Jong et al., 2011; Regeer et al., 2009; Spaapen, Dijkstra y Wamelink, 2007; Hemlin y Barlebo Rasmussen, 2006; Gibbons et al., 1994) para poder estimular la diversidad y la calidad en todas las áreas de conocimiento, atendiendo a sus diferencias, reconociendo y valorando todos sus resultados y productos de acuerdo a sus objetivos de investigación (Bianco, Gras y Sutz, 2014; Hicks, 2013; de Jong et al., 2011).

Sin embargo, de acuerdo con el debate internacional y los hallazgos de investigaciones pasadas, los sistemas de evaluación del desempeño de la investigación reconocen y valoran fundamentalmente un único tipo de producto y medio de difusión para los resultados de la producción de conocimiento, a saber: la producción de artículos científicos en inglés y su publicación en revistas – preferentemente con alto factor de impacto- e indexadas a bases de datos bibliográficas internacionales (WoS o SCOPUS).

De acuerdo con De Rijcke et al. (2016), eso ha inducido a prácticas imitativas de ese estilo de producción y difusión de conocimiento por parte de las disciplinas que gozan de menor prestigio en el sistema de ciencia. Leisyte y Westerheijden (2014) destacan que esos sistemas de evaluación presionan a que todos los temas y áreas de investigación imiten el modelo de producción de conocimiento y difusión de los resultados de investigación de las ciencias naturales y exactas o *ciencias duras*.

Consistentemente, el estudio de Bensusán et al. (2014) para el caso del SNI de México, señala que para algunas áreas de conocimiento existe cierta imposición de los estándares de productividad y calidad -propios de las ciencias exactas y naturales- que les son ajenos a sus prácticas de investigación. Es más, de acuerdo con Russell y Liberman (2002) el medio de difusión de los resultados de investigación de mayor relevancia para las ciencias exactas y naturales es la publicación de artículos científicos en revistas internacionales, para las ciencias sociales y las humanidades son los libros, capítulos de libros y artículos en revistas de circulación nacional y regional, y las ciencias aplicadas –ingenierías y biotecnologías- están en un lugar intermedio entre esos dos extremos. Para estos autores, esos modelos diferenciales de difusión de los resultados de investigación de los investigadores mexicanos, reflejan sus prácticas también diversas de producción de conocimiento.

Consistentemente, González-Brambila (2005) demuestra que los investigadores categorizados en el SNI de las distintas áreas de conocimiento exhiben un comportamiento diferencial respecto a la publicación en revistas internacionales indexadas en WoS. Su estudio muestra que los investigadores de ciencias exactas, biología, química y ciencias de la salud publican más y por más

tiempo en esas revistas, en comparación a los investigadores de ciencias sociales, humanidades, ciencias agrícolas, biotecnologías e ingenierías que trabajan sobre temas de investigación con intereses más locales, cuyos resultados tienden a difundirse en otros medios menos globales y más asociados a las audiencias locales/regionales. La autora concluye que los artículos científicos no son el único producto de la investigación y que para tener una imagen más completa de la productividad de los investigadores hay que considerar los restantes productos que la misma genera.

Así, los sistemas de evaluación centrados en indicadores bibliométricos pueden dar lugar a una imagen distorsionada del desempeño de investigación de los investigadores de las distintas áreas de conocimiento, ya que la contabilización de artículos puede ser un indicador que refleje adecuadamente el desempeño de las ciencias exactas, biología, química y ciencias de la salud pero no el de las ciencias sociales y las humanidades (Hicks, 2012) ni tampoco el de las ciencias agrícolas, biotecnologías e ingenierías.

Tal como se argumentó, de acuerdo con el debate internacional los procesos de evaluación cuyos criterios están prioritariamente centrados en indicadores bibliométricos que se relacionan débilmente con la productividad, calidad e impacto de la producción de conocimiento realizada por los investigadores de las distintas áreas de conocimiento, generan insatisfacción entre ellos porque pueden dar lugar a una imagen distorsionada de su desempeño real. Simultáneamente, los actores que participan de ese debate internacional realizan un conjunto de recomendaciones orientadas a mejorar las prácticas de evaluación de tal forma que la misma sea capaz de apreciar y valorar integralmente el desempeño de la investigación y el rendimiento de todos los productos que genera, y evitar con ello imágenes distorsionadas tanto de la productividad de la producción de conocimiento así como de su calidad e impacto. De acuerdo con estas ideas, se puede sostener que existe una asociación positiva entre la insatisfacción con los procesos de evaluación, cuyos criterios están prioritariamente centrados en indicadores bibliométricos, y las recomendaciones orientadas a mejorar las prácticas de evaluación, es decir orientadas a reformar dichos procesos de evaluación.

Teniendo presente que las definiciones conceptuales que se dan en esta tesis de los constructos *satisfacción con el proceso de evaluación* y *dirección de la reforma del sistema de evaluación* retoman las fuentes de insatisfacción que surgen del debate internacional y se inspiran en las recomendaciones que emergen de ese debate y los hallazgos de investigaciones pasadas respectivamente, es razonable esperar que ambos constructos estén relacionados entre sí. De aquí que en esta tesis se pone a prueba la siguiente hipótesis:

H6: la *satisfacción con el proceso de evaluación* se relaciona significativamente y negativamente con la *dirección de la reforma del sistema de evaluación*.

Es decir, el interés que tienen los investigadores categorizados en el sistema de evaluación de reformarlo en una dirección que cambia la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, mejor consideración de la investigación multi/inter/transdisciplinaria, y que valore más los resultados de la investigación aplicada u orientada a la resolución de problemas (*dirección de la reforma del sistema de evaluación*) depende negativamente y significativamente del nivel de *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el sistema.

Cuanto mayor sea la satisfacción que tienen los investigadores categorizados en el sistema con el proceso de evaluación que allí se realiza, porque el mismo es capaz de reconocer su desempeño, cuenta con la información necesaria para elaborar juicios académicos apropiados e incluye todas sus contribuciones importantes, garantiza evaluaciones académicas apropiadas en su área de conocimiento, respeta los equilibrios regionales e institucionales y da garantías de una evaluación objetiva, menor será el interés de estos investigadores de reformar dicho sistema de evaluación en una dirección que cambia el modelo de evaluación (hacia uno más cualitativo que cuantitativo) y sus parámetros o criterios de evaluación.

A la vez, de acuerdo con el debate internacional, el uso de indicadores bibliométricos como criterios centrales de evaluación representa una fuente de insatisfacción con los procesos de evaluación del desempeño de la investigación porque dichos criterios guardan escasa relación con los diversos productos que generan los procesos de producción de conocimiento.

Dichos indicadores no son sensibles a los objetivos y naturaleza de la investigación (Hicks et al., 2015) y por lo tanto no capturan la especificidad y diversidad de la producción de conocimiento (Hicks, 2004, DORA, 2012). Los criterios de evaluación informados por indicadores bibliométricos no consideran las especificidades de los productos y los resultados de investigación de los distintos campos científicos (Hicks, 2012), al tiempo que reconoce un tipo de producto -los artículos científicos publicados-, desconoce -o valora menos- los restantes productos también derivados de la producción de conocimiento.

Simultáneamente y de acuerdo con los hallazgos de investigaciones pasadas, los sistemas de evaluación del desempeño de la investigación informados principalmente por indicadores bibliométricos inducen a una estandarización u homogeneización de los tipos de investigación y sus productos (De Rijcke et al., 2016), donde la producción de artículos científicos en inglés y su publicación en revistas –preferentemente- con alto factor de impacto e indexadas a bases de datos bibliográficas internacionales (WoS o SCOPUS) ha sido jerarquizada y mejor valorada en la evaluación del desempeño (Leisyte y Westerheijden, 2014; De Ibarrola, 2012).

El mecanismo implícito por el cual los sistemas de evaluación centrados en indicadores bibliométricos tienden a reducir la diversidad de la producción de conocimiento se vincula con el comportamiento estratégico de los investigadores, quienes para alcanzar altas puntuaciones en los criterios de evaluación, tienden a focalizar su actividad hacia los tipos de investigación y formatos para la difusión de sus resultados que dicho sistema de evaluación reconoce y valora. Más allá de la insatisfacción que puedan generar los criterios de evaluación, los investigadores terminan adaptando su comportamiento a las presiones que impone el sistema de evaluación debido a que éstos están asociados a importantes incentivos materiales y simbólicos. De aquí, que igualmente exista una asociación positiva entre la insatisfacción con los procesos de evaluación del desempeño de la investigación -derivada de la escasa relación entre los criterios que utilizan y los diversos productos que generan los procesos de producción de conocimiento- y la estandarización u homogeneización de los tipos de investigación y productos a la que inducen esos sistemas de evaluación.

Teniendo presente que la definición conceptual que se da en esta tesis del constructo *satisfacción con el proceso de evaluación* retoma esa fuente de insatisfacción y que el nivel de diversidad de perfiles y trayectorias de investigación –en oposición a la estandarización u homogeneización de la



investigación- que permite el sistema de evaluación está contenido en la definición conceptual que se da de *flexibilidad epistemológica del sistema de evaluación*, es razonable esperar que ambos constructos estén relacionados entre sí. Es así que en esta tesis se pone a prueba la siguiente hipótesis:

H5: la *flexibilidad epistemológica del sistema de evaluación* se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.

Es decir, la *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el sistema depende positivamente y significativamente de la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores. Cuanto mayor sea la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores -porque dicho sistema permite diversidad de perfiles y de trayectorias de investigación-, mayor será la satisfacción que tienen con el proceso de evaluación porque el mismo es capaz de reconocer su desempeño, cuenta con la información necesaria para elaborar juicios académicos apropiados en su área de conocimiento e incluye todas sus contribuciones importantes.

En la misma línea argumentativa que la del debate, los hallazgos de investigaciones pasadas y la teoría, se puede sostener que en la medida que la insatisfacción con los procesos de evaluación centrados en indicadores bibliométricos tiene que ver con el hecho de que dichos criterios no capturan la especificidad y la diversidad de los tipos y productos derivados de la producción de conocimiento, es razonable esperar –simultáneamente- que la satisfacción con los procesos de evaluación varíe también dependiendo de los tipos o modos diferenciales de producción de conocimiento.

De acuerdo con la teoría, entre los atributos que caracterizan al Modo 1 de producción de conocimiento está que la difusión del conocimiento que se produce se da a través de canales formales institucionalizados en el ámbito de la ciencia académica tales como las publicaciones científicas, lo que representa la principal fuente de legitimidad y reconocimiento social del conocimiento científico producido (Gibbons et al., 1994). En concordancia con el debate internacional, la insatisfacción o cuestionamiento a los procesos de evaluación informados centralmente con indicadores bibliométricos es que solamente reconocen las publicaciones científicas, especialmente los artículos publicados en revistas científicas preferentemente con alto factor de impacto e indexadas a bases de datos bibliográficas internacionales. Los hallazgos de investigaciones pasadas documentan que dichos sistemas de evaluación han jerarquizado la producción de artículos científicos publicables en dichas revistas y han desvalorizado otros productos también derivados de la producción de conocimiento.

En la medida que los productos derivados de la producción Modo 1 de conocimiento son mejor valorados y reconocidos por los sistemas de evaluación informados por indicadores bibliométricos, es esperable entonces que los investigadores que predominantemente trabajan en ese modo de producción de conocimiento tengan una mayor satisfacción con el proceso de evaluación informado centralmente por indicadores bibliométricos, porque el mismo es capaz de reconocer su desempeño ya que valora los productos que genera. De aquí que en esta tesis se pone a prueba la siguiente hipótesis:

H2: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.

Es decir, la *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el sistema depende positivamente y significativamente de su *propensión modo 1* -en tanto tipo ideal- de producción de conocimiento de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento que realizan. Conforme aumente la *propensión modo 1* de producción de conocimiento que predominantemente orienta la generación de conocimiento que realizan los investigadores categorizados en el sistema de evaluación, mayor será su satisfacción con el proceso de evaluación.

De acuerdo con la teoría, entre los atributos que caracterizan al Modo 2 de producción de conocimiento está que la difusión del conocimiento que se produce se da a través de canales informales, en particular a través de las soluciones encontradas a los problemas dependientes del contexto o que emergen en los diversos contextos de aplicación (Gibbons et al., 1994).

De acuerdo con el debate internacional y los hallazgos de investigaciones pasadas, los sistemas de evaluación centrados en indicadores bibliométricos no son capaces de reconocer otros productos derivados de la investigación distintos de las publicaciones científicas, por ejemplo los productos/soluciones que se derivan de aquella producción de conocimiento orientada a resolver los problemas dependientes del contexto (ya sean sociales, productivos, de política pública, etc), tales como los reportes para informar la toma de decisiones de política pública, los desarrollos tecnológicos, la creación de productos técnicos, el desarrollo de técnicas e instrumentos, etc.

En la medida que los productos derivados de la producción Modo 2 de conocimiento son –en el mejor escenario- escasamente o poco valorados y reconocidos por los sistemas de evaluación informados por indicadores bibliométricos, es esperable entonces que los investigadores que predominantemente trabajan en el Modo 2 de producción de conocimiento tengan una escasa satisfacción con dicho proceso de evaluación, ya que el mismo no es capaz de reconocer su desempeño porque no valora satisfactoriamente los productos que genera ni es capaz de medir el impacto social, productivo y de política pública que puede tener este tipo o modo de producción de conocimiento. De aquí que en esta tesis se pone a prueba la siguiente hipótesis:

H4: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *satisfacción con el proceso de evaluación*.

Es decir, la *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el sistema depende negativamente y significativamente de su *propensión modo 2* -en tanto tipo ideal- de producción de conocimiento de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento que realizan. Conforme aumente la *propensión modo 2* de producción de conocimiento que predominantemente orienta la generación de conocimiento que realizan los investigadores categorizados en el sistema de evaluación, menor será su satisfacción con el proceso de evaluación informado centralmente por indicadores bibliométricos, ya que no es capaz o tiene dificultades para reconocer y valorar el desempeño de la producción Modo 2 de conocimiento.

Como se dijo, la evaluación del desempeño de la investigación incide en la dirección (Whitley, 2007), contenido y características epistemológicas de la producción de conocimiento (Gläser y Laudel, 2007b). A la vez, dependiendo del Modo de producción estará definido qué conocimiento se produce, cómo se produce, en qué contextos, cómo se organiza y difunde (Gibbons et al., 1994). Así, cada uno de esos modos de producción de conocimiento está asociado a contenidos y direcciones distintas de la producción de conocimiento.

Además, en la medida que los Modo 1 y 2 tienen prevalencias diferenciales respecto a su orientación predominante, la heterogeneidad en la combinación de conocimientos y los actores involucrados en los procesos de producción de conocimiento, los modos diferenciales de producción de conocimiento también están asociados a ciertas características epistemológicas de la producción de conocimiento. En particular, mientras que en el Modo 1 la producción de conocimiento se orienta por el avance del conocimiento disciplinar, su organización también es disciplinar e involucra únicamente a los investigadores; en el Modo 2 la producción de conocimiento avanza orientada por la búsqueda de soluciones a problemas específicos dependientes del contexto (sociales, productivos, de política pública, etc.), se organiza transdisciplinariamente –ya que los problemas que atiende son complejos y no vienen particionados en especialidades disciplinares- e involucra a actores heterogéneos (investigadores y contrapartes interesadas en la solución de problemas). Así el Modo 2 es más reflexivo o responsable socialmente que el Modo 1. En el contexto de esta tesis, los atributos que caracterizan a los modos diferenciales de producción de conocimiento son capturados a través de los constructos *propensión modo 1* y *propensión modo 2*.

Adicionalmente, como se planteó, los Modo 1 y Modo 2 requieren de sistemas de evaluación del desempeño de la investigación flexibles y contextualizados para estimular integralmente la “buena ciencia”. Es decir, sistemas de evaluación que sean capaces de valorar la calidad, desempeño e impacto del conocimiento producido en cada uno de esos modos, recompensando y reconociendo los múltiples objetivos de investigación que se plantean, los diversos productos que generan, las características propias del contexto en el que se produce el conocimiento, sus usos y la valoración que hacen de los resultados de investigación las audiencias para las cuáles ese conocimiento fue creado.

De lo contrario se pone en riesgo el desarrollo y avance de la ciencia, para Gibbons et al. (1994) porque los Modo 1 y Modo 2 de producción de conocimiento se relacionan entre sí y se dinamizan mutuamente; para Whitley (2007) porque los *sistemas de evaluación fuertes* del desempeño de la investigación, normados y estructurados en torno a los intereses disciplinarios de las élites científicas dominantes reducen la diversidad de la producción de conocimiento en términos de sus objetivos de investigación y enfoques, e inhiben el desarrollo de nuevos campos científicos, en particular de los no disciplinares o que combinan conocimientos y enfoques de diversas disciplinas.

Tal como fue señalado párrafos más arriba, de acuerdo con los hallazgos de investigaciones pasadas, los sistemas de evaluación del desempeño de la investigación, informados centralmente por indicadores bibliométricos, sugieren sesgos sistemáticos en contra de ciertas características epistemológicas de la producción de conocimiento. En particular en contra de la diversidad, los tipos diferenciales de producción de conocimiento, de la investigación inter/multi/transdisciplinar y de la investigación colaborativa. Más específicamente, dichos sistemas de evaluación exhiben sesgos sistemáticos en contra de aquella producción de conocimiento preocupada por atender los

problemas o desafíos que le impone su contexto social, económico y político y que se caracteriza por su interacción con actores externos al contexto de la ciencia académica, reforzando los estándares de excelencia disciplinares he induciendo a que los investigadores se preocupen menos por la responsabilidad social de la investigación.

En el contexto de esta tesis estas relaciones entre los sistemas de evaluación y esas características epistemológicas de la producción de conocimiento es capturada a través del constructo *flexibilidad epistemológica del sistema de evaluación*.

Estos hallazgos de investigaciones pasadas sugieren que dichos sistemas de evaluación son poco flexibles porque no permiten diversidad de perfiles y tipos de investigación, y al tiempo que parecen permitir más los atributos epistemológicos que caracterizan a la producción de conocimiento Modo 1, introducen sesgos sistemáticos en contra de los atributos epistemológicos que caracterizan a la producción Modo 2 de conocimiento. De aquí que en esta tesis se ponen a prueba las siguientes dos hipótesis:

H1: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *flexibilidad epistemológica del sistema de evaluación*

Es decir, la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el sistema de evaluación depende positivamente y significativamente de su *propensión modo 1* -en tanto tipo ideal- de producción de conocimiento. Conforme aumenta la *propensión modo 1* de producción de conocimiento que predominantemente orienta la generación de conocimiento que realizan los investigadores categorizados en el sistema de evaluación, mayor será su percepción respecto a la flexibilidad epistemológica del sistema de evaluación, ya que dicho sistema permite o facilita los atributos que caracterizan a la producción de conocimiento que realizan.

H3: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *flexibilidad epistemológica del sistema de evaluación*

Contrariamente, la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el sistema de evaluación depende negativamente y significativamente de su *propensión modo 2* -en tanto tipo ideal- de producción de conocimiento. Conforme aumente la *propensión modo 2* de producción de conocimiento que predominantemente orienta la generación de conocimiento que realizan los investigadores categorizados en el sistema de evaluación, menor será su percepción respecto a la flexibilidad epistemológica del sistema de evaluación, ya que dicho sistema impone diversas restricciones que limitan los atributos que caracterizan a la producción de conocimiento que realizan.

Adicionalmente, diversos estudios, entre ellos Müller y De Rijcke (2017), De Rijcke et al. (2016), sugieren que los sistemas de evaluación del desempeño de la investigación, informados principalmente por indicadores bibliométricos, imponen restricciones para el desarrollo de investigaciones de largo plazo (Leisyte y Westerheijden, 2014; Bensusán et al., 2014; Izquierdo, 2006), originales (Wang, Veugelers y Stephan, 2017; Bensusán et al., 2014) o disruptivas con la corriente principal (Martin y Whitley, 2010) y que la presión que ejercen sobre la productividad

medida a través de la cantidad de publicaciones en revistas indexadas en bases de datos bibliográficas internacionales (WoS o SCOPUS) induce a que los investigadores descuiden la calidad (Butler, 2003) de la producción de conocimiento por cumplir con el requisito de la cantidad y desarrollen diversas estrategias para maximizar el número de publicaciones (Leisyte y Westerheijden, 2014; Bensusán et al., 2014; van Dalen y Henkes, 2012; Heras, 2005).

Importa señalar que los signos de las relaciones establecidas en las últimas dos hipótesis (H1 y H3) son muy arriesgados. Por un lado, porque el constructo *flexibilidad epistemológica del sistema de evaluación* también captura las relaciones -descritas en el párrafo anterior- entre los sistemas de evaluación y esas otras características epistemológicas de la producción de conocimiento - investigación de largo plazo, original, relevante conceptualmente y confiable (calidad)-. Por otro lado, porque no es razonable suponer teóricamente ni existe evidencia que permita asignar prevalencias distintas de estas características epistemológicas a los modos diferenciales de producción de conocimiento. Es decir, dichas características están asociadas indistintamente a los dos modos de producción de conocimiento y por lo tanto eso podría alterar los signos de las relaciones hipotetizadas.

De todas maneras, si se encuentra evidencia que demuestre que los investigadores que predominantemente orientan su investigación en el Modo 1 de producción de conocimiento, perciben una mayor *flexibilidad epistemológica del sistema de evaluación* que los que lo hacen en el Modo 2, entonces esos resultados sugieren que el sistema de evaluación introduce sesgos que favorecen más la producción Modo 1 de conocimiento, que la que predominantemente se orienta en Modo 2.

En términos de Dahler-Larsen (2014) se puede argumentar que esos resultados sugieren que el sistema de evaluación genera *efectos constitutivos* que limitan la producción Modo 2 de conocimiento, ya que no es igualmente flexible desde el punto de vista de las características epistemológicas de la producción de conocimiento que permite. En términos de Gibbons et al. (1994) se puede argumentar que esos resultados sugieren que el sistema de evaluación está más alineado con las normas cognitivas y sociales que rigen la producción de conocimiento que se realiza en el contexto de la ciencia académica disciplinar -Modo 1-, que la que se realiza en el contexto de aplicación -Modo 2-.

Finalmente, de acuerdo con la teoría, el alcance de la incidencia de la evaluación del desempeño de la investigación sobre la dirección, contenido y características epistemológicas de la producción de conocimiento está mediado por el conjunto de características específicas que asuman tres dimensiones contextuales: i) el sistema público de ciencia; ii) el sistema de evaluación del desempeño de la investigación, en particular si se trata de un *sistema de evaluación fuerte*; y iii) los diversos campos científicos. (Whitley, 2007)

Influida por esas ideas, en esta tesis se retoman dos de las tres dimensiones propuestas por Whitley (2007) para alcanzar una comprensión contextualizada de la naturaleza de las relaciones hipotetizadas entre un sistema de evaluación fuerte como el SNI y los Modo 1 y Modo 2 de producción de conocimiento que predominantemente orientan el trabajo de los investigadores categorizados en ese sistema.

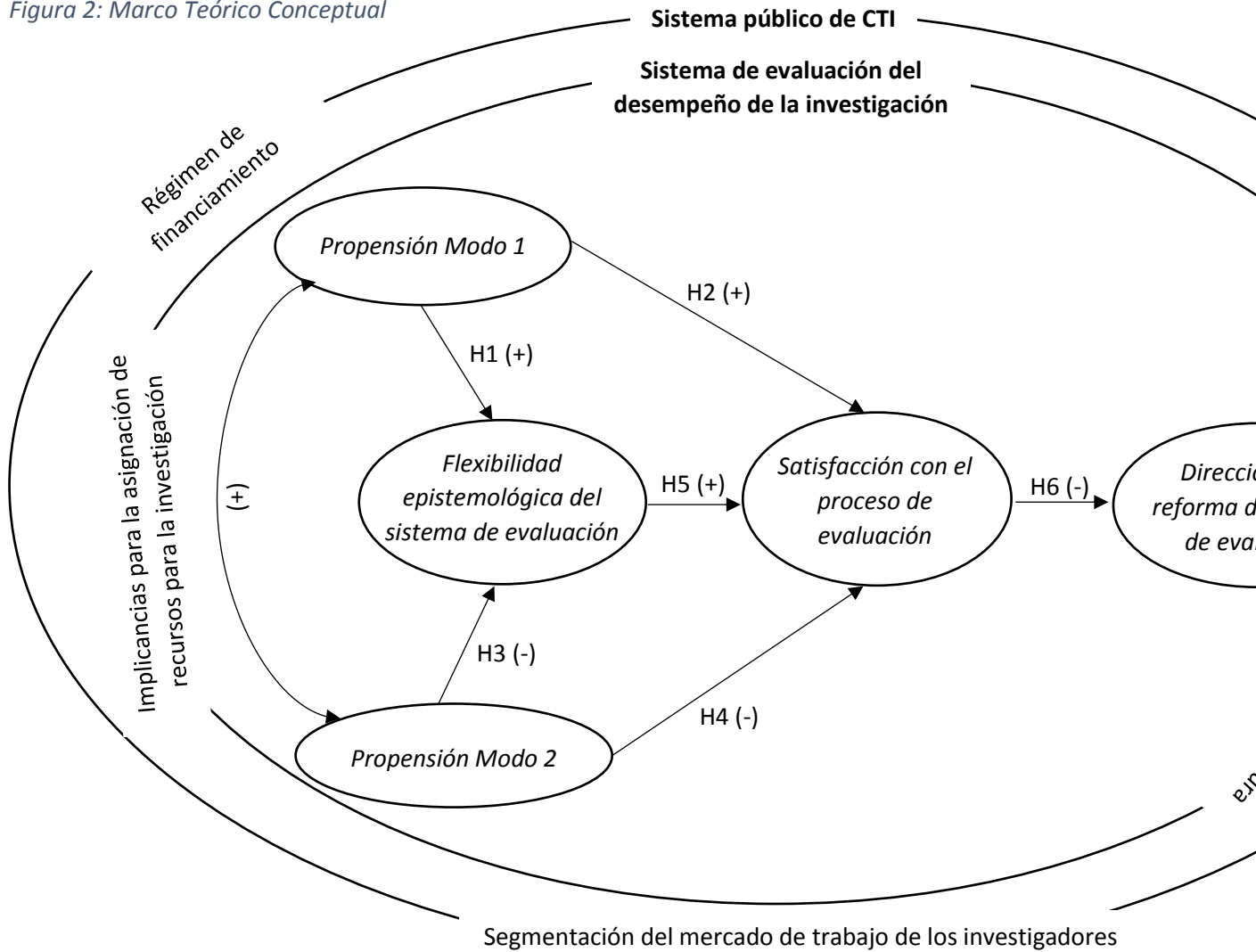
En esta tesis se caracterizan el sistema público de CTI de México -en términos de su régimen de financiamiento y la segmentación de las organizaciones productoras de conocimiento y del mercado de trabajo de los investigadores-, y el sistema de evaluación -en términos de su gobernanza y estructura y sus implicancias en la asignación de recursos para la investigación-. Dado el alcance de la investigación de esta tesis y que una caracterización de los campos científicos -como la que propone Whitley (2007)- para el caso del SNI supone la realización de múltiples esfuerzos de investigación en sí mismos, abordar esta dimensión contextual es materialmente imposible en el marco de esta investigación. Sin embargo, debido a la importancia de los campos científicos en las relaciones entre la evaluación de la investigación y la producción de conocimiento, en esta tesis se controlan los resultados de las estimaciones a través de la introducción en el modelo de un conjunto de variables independientes indicadoras del área de conocimiento. En el capítulo VI se caracterizan el sistema público de CTI de México y el SNI.

De acuerdo con la caracterización realizada del sistema público de CTI de México y del SNI, y siguiendo a la teoría, es razonable esperar consecuencias notables del SNI en el contenido y dirección de la producción de conocimiento mexicana, en particular en relación a sus modos diferenciales de producción de conocimiento. Si la teoría es cierta, entonces se deberían observar relaciones significativas entre los diversos constructos vinculados a la evaluación del desempeño de la investigación que se realiza en el SNI y los modos diferenciales de producción de conocimiento.

En la figura 2 se presenta gráficamente el marco teórico conceptual desarrollado a lo largo de este capítulo, se representan los conceptos relevantes para el análisis y sus relaciones teóricas hipotetizadas. Conjuntamente representan la teoría estructural que se propone y pone a prueba en esta tesis.

En el capítulo siguiente se caracteriza al sistema de CTI de México y al SNI en él. Simultáneamente, se detallan las características específicas de su diseño y sus implicancias para la asignación de los recursos para la investigación en el país. Posteriormente, en el capítulo VII, con base en la sistematización de dos grandes debates sobre el SNI, se retoman y profundizan los temas aquí tratados y otros que de allí emergen.

Figura 2: Marco Teórico Conceptual



Fuente: Elaboración propia

Notas: Las dos elipses exteriores representan las características específicas contextuales del Sistema público de CTI y del Sistema de evaluación. Las elipses interiores contienen los constructos o dimensiones conceptuales relevantes para el análisis, las flechas unidireccionales representan las relaciones de dependencia entre constructos y entre paréntesis el signo esperado de la relación. La flecha bidireccional representa la asociación positiva entre los dos modos de

## CAPITULO VI: Los sistemas de ciencia, tecnología e innovación y de evaluación a la investigación individual en México

De acuerdo con lo planteado en la sección III.3 del capítulo III, los cambios que se observen en la organización y producción de conocimiento asociados a la institucionalización de ciertos sistemas de evaluación de la investigación se ven reforzados o, contrariamente, amortiguados dependiendo de un conjunto de características contextuales-específicas que según Whitley (2007) se relacionan con: i) el sistema CTI; ii) el sistema de evaluación de la investigación; y iii) los campos científicos.

En ese sentido y con el objetivo de contextualizar las percepciones de los investigadores categorizados en el SNI sobre la satisfacción que tienen con el proceso de evaluación de dicho sistema, su flexibilidad epistemológica y la dirección de una eventual reforma al mismo, a continuación y a lo largo de este capítulo, se describen en primer lugar ciertas características específicas del sistema de CTI de México (sección VI.1), y en segundo lugar, las asociadas al sistema de evaluación del SNI<sup>27</sup> (sección VI.2). La sección VI.3 concluye.

En particular, y en relación al sistema de CTI de México, la caracterización que se realiza en este capítulo se centra en dos dimensiones contextuales. La primera, y siguiendo a Whitley (2007), se vincula con la *naturaleza del régimen de financiamiento de la investigación*. Específicamente con su alcance y periodicidad, con la incorporación o no de los objetivos de la política pública en el financiamiento a la investigación (a través por ejemplo de fondos específicos para el fomento de la investigación orientada a atender los principales problemas del desarrollo económico y social de México), y con la variedad de agencias de financiamiento (sub-sección VI.1.1). La segunda dimensión contextual del sistema de CTI se asocia con *el mercado de trabajo* de los investigadores (Whitley, 2007) en México (sub-sección VI.1.2).

Respecto a las características específicas del sistema de evaluación, en particular del SNI, la caracterización que se realiza se enfoca en dos dimensiones contextuales: la gobernanza y estructura, y las implicancias del sistema de evaluación en la asignación de recursos. Los elementos que interesan para esta tesis en la dimensión de la *gobernanza y estructura* (Whitley, 2007) del SNI se vinculan con la frecuencia o periodicidad de la evaluación, el grado de formalización de sus procedimientos, la estandarización –o no- de criterios de evaluación comunes entre las áreas de conocimiento y su estabilidad en el tiempo, el nivel de transparencia del sistema, y su unidad de evaluación. La segunda dimensión tiene que ver con las implicancias que tiene dicho sistema de evaluación para la asignación de recursos a la ciencia. Estas dimensiones contextuales específicas se

---

<sup>27</sup> Como se dijo, contextualizar las percepciones de los investigadores categorizados en el SNI de acuerdo con las características específicas de los campos científicos que propone Whitley (2007) no es factible en el contexto de este trabajo y además excede largamente los objetivos de investigación planteados. Sin embargo, y debido a que de acuerdo con la literatura sí son esperables variaciones en las percepciones de los investigadores dependiendo de los campos científicos en los que trabajan, en la estimación del modelo de ecuaciones estructurales se introducen un conjunto de variables independientes indicadoras del área de conocimiento como una forma de aproximar el campo científico. Estas variables independientes son introducidas –justamente- para controlar sus potenciales efectos perturbadores sobre las dimensiones de análisis de interés (satisfacción con el proceso de evaluación, flexibilidad epistemológica del sistema de evaluación y la dirección de una eventual reforma al mismo). (Ver Capítulo X sobre modelos de medida y estructural, sección X.5 el modelo teórico estructural condicional).



abordan en base a la información contenida en el Acuerdo de creación del SNI y sus modificaciones (sub-sección VI.2.1), en los Reglamentos de operación y funcionamiento del sistema (sub-sección VI.2.2) y en los Criterios Internos de evaluación de cada una de las Comisiones Dictaminadoras por áreas de conocimiento (sub-sección VI.2.3).

En última instancia se trata de caracterizar al sistema de evaluación del SNI, de forma tal que pueda ser asociado, ya sea, a un *sistema de evaluación fuerte* o, por el contrario, a uno débil (Whitley, 2007), ya que de ello dependen las eventuales consecuencias del SNI sobre la dirección y contenido de la producción de conocimiento. A la vez, dichas consecuencias se verán reforzadas o, por el contrario, amortiguadas, dependiendo de ciertas características específicas del sistema de CTI de México. De alguna manera, este enfoque analítico permite contextualizar los resultados de este trabajo, y también, parte de los principales problemas y/o desafíos identificados por los estudios realizados sobre el SNI de México.

## VI.1 El sistema de Ciencia, Tecnología e Innovación en México: algunas características importantes

### VI.1.1 La producción de conocimiento y su financiamiento

En la literatura está ampliamente reconocido la importancia de la CTI para el crecimiento económico. Sin embargo, si bien el desempeño de la economía de un país depende fuertemente de las capacidades que éste tenga para desarrollar CTI; su dinamismo y sostenibilidad dependen simultáneamente del uso, absorción y/o aplicación de esas capacidades con fines productivos.

Tal como señalan Dutrénit et al. (2010), las actividades económicas más dinámicas y que generan mejores y mayores pagos a los factores de producción son las que incorporan CTI, las de mayor productividad (Dutrénit et al., 2013; De Fuentes et al., 2015). Así se generan círculos virtuosos entre CTI y producción de bienes y servicios; conforme aumenta la productividad -derivada de la agregación de valor basada en el conocimiento científico y tecnológico-, mayor es la demanda, que a su vez se expresa en la capacidad de inversión que tiene el sector productivo –las empresas- en la creación de capacidades de CTI. Dicho de otro modo, los sectores económicos dinámicos y de alta productividad tendrán mayores capacidades de financiamiento de las actividades de investigación científica y tecnológica.

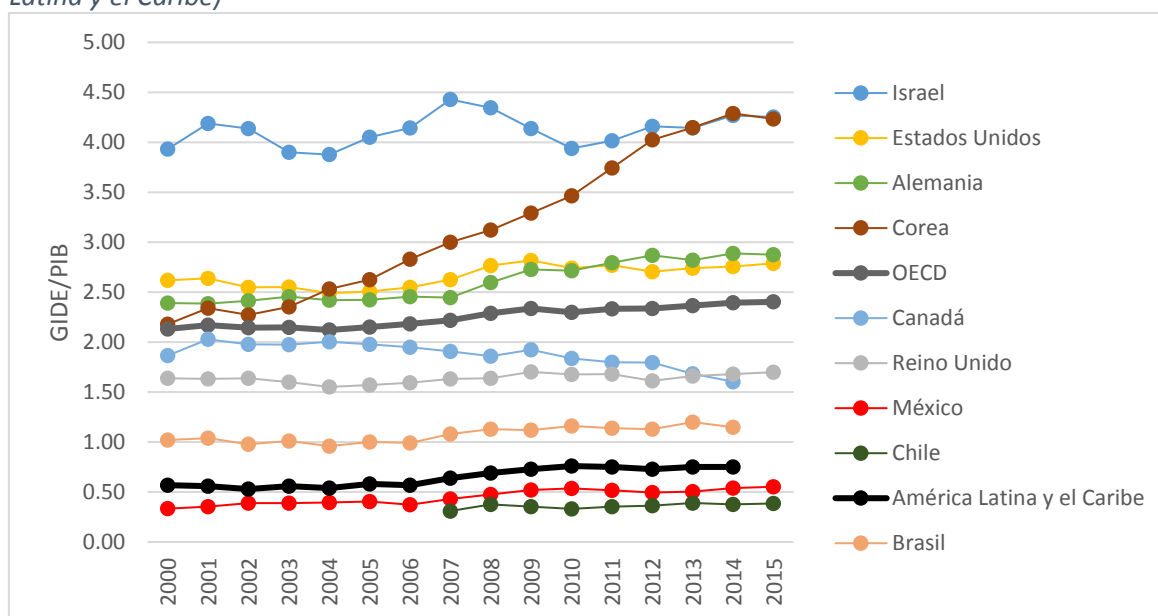
Sin embargo, históricamente en México -al igual que en otros países de América Latina y el Caribe-, el Sistema Nacional de Innovación o, como se expresa en los diversos PECiTI (2008-2012; 2014-2018), el Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI) presenta una débil articulación entre los actores que lo componen. Incluso desde antes, en el PECyT (2001-2006) y como parte del diagnóstico que da fundamento a ese plan estratégico de Ciencia y Tecnología, se indicaba que el sistema era reducido y estaba representado por la agregación de un conjunto de instituciones y organizaciones (públicas y privadas) que desarrollaban sus actividades en un ambiente escasamente articulado. Seis años después, Dutrénit et al. (2010) sostienen que más allá de ciertos logros importantes, el sistema –estructuralmente- continua inalterado: escasos niveles de interacción y articulación entre sus agentes, capacidad limitada para producir conocimiento sumado a una también muy baja demanda de conocimiento por parte de los agentes locales.

De acuerdo con el PECiTI (2014-2018) y desde el punto de vista de los recursos que México asigna al desarrollo de sus capacidades en CTI, en el año 2012, el Gasto Nacional en Ciencia Tecnología e

Innovación (GNCTI) fue de 114,474.1 millones de pesos, lo que representó un 0.74% del PIB. De ese total, el 58.3% estuvo dirigido al financiamiento de la investigación y desarrollo experimental (IDE), el 20.6% fue para financiar la educación de posgrado, el 19.1% para los servicios científicos y tecnológicos, y apenas un 2% para apoyar a la innovación. La distribución del GNCTI en 2014 fue algo distinta a la de 2012, ya que perdió peso relativo el gasto en investigación y desarrollo experimental (GIDE) -ubicándose en un 51%- , a favor de un aumento en la participación del gasto en innovación, que alcanzó un 9% del GNCTI. Simultáneamente, la participación de los gastos asociados a los servicios científicos y tecnológicos y la educación de posgrado se mantuvieron constantes. (CONACYT, 2014)

De acuerdo con los datos reportados por la OECD (2017) y la Red de Indicadores de Ciencia y Tecnología (RICYT)<sup>28</sup>, el GIDE en relación al PIB de México se ubica por debajo del promedio de los países que integran la OECD y también por debajo del promedio de los países de América Latina y el Caribe (ver Gráfico 1). En el período, el GIDE en relación al PIB de México pasó de 0.33% en el año 2000, al 0.55% en 2015. Entre los países que componen la OECD, México se ubica únicamente apenas por encima de Chile; y siempre por debajo de la meta del 1% del PIB planteada en los programas especiales de gobierno.

Gráfico 1: GIDE en relación PIB (Algunos países seleccionados, promedio de OECD y de América Latina y el Caribe)



Fuente: Elaboración propia en base a *Main Science and Technology Indicators* (OECD, 2017) y RICYT.

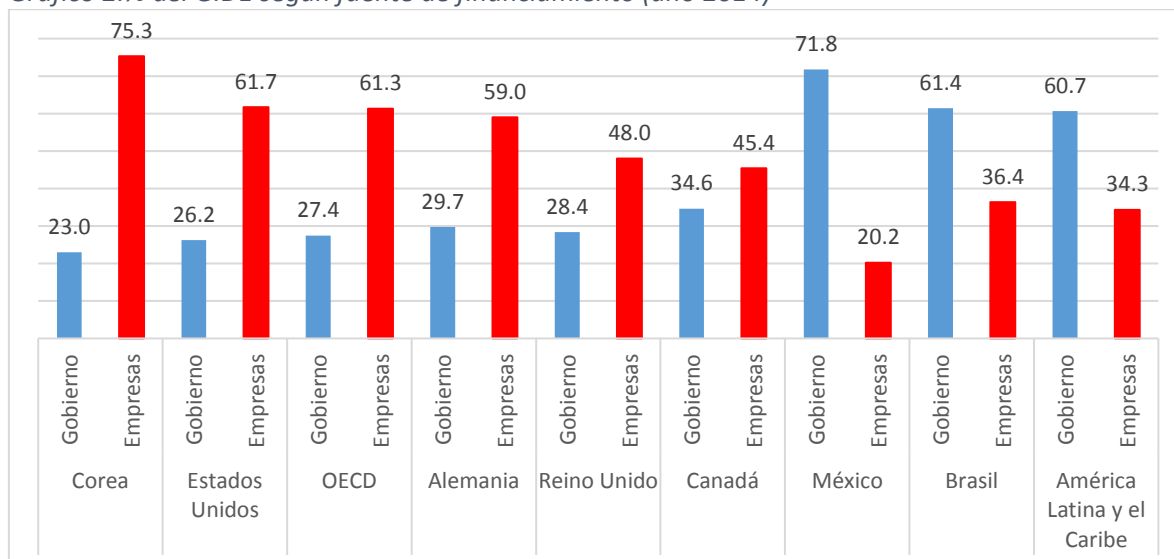
Simultáneamente, y desde el punto de vista de las fuentes de financiamiento del GIDE, es posible observar una estructura diferencial entre los países desarrollados y en desarrollo. Mientras que en Corea, Estados Unidos, Alemania, Reino Unido y Canadá son las empresas del sector industrial las que financian la mayor parte del gasto en I+D de esos países, la principal fuente de financiamiento

<sup>28</sup> Disponibles en

<http://db.ricyt.org/query/AR,BO,BR,CA,CL,CO,CR,CU,EC,ES,GT,HN,JM,MX,NI,PA,PE,PR,PT,PY,SV,TT,US,UY,VE,AL,IB/1990%2C2014/GASTOxPBI>. Fecha de consulta 31/05/2017.

del GIDE en México y Brasil es pública y corresponde al Gobierno (71.8% y 61.4% respectivamente). En particular, al Gobierno Federal en el caso de México (CONACYT, 2014:37). El gráfico 2 muestra la distribución del GIDE según las principales fuentes de financiamiento para algunos países seleccionados. Importa señalar que si bien los datos del gráfico 2 son para el año 2014, la estructura de financiamiento se mantiene relativamente inalterada desde el año 2010, incluso desde antes, ver *Main Science and Technology Indicators* (OECD, 2017)<sup>29</sup> y RICYT<sup>30</sup>.

Gráfico 2: % del GIDE según fuente de financiamiento (año 2014)



Fuente: Elaboración propia en base a *Main Science and Technology Indicators* (OECD, 2017) y RICYT.

Debido a que el GIDE de México es financiado fundamentalmente por el Gobierno Federal, a continuación se presenta en la tabla 2 la distribución del Gasto Federal en Ciencia y Tecnología (GFCyT) para el período 2005-2014, según la participación de los diversos sectores en su distribución. Tal como puede observarse en la tabla 2, cerca de las dos terceras partes del GFCyT en el 2014 se destinaron al presupuesto del CONACYT (40.3%) y al ejercido por la Secretaría de Educación Superior en CTI (24.0%). En particular, aproximadamente un 70% del GFCyT destinado al CONACYT es para el estímulo y financiamiento de las capacidades científicas y tecnológicas del país, a través de la implementación de fondos competitivos en los diversos instrumentos de fomento de las actividades de CTI, y el restante 30% es para el financiamiento de los Centros Públicos de Investigación (CPI) del CONACYT. (CONACYT, 2014)

Más específicamente, un análisis detallado del financiamiento a diversos instrumentos de fomento a las capacidades de CTI en 2016 muestra que el 26.8% del presupuesto del CONACYT se destinó al financiamiento de becas de formación de recursos humanos (maestría y doctorado), el 12.0% al financiamiento del SNI, el 14.1% al Programa de Estímulos a la Innovación, aproximadamente el

<sup>29</sup> Disponible en [http://www.oecd-ilibrary.org/science-and-technology/main-science-and-technology-indicators-volume-2016-issue-2/percentage-of-gross-domestic-expenditure-on-r-amp-d-gerd-financed-by-industry\\_msti-v2016-2-table13-en](http://www.oecd-ilibrary.org/science-and-technology/main-science-and-technology-indicators-volume-2016-issue-2/percentage-of-gross-domestic-expenditure-on-r-amp-d-gerd-financed-by-industry_msti-v2016-2-table13-en). Fecha de consulta 31/05/2017.

<sup>30</sup> Disponible en <http://db.ricyt.org/query/AR,BO,BR,CA,CL,CO,CR,CU,EC,ES,GT,MX,PA,PT,PY,SV,US,UY,AL,IB/1990%2C2014/GASIDSPER>. Fecha de consulta 31/05/2017.

10.0% corresponde a gastos de operación (FCCyT, 2016).

En México, los recursos para el financiamiento de la investigación provienen esencialmente del sector público, y dentro de éste, principalmente del CONACYT. A la vez, y como se mostró, son relativamente escasos y, por lo mismo, altamente competitivos. Dichos recursos son asignados –en su mayoría– a través de proyectos de investigación que van de 1 a 3 años de duración, presentados a las diversas convocatorias en los distintos instrumentos de fomento a la investigación científica y tecnológica del CONACYT. En general, para acceder a esos recursos para la investigación, se valora positivamente que el responsable de la propuesta de investigación esté categorizado en el SNI como *investigador nacional*<sup>31</sup>.

Además, la mayoría de los recursos públicos para el fomento de la CTI, se dirigen a la formación y estímulo de recursos humanos altamente calificados (Becas y SNI), una parte va al apoyo de la investigación básica o fundamental, una parte se destina a fondos sectoriales orientados a temáticas de las Secretarías de Estado, y una pequeña proporción está destinada al fomento de la investigación orientada a atender los principales problemas del desarrollo social y económico de México. Finalmente, existe una muy baja variedad de agencias de financiamiento de la producción de conocimiento en México, de hecho, el grueso de los recursos destinados a la investigación científica y tecnológica del país, se concentran en una única agencia de financiamiento, que es el CONACYT.

En ese sentido, y debido a que los recursos para la investigación son asignados fundamentalmente a través de proyectos de investigación competitivos, de 1 a 3 años de duración y preferentemente presentados por investigadores SNI, de acuerdo con Whitley (2007), es razonable esperar un horizonte temporal de la investigación de corto plazo.

Por otra parte, si bien la asignación de recursos por proyectos facilita la orientación de la investigación hacia la atención de los objetivos de política pública (Whitley, 2007), en México sólo una porción menor del presupuesto del CONACYT está destinada al fomento de la investigación orientada a atender los principales problemas del desarrollo social y económico del país -aunque este es un objetivo explícito de política pública en CTI-. El trabajo de Dutrénit et al. (2010) realiza un análisis detallado, que permite observar esta inconsistencia entre el diseño de la política pública en CTI y su implementación.

Finalmente, debido a la escasa variedad de agencias u organizaciones del ámbito de la producción de bienes y servicios u otras organizaciones orientadas al financiamiento de la investigación, es esperable de acuerdo con Whitley (2007) y como sucede en México, por un lado, una alta dependencia de los fondos públicos competitivos de investigación, y por otro, una menor diversidad en la orientación y objetivos de la investigación. La mayoría de la producción de conocimiento en México parece concentrarse en investigación con orientación fundamental o básica, y con objetivos que permitan la obtención de resultados de investigación en el corto plazo.

---

<sup>31</sup> Ver Términos de Referencia de las diversas convocatorias del CONACYT. Disponible en <http://www.conacyt.gob.mx/index.php/el-conacyt/convocatorias-y-resultados-conacyt>. Fecha de consulta 31/05/2017.

Tabla 2:: Distribución del Gasto Federal en Ciencia y Tecnología (GFCyT) según participación sectorial (2005-2014)

Sectores	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
CONACYT	29.2%	30.9%	30.6%	31.8%	36.8%	34.9%	35.4%	36.0%	40.3%	40.3%
Educación Pública	36.6%	35.7%	33.8%	29.4%	29.4%	29.1%	27.6%	29.0%	25.8%	24.0%
Energía	16.9%	14.9%	14.8%	15.2%	13.0%	17.6%	18.3%	17.3%	15.6%	15.2%
Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación	5.5%	6.3%	6.5%	5.8%	5.6%	4.7%	4.1%	4.9%	4.9%	8.5%
Salud y Seguridad Social	6.2%	6.1%	7.3%	9.3%	9.2%	7.5%	8.9%	7.1%	8.6%	7.3%
Economía	2.6%	3.3%	4.1%	5.3%	3.2%	3.3%	3.5%	2.7%	2.3%	2.5%
Medio Ambiente y RRNN	1.8%	1.7%	1.7%	1.3%	1.4%	1.4%	0.9%	1.2%	0.9%	1.0%
Marina	0.6%	0.6%	0.7%	0.9%	0.8%	0.7%	0.8%	1.0%	0.8%	0.5%
Comunicaciones y Transporte	0.3%	0.4%	0.3%	0.4%	0.2%	0.3%	0.3%	0.3%	0.3%	0.4%
Otros	0.2%	0.1%	0.2%	0.3%	0.2%	0.4%	0.1%	0.3%	0.3%	0.2%
Procuraduría General de la República	0.0%	0.0%	0.0%	0.2%	0.2%	0.2%	0.2%	0.3%	0.2%	0.2%
<b>Total</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>

Fuente: Elaboración propia en base a CONACYT (2014).

#### VI.1.2 El mercado de trabajo de los investigadores

La estructura del mercado de trabajo de los investigadores está poco segmentada y fundamentalmente concentrada en las instituciones de educación superior de México.

En el período comprendido entre los años 2010 y 2013, el número promedio de investigadores fue de 48,664, lo que equivale a 34,709 investigadores de tiempo completo, de los cuales en el año 2010, el 48.6% era empleado por alguna institución de educación superior, el 23.7% trabajaba en alguna dependencia del Gobierno y el restante 27.7% desarrollaba sus actividades en empresas del ámbito de la producción de bienes y servicios. Hacia el año 2013 dicha estructura de empleo se mantuvo relativamente inalterada, con un 47.6% de investigadores tiempo completo trabajando en instituciones de educación superior, un 30.5% en dependencias del Gobierno y el restante 23.8% en empresas del sector productivo (ver tabla 3). (OECD, 2017)

A la vez, el tamaño de la comunidad científica de México es muy pequeña, con valores similares a los de los restantes países de América Latina y el Caribe, pero con valores muy por debajo si se los compara con el promedio de los países que integran la OECD. Mientras que, para el período 2010-2013, la proporción promedio de investigadores de tiempo completo cada 1000 integrantes de la PEA (Población Económicamente Activa) fue de 7.7 en los países de la OECD, en México esa proporción no alcanza a un investigador por cada 1000 integrantes de la PEA de ese país (ver tabla 3).

Tabla 3: Investigadores en México según ámbito de empleo y otras medidas relativas

	2010	2011	2012	2013	2010	2011	2012	2013
	Número				Porcentaje			
Educación Superior	18,668	18,881	13,850	14,422	48.6%	47.4%	45.8%	47.6%
Gobierno	9,126	9,279	9,219	8,582	23.7%	23.3%	30.5%	28.3%
Empresas	10,641	11,652	7,194	7,323	27.7%	29.3%	23.8%	24.1%
<b>Total (Equivalente Jornada Completa –EJC-)</b>	<b>38,435</b>	<b>39,812</b>	<b>30,263</b>	<b>30,327</b>	100.0%	100.0%	100.0%	100.0%
<b>Total investigadores SNI<sup>4</sup></b>	<b>16,519</b>	<b>17,568</b>	<b>18,476</b>	<b>19,634</b>	(30.3%)	(31.1%)	(44.6%)	(46.5%)
<b>Total (Individuos)</b>	<b>54,532</b>	<b>56,481</b>	<b>41,419</b>	<b>42,222</b>				
Investigadores (EJC) en México por cada 1000 integrantes de la PEA <sup>1</sup>	0.8	0.8	0.6	0.6				
Investigadores (EJC) en ALC <sup>2</sup> por cada 1000 integrantes de la PEA	0.75	0.8	0.78	0.81				
Investigadores (EJC) en OECD <sup>3</sup> por cada 1000 integrantes de la PEA	7.5	7.7	7.8	7.9				

Fuente: Elaboración propia en base a *Main Science and Technology Indicators* (OECD, 2017), RICYT y Bensusán et al. (2014:15). Notas: 1/La PEA es la Población Económicamente Activa. 2/Promedio de países de América Latina y el Caribe (ALC). 3/Promedio de países que componen la OECD. 4/Incluye a los Candidatos y entre paréntesis la proporción de investigadores SNI en relación al total de investigadores (individuos).

Si se consideran a los investigadores categorizados en el SNI (incluidos los candidatos), el tamaño de la comunidad científica no sólo se reduce significativamente sino que se concentra aún más en las instituciones de educación superior. De hecho, la proporción de investigadores SNI respecto al total de investigadores no llega ni a la mitad, pasando de 30.3% en el año 2010 a un 46.5% en el año 2013 (ver tabla 3). Cabe destacar que los datos del número de investigadores categorizados en el SNI en el 2017 ascendió a 27066. En el año 2014, el 71% de los investigadores categorizados en el SNI estaba empleado en alguna institución de educación superior, el 10% en dependencias del Gobierno, el 8% en los CPI, el 3.5% en el CINVESTAV (Centro de Investigación y de Estudios Avanzados), el 3% en instituciones extranjeras, apenas un 0.5% en empresas y el restante 4% en otras instituciones sin clasificar (CONACYT, 2014:62). Más aún, en el año 2014, el 46% de los investigadores categorizados en el SNI se concentran únicamente en 10 instituciones (CONACYT, 2014:62), la amplia mayoría de ellas corresponden a universidades, siendo la Universidad Nacional Autónoma de México (UNAM) la que lidera, concentrando al 18.5% de los investigadores SNI, seguida de la Universidad Autónoma Metropolitana (UAM) con un 5%.

Dada la alta concentración del mercado de trabajo de los investigadores, en particular en las instituciones de educación superior y dentro de éstas, en las universidades públicas, sumado al hecho de que en los últimos años la participación de los investigadores SNI se ha incrementado en el total de investigadores de México; de acuerdo con Whitley (2007), resulta razonable esperar una menor diversidad de los criterios para evaluar el desempeño de los investigadores, una mayor convergencia y homogeneización de sus trayectorias académicas y una mayor competencia entre ellos para acceder a los puestos de trabajo que están concentrados en el mismo mercado.

De hecho, el trabajo realizado por Bensusán et al. (2014) señala que las tres cuartas partes de los investigadores del SNI considera que las instancias de evaluación externas son complementarias a las de su institución de adscripción, y además, el 58% indica que los criterios de evaluación aplicados

externamente como al interior de sus instituciones, no son contradictorios entre sí (Bensusán et al., 2014:71).

## VI.2 El sistema de evaluación: creación y evolución del SNI desde su marco regulatorio

### VI.2.1 Acuerdo de creación del SNI y sus modificaciones

En 1983 en una reunión organizada por la actual Academia Mexicana de Ciencias (AMC), en aquel momento Academia de la Investigación Científica (AIC), se discutía la necesidad de apoyar el trabajo de investigación realizado por los investigadores mexicanos para que pudieran continuar y desarrollar sus tareas en un contexto caracterizado por la enorme crisis económica y social que atravesaba México en esa época, que además alentaba la fuga de cerebros. Es a partir de este momento, a instancias de los investigadores nucleados en la AIC que tenían vínculos con la función pública y el gobierno federal, que se observan las primeras ideas e iniciativas orientadas a la creación del SNI de México. Mismas que culminan a fines de ese año con el anuncio del presidente del país, Lic. Miguel de la Madrid, sobre la creación del SNI y le encarga a la AIC la elaboración de sus estatutos. (FCCyT y AMC, 2005).

El 26 de julio de 1984 se publica en el Diario Oficial de la Federación el Acuerdo Presidencial por el que se establece el SNI<sup>32</sup>. Entre las consideraciones que fundamentan el Acuerdo para crear el SNI están: i) el reconocimiento de la importancia de la ciencia y la tecnología como determinante central para mejorar la calidad de vida y la independencia tecnológica; ii) la convicción de que es solo a través de la investigación científica, tecnológica y humanística que se puede contribuir y acceder al conocimiento, y que además son centrales en la estructura económica del país por sus capacidades transformadoras; iii) la inscripción del SNI como un instrumento fundamental para promover la política de desarrollo científico y tecnológico del país, establecida en el marco más general del Plan Nacional de Desarrollo de la época; y, iv) el entendimiento de que para alcanzar investigación de calidad es necesario establecer un sistema que estimule a los investigadores, tanto a los que son reconocidos por su experiencia y trayectoria como a los que se inician en dicha actividad. Son seis los objetivos establecidos en el Acuerdo de creación del SNI, a saber: i) fomentar el desarrollo científico y tecnológico y el fortalecimiento de la investigación en todas sus áreas mediante el apoyo a los investigadores de las instituciones de educación superior (IES) y de investigación del sector público; ii) aumentar el número y el nivel de los investigadores de tiempo completo; iii) promover la investigación eficiente y de calidad; iv) estimular la investigación del sector público según las prioridades establecidas en el Plan Nacional de Desarrollo; v) propiciar la conformación de grupos de investigación en todo el territorio mexicano; y, vi) contribuir al desarrollo de sistemas nacionales de información científica y tecnológica integrados y por disciplinas.

Desde el punto de vista de la estructura del sistema, el Acuerdo de creación del SNI formaliza sus procedimientos indicando que tendrá un Consejo Directivo cuyas funciones serán generar políticas y programas para el SNI alineados con los objetivos y prioridades del Plan Nacional de Desarrollo y el Programa Nacional de Desarrollo Tecnológico y Científico; tomar decisiones sobre las propuestas del Secretariado Técnico; supervisar los mecanismos de evaluación y la operación del SNI; aprobar los criterios de evaluación que se aplicarán y el reglamento que determina la organización y funcionamiento del sistema. Dicho Consejo Directivo estará integrado por el Secretario de

---

<sup>32</sup> Disponible en [http://www.dof.gob.mx/nota\\_detalle.php?codigo=4680072&fecha=26/07/1984](http://www.dof.gob.mx/nota_detalle.php?codigo=4680072&fecha=26/07/1984). Fecha de consulta 12/05/17.

Educación Pública que será su Presidente; el Director General del CONACYT que será su Vicepresidente, y tres investigadores –nivel III del SNI- consejeros y uno de ellos actuará como Secretario del Consejo. Los consejeros serán designados por el Presidente del Consejo Directivo, renovados cada tres años y pueden ser designados nuevamente una sola vez. Al menos uno de los consejeros tiene que desarrollar sus actividades fuera del área metropolitana del Valle de México.

Además del Consejo Directivo, el SNI también tendrá un Secretariado Técnico, cuyas funciones serán: recibir las postulaciones al SNI y redireccionarlas a las Comisiones Dictaminadoras para su evaluación; coordinar las actividades de las Comisiones Dictaminadoras; poner a consideración del Consejo Directivo los dictámenes realizados por las Comisiones Dictaminadoras; informar al Consejo Directivo sobre los mecanismos de evaluación y operación del SNI; realizar convocatorias anuales para la selección de los *investigadores nacionales*; y, elaborar la propuesta de reglamento del SNI y ponerla a consideración del Consejo Directivo. Este Secretariado Técnico estará integrado por los Subsecretarios de Planeación Educativa, de Educación e Investigación Tecnológica y de Educación Superior e Investigación Científica de la SEP, por el Secretario General del CONACYT y un miembro propuesto por la AIC.

Asimismo, el Acuerdo crea tres Comisiones Dictaminadoras, una para la evaluación de las postulaciones de los investigadores provenientes de las áreas de Ciencias, Fisicomatemáticas e Ingenierías; otra para las postulaciones de Ciencias Biológicas, Biomédicas, Agropecuarias y Químicas, y una tercer Comisión Dictaminadora para la evaluación de los investigadores de las áreas de Ciencias Sociales y Humanidades. Las Comisiones Dictaminadoras estaban integradas por nueve *investigadores nacionales* Nivel III, cuatro de ellos a propuesta de la AIC y todos designados por el Consejo Directivo. Estaba previsto que todos los integrantes de las Comisiones Dictaminadoras se desempeñarían en el cargo durante cuatro años. Las Comisiones Dictaminadoras tenían a su cargo la evaluación de los méritos académicos de los investigadores aspirantes, y en función de ello, proponía al Consejo Directivo –a través del Secretariado Técnico- el nivel de categorización en el SNI para cada uno de los postulantes.

Según el Acuerdo de creación del SNI, los criterios de evaluación específicos por disciplina eran establecidos a propuesta de cada una de las tres Comisiones Dictaminadoras y debían ser publicados y dados a conocer en cada convocatoria anual del SNI. En términos generales<sup>33</sup> y según lo establecido en el Acuerdo de creación del sistema, los criterios de evaluación para determinar la inclusión o no de los investigadores en el SNI consideraban por un lado la productividad del postulante respecto a sus trabajos y la formación de investigadores y de personal de alto nivel; y por otro, la contribución de sus actividades de investigación al desarrollo científico, tecnológico, social y cultural de México según los objetivos y lineamientos del Plan Nacional de Desarrollo.

Según lo establecido en el Acuerdo de creación del SNI, podían postular todos los investigadores que realicen investigación de tiempo completo y que desempeñan sus actividades en las instituciones de educación superior y de investigación del sector público. Luego de la evaluación de

---

<sup>33</sup> El primer Reglamento del SNI disponible en el Diario Oficial de la Federación (DOF) es del año 2006. En la página web del CONACYT sólo están disponibles los Reglamentos de 2008, 2012, 2016 (disponibles en <http://www.conacyt.gob.mx/index.php/el-conacyt/sistema-nacional-de-investigadores/otros/reglamentos-antteriores>) y el vigente desde 2017 (disponible en <http://www.conacyt.gob.mx/index.php/el-conacyt/sistema-nacional-de-investigadores/marco-legal/reglamento-sni>). Fecha de consulta 12/05/2017.



los méritos académicos de los investigadores -realizada por pares-, ellos podían quedar categorizados en dos grandes categorías: i) *Investigador Nacional*, orientada al estímulo de los investigadores activos, que a su vez está dividida en tres sub-categorías que van desde el *Nivel I*, pasando por el *Nivel II*, hasta *Nivel III*, que es la categoría de mayor reconocimiento dentro del sistema; y, ii) *Candidato a Investigador Nacional*, orientada a estimular a los que inician su carrera de investigación.

Desde el punto de vista de la asignación de recursos económicos, los investigadores que hayan sido nombrados ya sea como *investigador nacional* –en cualquiera de sus niveles- o como *candidato*, recibirán un estímulo económico –una beca- cuyo monto era de un salario mínimo para los *candidatos* y de dos, tres y cinco salarios mínimos para los *investigadores nacionales* categorizados en los niveles I, II y III respectivamente (De Ibarrola, 2005). Además, las becas otorgadas, de carácter temporal, según la permanencia en el sistema -de tres a cuatro años- y sujeta a evaluación, no afectarán en ningún caso la relación contractual del investigador con la institución en la que desempeña sus actividades y tampoco será considerada como salario ni como otra remuneración por servicios prestados. Finalmente, las convocatorias a formar parte del SNI serán anuales, y los resultados de cada promoción serán inapelables. Los resultados de evaluación serán públicos, en cada promoción se publicará la lista de los investigadores categorizados para el conocimiento de los interesados.

Desde su creación hasta la actualidad, el SNI ha sido modificado en diversas ocasiones. De Ibarrola (2005) señala que su marco legal -el Acuerdo de creación y el Reglamento-, ha estado en revisión permanente y cambio, aunque sus características centrales se mantienen. El Acuerdo de creación del SNI ha sido modificado cinco veces. En 1986 se registra la primera modificación, que sub-divide la Comisión Dictaminadora de Ciencias, Física, Matemáticas e Ingenierías en dos, creando una cuarta Comisión Dictaminadora, denominada Ingeniería y Tecnología, para una evaluación más adecuada y específica de los méritos académicos de los investigadores asociados a ese campo científico. (DOF, 6/02/1986). Canales<sup>34</sup> (2013) atribuye –convincientemente- este cambio al malestar que se generó entre los ingenieros el hecho de ser evaluados junto con los matemáticos y los físicos. Es decir, con criterios de evaluación y ponderación de la producción de conocimiento, propias de otras tradiciones disciplinarias y áreas del conocimiento.

En 1988 tiene lugar una segunda modificación al Acuerdo de creación del SNI. Esta modificación estuvo orientada en dos direcciones. Por un lado amplía la población objetivo del SNI, permitiendo la postulación de los investigadores que se desempeñan en el ámbito privado; y por otro, habilita a actuar como evaluadores en la Comisión Dictaminadora de Ingeniería y Tecnología a expertos de reconocida trayectoria en esos campos, aunque no hayan sido reconocidos como *investigadores nacionales nivel III*. (DOF, 24/03/1988).

En 1993 y con los fundamentos de: “Que es necesario abrir la participación de la comunidad científica en la integración de las Comisiones Dictaminadoras así como renovar periódicamente al cuerpo de investigadores miembros de dichas Comisiones, para fortalecer la imparcialidad y transparencia e imprimir mayor dinamismo en sus decisiones; [y] Que es conveniente estimular a aquellos investigadores que han realizado una labor extensa, relevante y permanente en favor de

---

<sup>34</sup> Ver “Las comisiones en el SNI”, publicado el 21/2/2013 en Milenio.

la investigación (...)” (DOF, 4/06/1993, énfasis añadido); se realizan tres cambios al Acuerdo de creación, a saber: i) pasa de cuatro a tres los miembros propuestos por la AMC para integrar las Comisiones Dictaminadoras, y los restantes seis -desde este momento- serán sugeridos por el CONACYT, a propuesta de otras instituciones y asociaciones científicas con alto prestigio académico; ii) reduce el período de actuación de los evaluadores en las Comisiones Dictaminadoras, pasando de cuatro a tres años; y, iii) crea la categoría de *Investigador Nacional Emérito*<sup>35</sup> (DOF, 4/06/1993).

Con un fundamento similar al del año 1993, en 1995 tiene lugar la cuarta modificación al Acuerdo de creación del SNI. Esta modificación supuso una ampliación de las Comisiones Dictaminadoras, aumentando de nueve a doce la cantidad de investigadores nivel III que actúan como evaluadores (DOF, 14/09/1995). En ese momento la fundamentación para tal modificación sostenía: “Que en este contexto, se considera necesario actualizar al Sistema Nacional de Investigadores para que continúe apoyando a la comunidad científica que lo amerite mediante procedimientos que aseguren un acceso equitativo al Sistema, y, [que para lograrlo] resulta necesario ampliar aún más la participación de la comunidad científica en la integración de las Comisiones Dictaminadoras, para fortalecer su imparcialidad y transparencia e imprimir mayor dinamismo en sus decisiones (...)” (DOF, 14/09/1995, énfasis añadido). Los tres miembros adicionales también serán sugeridos por el CONACYT, a propuesta de otras instituciones y asociaciones científicas con alto prestigio académico.

Finalmente, en el año 1999 se registra la última modificación al Acuerdo de creación del SNI de 1984, siendo así el vigente hasta la actualidad. Entre los fundamentos para realizar las modificaciones que se detallan a continuación está “Que con el propósito de garantizar la transparencia del proceso de evaluación de los aspirantes a ingresar o reingresar en el Sistema resulta conveniente establecer una instancia a través de la cual éstos puedan plantear sus inconformidades” (DOF, 13/04/1999, énfasis añadido).

En ese sentido y a partir de este momento, han habido dos cambios significativos. Por un lado, los resultados de evaluación no solamente serán publicados en diversos medios de comunicación, sino que además a los interesados se les hará llegar una notificación con la resolución y el dictamen de evaluación “debidamente razonado”. Por otro lado, se crea un mecanismo por el cual los investigadores insatisfechos con los resultados de evaluación pueden apelar -de forma escrita y con la documentación que sustenta su argumento- la resolución, dirigiéndose a la Comisión Dictaminadora Revisora -del área correspondiente-, creada en este momento para tales efectos.

Otro cambio significativo, que también se da en este momento, se vincula nuevamente con la ampliación de cuatro a siete la cantidad de Comisiones Dictaminadoras asociadas a las diversas áreas del conocimiento y sus especificidades, que a partir de este año se establecen en el Reglamento del SNI.

Por lo dicho hasta aquí, ya se distinguen algunas características que permiten asociar al SNI, con lo que en la literatura se conoce como un *sistema de evaluación fuerte*. A saber, la evaluación está institucionalizada y es sistemática, se realiza en períodos regulares de acuerdo a ciertas reglas y procedimientos altamente formalizados, y bastante transparentes –por lo menos en lo referido a

---

<sup>35</sup> Los requisitos para recibir la distinción de *investigador nacional emérito* implican ser investigador nacional nivel III, haber sido renovado en esa categoría por dos períodos consecutivos y tener 60 años o más. (DOF, 4/06/1993).

los resultados de la evaluación-; además está organizada en torno a las disciplinas existentes (las que son agrupadas en comisiones dictaminadoras según áreas de conocimiento).

Otros elementos importantes para identificar al SNI como a un sistema de evaluación fuerte -tales como: los criterios generales y específicos de evaluación, la vigencia de las distinciones que otorga el sistema y los estímulos económicos previstos como premios al desempeño de investigación-, se describen a continuación.

Para lo que sigue, importa señalar que si bien se mencionan algunos cambios ocurridos antes del 2006 y después del año 2012, el énfasis se pone en el período comprendido entre esos dos años, debido por un lado, al acceso a la información y por otro, a que el Reglamento de 2012 es el último vigente a la fecha en que se recoge la información contenida en la encuesta que se analiza en este trabajo.

### VI.2.2 Reglamentos de operación y funcionamiento del SNI

Según De Ibarrola (2005) desde la primera convocatoria del SNI hasta el año 2005 se emitieron trece Reglamentos -1984, 1988, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1999, 2000, 2003 y 2004-, y desde ese momento hasta la actualidad, se emitieron cuatro Reglamentos más, a saber:

- i) en 2006 (DOF, 26/09/2006), que fue modificado con cambios menores en dos oportunidades durante el año 2007 (DOF, 07/02/2007; 10/08/2007);
- ii) en 2008 (DOF, 21/03/2008), que también fue modificado con cambios menores tres veces (DOF, 21/10/2009; 25/07/2011; 02/08/2011);
- iii) en 2012 (DOF, 26/12/2012), que fue modificado en cuatro ocasiones (DOF, 11/02/2014; 14/11/2014; 21/05/2015; 26/07/2016), las dos primeras orientadas a reglamentar la postulación de los investigadores asociados a las “Cátedras CONACYT”, la tercera es un cambio menor, y en la cuarta modificación aparecen por primera vez el Comité de investigadores eméritos y la figura de las Comisiones Transversales; y,
- iv) en 2017 (DOF, 27/01/2017) que es el vigente.

Reyes y Suriñach (2012) señalan que en el año 1996 aumentaron las exigencias para ingresar al SNI, debido a que a partir de ese momento se comenzó a exigir -como requisito- el grado académico de doctor. Otro cambio en ese año fue que los aspirantes que ingresaran al sistema como *candidatos*, podían permanecer en esa categoría solo una vez por un período de vigencia de tres años.

Como se adelantó, de acuerdo con De Ibarrola (2005) y Canales (2013), en el Reglamento de 1999 se establecen las siete Comisiones Dictaminadoras asociadas a cada una de las áreas del conocimiento, que estarán vigentes hasta el año 2016, a saber: i) Físico-Matemáticas y Ciencias de la Tierra; ii) Biología y Química; iii) Medicina y Ciencias de la Salud; iv) Humanidades y Ciencias de la Conducta; v) Ciencias Sociales; vi) Biotecnología y Ciencias Agropecuarias; e vii) Ingenierías. A partir del año 2017 además de las siete Comisiones Dictaminadoras señaladas, aparece la Comisión Transversal de Tecnología (CTT) (Reglamento 2017), cuyas funciones centrales son evaluar – transversalmente- la producción tecnológica de los postulantes del SNI y en función de ello, asesorar

e informar a cada una las siete Comisiones Dictaminadoras sobre la pertinencia de dicha producción (Criterios internos y específicos CTT, 2017<sup>36</sup>).

Además, de esa ampliación en el número de Comisiones Dictaminadoras, en el Reglamento del año 2006 se vuelve a incrementar el número de investigadores -de doce a catorce miembros- que actúan como evaluadores en dichas comisiones (Reglamento SNI, 2006 en DOF, 26/09/2006).

De acuerdo con lo planteado en el Reglamento de 2006, las Comisiones Dictaminadoras se integrarán procurando “el mayor equilibrio y paridad posible entre disciplinas, instituciones y regiones” (DOF, 26/09/2006 énfasis añadido). Al igual que en años anteriores, los evaluadores se desempeñarán en sus actividades durante tres años, pero cada año se renovarán cuatro de ellos y su designación tomará en cuenta la propuesta realizada por el Foro Consultivo Científico y Tecnológico (FCCyT). Esta propuesta del FCCyT se basa en una consulta abierta, en la que la comunidad académica nucleada en el SNI, elige -mediante votación- a los investigadores que serán propuestos por el FCCyT a formar parte de las Comisiones Dictaminadoras.

En marzo de 2013, en la prensa mexicana<sup>37</sup> se informaba que el FCCyT registraba la mayor concurrencia de los miembros del SNI en el proceso de renovación de las Comisiones Dictaminadoras, alcanzando una participación del 64,5% de ellos. En dicha publicación, la Dra. Gabriela Dutrénit –investigadora y coordinadora del FCCyT- destacaba y reafirmaba que: *“... lo que se busca desde el FCCyT es contribuir con la mayor representatividad de acuerdo con las diversas disciplinas, instituciones, regiones y género de los investigadores categorizados en el sistema”*.

Los integrantes de las Comisiones Dictaminadoras son de conocimiento público y sus nombres son publicados en la página web del CONACYT. También, y a propuesta de las Comisiones Dictaminadoras, está previsto que se integren sub-comisiones para la evaluación de solicitudes en disciplinas específicas (Reglamento 2006).

Desde el punto de vista de los criterios generales de evaluación, en el Reglamento del año 2006 se establece que las Comisiones Dictaminadoras reconocen distintas tradiciones disciplinares respecto a la difusión de los resultados de investigación científica y tecnológica; y que además, la difusión también varía según la institución –políticas editoriales, normas y exigencias-, su ubicación geográfica, la capacidad financiera de los ámbitos laborales de los investigadores y de sus redes de investigación; en ese sentido se indica que la evaluación de los trabajos se basa en su calidad, trascendencia e impacto.

Las Comisiones Dictaminadoras, integradas por pares y en base a la revisión del *curriculum vitae* y la documentación probatoria presentados por el postulante, analizarán su calidad y la cantidad de productos presentados (productividad), y en función de ello emitirán su dictamen, proponiendo al Consejo de Aprobación –a través del Secretario Ejecutivo- el ingreso o no al SNI y su nivel de categorización en el Sistema. Las solicitudes serán evaluadas por al menos dos integrantes de la Comisión Dictaminadora que corresponda y la resolución final es decidida por toda la comisión. Los

---

<sup>36</sup> Criterios internos de la Comisión Transversal de Tecnología, disponibles en <http://conacyt.gob.mx/index.php/sni/otros/marco-legal-sni/criterios-sni/13721-criterios-especificos-ctt/file>. Fecha de consulta, 12/05/2017

<sup>37</sup> Ver “Buscan comisiones más equilibradas en el SNI”, publicado el 7/3/2013 en El Economista.

resultados de la evaluación serán publicados en la página web del CONACYT y los dictámenes de evaluación serán enviados a cada uno de los postulantes.

Los criterios generales de evaluación del 2006 sostienen que se evaluará prioritariamente la cantidad y la calidad de la producción del último período de evaluación -cuando se trata de reingreso vigente-, o de los últimos tres años -si se trata de nuevo ingreso o reingreso no vigente-. Además, se tendrá en cuenta la producción global del postulante.

Según el Reglamento de 2006, la calidad de la producción se valora en términos de su originalidad, impacto, trascendencia y repercusión en la solución de los problemas prioritarios del país; además se examina la consolidación de la(s) línea(s) de investigación, la creación de empresas de alto valor agregado y el liderazgo y reconocimiento nacional e internacional del aspirante a formar parte del SNI. Estos criterios generales de evaluación y el procedimiento en su conjunto se mantienen en los Reglamentos del SNI de los años 2008, 2012 y 2017.

Sin embargo, a partir del Reglamento del año 2008 y en los siguientes Reglamentos, los insumos centrales de evaluación se reducen a dos tipos de actividades académicas, las asociadas a la investigación científica y/o tecnológica y a la de formación de científicos y tecnólogos; también disminuye el tipo de productos que son considerados en la evaluación vinculados a cada una de esas actividades académicas, y a partir del año 2017 se hace explícito -en el Reglamento de ese año- la importancia del prestigio del órgano de publicación, que hasta entonces se señalaba únicamente en los *criterios específicos e internos* de evaluación de las distintas Comisiones Dictaminadoras (en la siguiente sub-sección se profundiza sobre este y otros criterios específicos de evaluación). En los diversos Reglamentos está previsto que cada una de las Comisiones Dictaminadoras defina -en base a los criterios generales de evaluación-, criterios de evaluación específicos.

Las otras actividades académicas realizadas por los postulantes, tales como: el trabajo de divulgación, la vinculación de la investigación con los sectores público, social y privado, la participación en el desarrollo institucional y en cuerpos editoriales o colegiados de evaluación científica y tecnológica serán consideradas de manera complementaria para la permanencia o ascenso en el SNI, pero en ningún caso sustituyen o compensan a los productos centrales de evaluación asociados a las dos actividades antes mencionadas.

En la tabla 4 se listan los criterios generales de evaluación, las actividades y productos centrales de evaluación considerados en los distintos Reglamentos del SNI en el período comprendido entre los años 2006 y 2017.

Tabla 4: Criterios Generales: actividades y productos centrales considerados en la evaluación del SNI

Actividades y productos de Evaluación 2006	Actividades y productos de Evaluación 2008 y 2012	Actividades y productos
<p><b>Investigación científica y/o tecnológica:</b> <u>artículos</u>; memorias in extenso; <u>libros</u>; <u>capítulos de libros</u>; reseñas; opúsculos; <u>patentes</u>; <u>desarrollos tecnológicos</u>; <u>innovaciones</u>; <u>transferencias tecnológicas</u>; participación en proyectos de largo aliento.</p>	<p><b>Investigación científica y/o tecnológica:</b> artículos; libros; capítulos de libros; patentes; desarrollos tecnológicos; innovaciones; transferencias tecnológicas.</p>	<p><b>Investigación científica y/o tecnológica:</b> que hayan sido sujetos de estudio de comités editoriales de revistas de reconocido prestigio; artículos de investigación; capítulos de libros publicados por editoriales de reconocido prestigio en el ámbito de la investigación; becas concedidas en México y en el extranjero; desarrollos tecnológicos; innovaciones científicas; transferencias tecnológicas de quien la llevó a cabo.</p>
<p><b>Formación de científicos y tecnólogos:</b> <u>dirección de tesis profesionales y de posgrado terminadas</u>; <u>impartición de cursos en licenciatura y posgrado</u>; libros de texto (educación superior); artículos en revistas de docencia; diaporamas o programas de cómputo de carácter educativo; <u>formación de investigadores</u>; formación de especialistas; formación de profesionales; formación de técnicos especializados; tutoría de estudiantes.</p>	<p><b>Formación de científicos y tecnólogos:</b> dirección de tesis profesionales y de posgrado terminadas; impartición de cursos en licenciatura y posgrado; formación de investigadores y grupos de investigación.</p>	<p><b>Formación de científicos y tecnólogos:</b> de tesis profesionales y de posgrado; codirecciones de tesis; participación en comités de tesis; cursos en licenciatura y posgrado; formación de investigadores y grupos de investigación.</p>
<p><b>Divulgación:</b> conferencias y seminarios de divulgación; participación en congresos de divulgación; libros de divulgación; capítulos en libros de divulgación; artículos (ediciones formales) de divulgación; ensayos (ediciones formales) de divulgación; diseño de equipamientos y exposiciones museográficas.</p>		

<p><b>Vinculación de la investigación con los sectores público, social y privado:</b> reportes de proyectos específicos bajo contrato; desarrollo de posgrados orientados; reportes de trabajos solicitados por terceros.</p>		
<p><b>Desarrollo institucional:</b> generación, consolidación o fortalecimiento de centros, unidades, talleres, bancos de información, laboratorios de investigación o acervos documentales, bibliográficos, de especies biológicas o colecciones científicas; creación y montaje de equipo para las actividades de investigación; creación de grupos de investigación; desarrollo y consolidación de posgrados; informes de resultados de promoción, formación y coordinación en redes nacionales de investigación; tutoría de grupos nacionales emergentes; participación en consorcios de innovación; obtención de financiamiento nacional e internacional para proyectos de investigación; desarrollo de programas de investigación y docencia interinstitucionales en el sector educativo y científico; participación en mesas directivas de asociaciones científicas.</p>		
<p><b>Participación en cuerpos editoriales o colegiados de evaluación científica y tecnológica:</b> participación en Comisiones Dictaminadoras; arbitraje de artículos en revistas especializadas de circulación internacional con alto impacto, solicitados por el comité editorial correspondiente; evaluación de propuestas de investigación; evaluación de programas de posgrado; participación en órganos de evaluación del trabajo científico y tecnológico; promoción editorial y participación en consejos o comités editoriales</p>		

Fuente: Elaboración propia en base a los Reglamentos del SNI de 2006, 2008, 2012 y 2017. Disponibles en DOF 26/09/2006; 21/03/2008, 26/12/20

Hasta este punto se han descrito diversos elementos del sistema de evaluación del SNI, que en el marco de este trabajo han sido aproximados empíricamente a través de la dimensión conceptual definida como *satisfacción con el proceso de evaluación*. Esta dimensión busca aproximar una medida sobre la satisfacción que tienen los investigadores categorizados en el SNI, en relación a múltiples aspectos vinculados a las Comisiones Dictaminadoras -tales como: la capacidad de los mecanismos de designación de los evaluadores para seleccionar perfiles que faciliten evaluaciones apropiadas en las diversas áreas de conocimiento, la adecuación y equilibrios en la integración de dichas comisiones, y la pertinencia de los requisitos para ser evaluador-; y además, sobre la relevancia de la información solicitada, tanto para el desarrollo de una evaluación que logre juicios académicos apropiados, como en la inclusión de todas las contribuciones relevantes realizadas por los investigadores.

En este sentido, dar cuenta de las características específicas del sistema de evaluación del SNI, permite una mejor comprensión de los resultados empíricos obtenidos, tanto en relación a la satisfacción con el proceso de evaluación, como también respecto a los observados en los ítems específicos explicados por dicha dimensión conceptual. Estos resultados se presentan en el capítulo IX.

Desde el punto de vista de la frecuencia de la evaluación, de acuerdo con lo establecido en los Reglamentos del SNI (2006, 2008, 2012 y 2017) los períodos de tiempo entre evaluaciones han permanecido constantes entre 2006 y 2017 para el caso de los *investigadores nacionales* (niveles I, II y III), mientras que para los *candidatos*, si bien la vigencia del nombramiento al ingreso se ha mantenido constante, su prórroga ha variado. (Ver tabla 5)

En particular y dependiendo de los criterios internos de evaluación de cada Comisión Dictaminadora, para los candidatos, en el año 2006, la prórroga se daba una sola vez, por un año consecutivo; en 2008 pasaron a otorgarse dos prórrogas únicas -de un año cada una y consecutivas-; en 2012 se regresó a una única prórroga pero de hasta dos años; y a partir de 2017, dicha prórroga se torna “excepcional”, única y sólo por un año consecutivo para aquellos casos cuya productividad permite predecir que al cabo de ese año el *candidato* ascenderá a la categoría de *investigador nacional*.

Este cambio de 2017 es un retroceso significativo del sistema de evaluación, en la medida que implica una severa restricción al ingreso para los jóvenes investigadores que intentan construir su trayectoria académica.

Adicionalmente, la amplia mayoría de los miembros del SNI (*candidatos y nivel I*) disponen de un período entre evaluaciones muy restringido, de apenas tres años hasta la segunda evaluación. La importancia que tienen los períodos de tiempo disponibles entre cada evaluación -de los productos de la investigación, radica en que de ellos depende -aunque no de manera exclusiva- la calidad, originalidad, relevancia conceptual y pertinencia de las contribuciones derivadas de la producción de conocimiento. El mecanismo es relativamente evidente, si los períodos entre evaluaciones son demasiado acotados, el horizonte temporal de la investigación también debe ser menor, y por lo tanto, las actividades de investigación deben acotarse en términos de su alcance, ambición de sus objetivos, dificultad, riesgo y tiempo disponible para lograr no solamente resultados, sino productos concretos.



Tabla 5: Vigencia y Estímulos: tiempo entre evaluaciones y el incentivo económico

	2006		2008		2012		2017	
	Vigencia <sup>1</sup>	Estímulo <sup>4</sup>	Vigencia <sup>1</sup>	Estímulo <sup>4</sup>	Vigencia <sup>1</sup>	Estímulo <sup>4</sup>	Vigencia <sup>1</sup>	Estímulo <sup>5</sup>
Candidato	3 (1)	3	3 (1) (1)	3	3 (2) <sup>2</sup>	3	3 (1) <sup>3</sup>	6800
Nivel I	3 (4)	6	3 (4)	6	3 (4)	6	3 (4)	13600
Nivel II	4 (5)	8	4 (5)	8	4 (5)	8	4 (5)	18200
Nivel III	5 (5) (10)	14	5 (5) (10)	14	5 (5) (10)	14	5 (5) (10)	31900

Fuente: Elaboración propia en base a los Reglamentos del SNI de 2006, 2008, 2012 y 2017. Disponibles en DOF 26/09/2006; 21/03/2008, 26/12/2012 y 27/01/2017 respectivamente. Notas: 1. Se trata de la vigencia de la distinción medida en años, y entre paréntesis, los años de prórrogas consecutivas en el mismo nivel. Los Candidatos sólo pueden tener una o dos prórrogas en el mismo nivel; en las restantes categorías no hay restricción en la cantidad de prórrogas en el mismo nivel. 2. Queda a juicio de la Comisión Dictaminadora si la prórroga la otorga por dos años o uno. 3. La prórroga es excepcional y sólo por un año. Será otorgada a juicio de la Comisión Dictaminadora y en los casos en que sea altamente probable que, al finalizar el período, el Candidato haya alcanzado los requisitos suficientes para ascender o ingresar a la categoría investigador nacional. 4. Los estímulos económicos son mensuales y están medidos en unidades de Salario Mínimo mensual, es decir, se ajustan según la variación del Índice Nacional de Precios al Consumidor y no son pasibles de ser gravados impositivamente. 5. A partir del 2017 los estímulos económicos mensuales se miden en pesos mexicanos corrientes y se ajustan anualmente según la variación del Índice Nacional de Precios al Consumidor, pero nunca por encima del 3% y tampoco son pasibles de ser gravados impositivamente.

En el momento de la segunda evaluación y a diferencia de lo dicho para los candidatos, los investigadores nivel I que permanecen en el sistema y en ese nivel, tendrán evaluaciones sistemáticas cada cuatro años. Esa ampliación en el período entre evaluaciones representa una cierta mejora para el desarrollo de investigaciones más ambiciosas, sin embargo para afirmar que ese tiempo es suficiente en el contexto de los académicos en México, es preciso realizar estudios específicos en esa dirección.

Los períodos de tiempo al ingreso y entre evaluaciones aumentan para los niveles II y III. Mientras que para los primeros su primer nombramiento es por cuatro años, y las siguientes evaluaciones cada cinco años; para los segundos, el primer nombramiento es por cinco años, la segunda evaluación también tendrá una vigencia de cinco años, pero las siguientes evaluaciones sistemáticas serán cada diez años. (Ver tabla 5).

Sin embargo, y como se verá en la sub-sección siguiente, estos períodos de tiempo son relativos en la medida que de acuerdo con los criterios específicos de evaluación de las Comisiones Dictaminadoras, algunas de ellas exigen -para el ascenso o permanencia en el sistema- demostrar *continuidad* en el período entre evaluaciones a través de la presentación de productos anuales de investigación y formación de recursos humanos.

Desde el punto de vista de la asignación de recursos económicos, además de las distinciones y el prestigio que representan, el SNI también otorga estímulos económicos –provenientes de fondos públicos- mensuales en cada una de las categorías y niveles durante el período de vigencia de los nombramientos. En la tabla 5 se pueden observar los tamaños absolutos de dichos estímulos. Los estímulos se han mantenido constantes hasta el año 2017, cuando se aprueba el último Reglamento vigente. En ese momento cambia la unidad de medida del estímulo, pasando de unidades de salario mínimo a pesos mexicanos corrientes y su ajuste, que si bien sigue siendo según la variación de la inflación, el ajuste máximo posible será de 3%.

Los estímulos económicos otorgados son considerados como una *beca* para premiar el desempeño de los investigadores distinguidos por el sistema. Como tal, la *beca* es temporal y no está sujeta a gravámenes impositivos ni realiza aportes jubilatorios.

De acuerdo con un estudio realizado en 2013, y considerando el peso relativo de todos los estímulos económicos<sup>38</sup> en el total de ingresos anuales de los investigadores categorizados en el SNI, Bensusán, Ahumada e Inclán (2013) muestran que para el 75% de dichos investigadores, los estímulos económicos que perciben multiplican por dos o más sus ingresos salariales anuales.

Adicionalmente, pertenecer al SNI facilita el acceso a fuentes externas de financiamiento público para la investigación y/o para el desarrollo de programas de posgrado. Algunos ejemplos –aunque no únicos- son a través de los Proyectos de Desarrollo Científico para atender Problemas Nacionales<sup>39</sup> y los Proyectos de Investigación Científica Básica<sup>40</sup>, ambos son fondos competitivos del CONACYT y entre los requisitos para postular se exige ser doctor o miembro del SNI. Otro ejemplo, pero vinculado a la formación de posgrado, se puede encontrar en el Programa Nacional de Posgrados de Calidad del CONACYT<sup>41</sup>, ya que entre los criterios utilizados para evaluar la calidad de los diversos posgrados, y la consecuente asignación de recursos -becas para sus estudiantes-, se considera muy positivamente que los profesores de los cuerpos académicos del posgrado hayan sido nombrados miembros del SNI; a la vez, otra medida para ponderar la calidad y el impacto del posgrado, es que sus egresados también ingresen al SNI.

De esta forma, se puede afirmar la gran importancia que tiene mantener la pertenencia al SNI, no solamente por el prestigio y el reconocimiento social en el ámbito académico que representa, sino también por la relevancia, tanto de su peso económico en los ingresos de los investigadores, como en el acceso a otras fuentes de financiamiento externo para la investigación y la formación de recursos humanos del más alto nivel. En ese sentido, resulta razonable esperar una gran y fuerte adaptación de los investigadores respecto a sus actividades y productos de investigación para ajustarse a lo que dicho sistema premia y valora, reforzando aún más algunos de sus efectos.

---

<sup>38</sup> Incluye otros estímulos económicos extra-salariales además del otorgado por el SNI, sin embargo este último es el más importante.

<sup>39</sup> Ver convocatoria 2016. Disponible en <http://www.conacyt.gob.mx/index.php/el-conacyt/convocatorias-y-resultados-conacyt/convocatorias-desarrollo-cientifico-1/convocatoria-de-proyectos-de-desarrollo-cientifico-para-atender-problemas-nacionales-2016/13264-convocatoria-problemas-nacionales-2016/file>. Fecha de consulta 20/05/2017

<sup>40</sup> Ver convocatoria 2013-2014. Disponible en <http://www.conacyt.mx/index.php/el-conacyt/convocatorias-y-resultados-conacyt/convocatorias-fondos-sectoriales-constituidos/convocatoria-sep-conacyt/investigacion-basica-sep/cerrada-investigacion-basica/convocatoria-2013-2014-investigacion-basica/8964-bases-de-la-convocatoria-12/file>. Fecha de consulta 20/05/2017

<sup>41</sup> Ver Marco de Referencia para la evaluación de los Programas de Posgrado de Calidad. Disponible en [www.conacyt.mx/index.php/becas-y-posgrados/programa-nacional-de-posgrados-de-calidad/convocatorias-avisos-y-resultados/convocatorias-cerradas-pnpc/9005-marco-de-referencia-modalidad-escolarizada/file](http://www.conacyt.mx/index.php/becas-y-posgrados/programa-nacional-de-posgrados-de-calidad/convocatorias-avisos-y-resultados/convocatorias-cerradas-pnpc/9005-marco-de-referencia-modalidad-escolarizada/file). Fecha de consulta 20/05/2017

### VI.2.3 Criterios de evaluación específicos e internos de las Comisiones Dictaminadoras por área de conocimiento

Los criterios de evaluación específicos de cada una de las siete Comisiones Dictaminadoras, que se presentan en esta sub-sección, refieren a los criterios establecidos –como mínimo- a partir de la publicación de los Reglamentos del SNI de los años 2006 y 2008<sup>42</sup>.

En 2017, y junto con el Reglamento de ese año, se revisaron y actualizaron los criterios específicos por área de conocimiento<sup>43</sup>. Dicha revisión y actualización, se centra en la articulación de los criterios específicos de cada Comisión Dictaminadora con los de la recientemente creada Comisión Transversal de Tecnología. El resto de los elementos de la evaluación, tales como: los productos, las actividades de enseñanza y de divulgación de la ciencia que se consideran, los criterios utilizados para aproximar la calidad e impacto de las contribuciones realizadas por los investigadores, así como su liderazgo y su aporte en la creación de infraestructura permanecen constantes. También permanecen inalterados los criterios de evaluación para determinar la permanencia y/o ascenso en el sistema.

En ese sentido se puede sostener que los criterios específicos de evaluación y referencia para cada una de las siete Comisiones Dictaminadoras por área son altamente estables en el tiempo, cuya vigencia data desde al menos una década. Esto representa un reforzamiento importante de la institucionalización, reglas y normas de la evaluación, porque los criterios que la rigen son sistemáticos y sostenidos a lo largo del tiempo.

De acuerdo con los criterios internos de las Comisiones Dictaminadoras, los productos que se consideran en la evaluación son relativamente homogéneos para casi todas las áreas, la única excepción es el área de Humanidades y Ciencias de la Conducta. En esta área los libros, capítulos de libros, traducciones y ediciones críticas y anotadas de autores y de documentos de textos en lenguas clásicas, indígenas o muertas publicados en editoriales de reconocido prestigio son centrales en la evaluación; le siguen los artículos con las especificidades que se detallan a continuación; complementariamente consideran los desarrollos tecnológicos, la coordinación y edición de libros publicados en editoriales con arbitraje, y obras de divulgación y guiones museográficos.

Para las restantes áreas, el principal producto es el artículo publicado en revistas caracterizadas de la siguiente forma:

---

<sup>42</sup> Si bien los documentos que contienen los criterios específicos e internos de las Comisiones Dictaminadoras por área no tienen fecha –de publicación ni de la convocatoria a la que corresponden-, por las referencias que hacen a los diversos Reglamentos del SNI, se asume que como mínimo se aplican desde los años 2006 y 2008.

<sup>43</sup> Los criterios específicos de evaluación internos de las Comisiones Dictaminadoras por área del año 2017 están disponibles en <http://conacyt.gob.mx/index.php/sni/convocatorias-conacyt/convocatorias-sistema-nacional-de-investigadores-sni/marco-legal-sni/criterios-sni>. Fecha de consulta 20/05/2017. Los criterios específicos de evaluación anteriores y aplicados desde el 2006-2008, fueron proporcionados por la Dra. Graciela Bensusán para este trabajo, quien los obtuvo del CONACYT en el marco del proyecto “*Reflexiones sobre la evaluación a los investigadores: una mirada desde diferentes perspectivas*” (Bensusán et al., 2014).

- i) indexadas en el JCR<sup>44</sup> en el caso de las áreas de Físico, Matemáticas y Ciencias de la Tierra, Biología y Química, Medicina y Ciencias de la Salud, y Biotecnologías y Ciencias Agropecuarias;
- ii) especializadas de alta calidad, con arbitraje riguroso y amplia circulación; los evaluadores de las áreas de Humanidades y Ciencias de la Conducta, y los de Ciencias Sociales podrán privilegiar revistas indexadas nacionales o internacionales; y,
- iii) de prestigio, se indica en el área de las Ingenierías.

El segundo producto considerado por la mayoría de las áreas son los libros y capítulos de libros publicados por editoriales de prestigio, la única excepción es Medicina y Ciencias de la Salud que en sus criterios internos los señala como productos secundarios y complementarios. Finalmente, el tercer tipo de producto considerado por todas las áreas son los desarrollos tecnológicos que incluyen: las patentes concedidas, certificados de invención, de registro de nuevos productos, transferencias tecnológicas, programas de cómputo con derechos de autor, reportes técnicos, consultorías especializadas, innovaciones y otras aplicaciones relevantes en el campo de estudio. A partir de 2017, los desarrollos tecnológicos de investigadores en cualquier área son valorados por la Comisión Transversal de Tecnología, y a partir de ello, esta Comisión Transversal asesora a cada una de las Comisiones Dictaminadoras sobre la pertinencia e impacto de cada uno de ellos.

Las actividades de enseñanza que se consideran en todas las áreas se vinculan fundamentalmente con la dirección y co-dirección de tesis de licenciatura, maestría y doctorado, y con el dictado de cursos en todos esos niveles. Sin embargo, la norma es que conforme avanza el nivel, la exigencia es la dirección de tesis de doctorado; mientras que para los niveles de menor rango (candidato y nivel I) se les valora dirección y co-dirección de tesis y el dictado de cursos en la licenciatura, a los niveles II y III se les exige que sus actividades de enseñanza incluyan el posgrado y, fundamentalmente el doctorado para el caso de los investigadores nivel III. Otras actividades de enseñanza, como la elaboración de libros de texto u otros materiales didácticos, la participación en tribunales de tesis –lectores, sinodales-, etc. no son consideradas en la evaluación. Asimismo, para todas las áreas, las actividades de divulgación de la ciencia y la tecnología serán consideradas como actividades complementarias y valoradas en el caso de los niveles II y III, siempre que se compruebe su calidad, impacto e eficacia.

Respecto a los criterios utilizados para aproximar la calidad e impacto de las contribuciones realizadas por los investigadores, existe una alta estandarización, ya que todas las Comisiones Dictaminadoras asociadas a las distintas áreas utilizan los mismos indicadores, a saber:

- i) el factor de impacto de la revista y el número de citas recibidas, excluyendo las auto-citas y las citas de tesis, para el caso de los artículos publicados. Las Comisiones Dictaminadoras de las áreas de Ciencias Sociales y de Humanidades y Ciencias de la

---

<sup>44</sup> Las áreas de Física, Matemáticas y Ciencias de la Tierra, y de Biotecnologías y Ciencias Agropecuarias también consideran los artículos publicados en las revistas incluidas en el Índice de Revistas Mexicanas de Investigación Científica y Tecnológica del CONACYT. Por su parte, el área de Medicina y Ciencias de la Salud sólo los considera en segundo lugar de importancia. Para el caso de las áreas de Humanidades y Ciencias de la Conducta y de Ciencias Sociales, la consideración –o no- de esos artículos queda a juicio de los evaluadores en sus respectivas áreas de especialización. Las áreas de Biología y Química, y de Ingeniería, no mencionan los artículos publicados en las revistas incluidas en el Índice de Revistas Mexicanas de Investigación Científica y Tecnológica del CONACYT, lo que hace suponer que no son considerados.

- Conducta, relativizan estos indicadores bibliométricos, indicando que se considera el prestigio de la revista, calidad, impacto, grado de arbitraje riguroso y de circulación nacional e internacional pero no especifican cómo;
- ii) el prestigio de la editorial, el tiraje, la cantidad de re-ediciones, y su eventual traducción a otros idiomas, para el caso de los libros de investigación. Los evaluadores de la Comisión Dictaminadora de Humanidades y Ciencias de la Conducta además agregan que se considera el dictamen del libro, y los comentarios y reseñas a la obra;
  - iii) la calidad del contenido, el uso en el sector productivo y el prestigio del órgano emisor, para valorar la calidad de los desarrollos tecnológicos tales como: las patentes concedidas –las solicitadas no son consideradas –, los certificados de invención y de registro, y/o la elaboración de normas y reglamentos; y,
  - iv) el impacto de esos desarrollos tecnológicos: en el ámbito industrial, social o de la política pública, la contribución a la solución de los problemas sociales, productivos y/o de política pública del país, o la generación de valor económico, se demuestra a través de cartas de los usuarios.

La única Comisión Dictaminadora que en sus criterios internos y específicos indica que, además de lo señalado para aproximar la calidad e impacto de las contribuciones de los investigadores, tendrá en cuenta el contenido de la investigación realizada, es la asociada al área de Física, Matemáticas y Ciencias de la Tierra. Asimismo, señala que para ascender a los niveles II y III, se realiza un análisis de las citas recibidas, distribuyéndolas en la obra global del investigador. Por su parte, la Comisión Dictaminadora asociada al área de Humanidades y Ciencias de la Conducta, indica que para ascender a los niveles más altos del sistema, es central una productividad destacada de publicaciones en revistas con arbitraje, y dependiendo de la disciplina, los libros de autor son muy importantes también, además de una trayectoria que refleje el trabajo de calidad.

Además de lo dicho y por igual para todas las áreas asociadas a las distintas Comisiones Dictaminadoras, los investigadores que busquen ascender a los niveles II o III deben, complementariamente y por un lado, contribuir a la creación de infraestructura para la investigación, lo que de acuerdo con los criterios internos de evaluación significa: i) la generación, consolidación o fortalecimiento de unidades, laboratorios, grupos de investigación y/o de desarrollo tecnológico; y ii) la creación de nuevos instrumentos para la investigación y/o para el desarrollo tecnológico, mismos que serán valorados dependiendo del reconocimiento explícito en artículos científicos, patentes y/o cartas de los usuarios.

Por otro lado, y de acuerdo con los criterios internos de todas las Comisiones Dictaminadoras por área, los investigadores que aspiren ascender a los niveles II y III, también deben demostrar su liderazgo académico en la generación y consolidación de líneas de investigación, y en la formación de grupos de investigación y de investigadores independientes. De acuerdo a lo planteado en los criterios internos de evaluación de las diferentes Comisiones Dictaminadoras, dicho liderazgo se demuestra fundamentalmente a través de una cantidad significativa de productos de investigación de alta calidad y su impacto –gran cantidad de citas o interés sustantivo del desarrollo tecnológico para su explotación comercial- en una determinada disciplina o especialidad; y complementariamente, por el reconocimiento nacional e internacional que el investigador haya recibido a través de la participación en comités editoriales, comisiones evaluadoras, premios,

distinciones y/o invitaciones como ponente en conferencias magistrales y encuentros internacionales.

Además de lo dicho, los criterios internos de evaluación de las distintas Comisiones Dictaminadoras por área, también especifican exigencias diferenciales y mayores según aumentan los niveles de categorización (*candidato, nivel I, nivel II, nivel III*) y, también varían, dependiendo de si se trata de un nuevo ingreso al sistema, o de la permanencia en el mismo nivel.

Al ingreso y para quedar categorizado como *Candidato* en el sistema, las distintas Comisiones Dictaminadoras exigen por igual que el postulante demuestre tener el grado académico de doctor o su equivalente, y también que muestre productos generados en los últimos tres años. Las diferencias en los criterios de evaluación -para ingresar como *candidato*- entre las Comisiones Dictaminadoras por área, aparecen cuando se observa qué incluye o considera cada una de ellas por productos generados.

En particular, mientras que las Comisiones Dictaminadoras de las áreas de Física, Matemáticas y Ciencias de la Tierra e Ingeniería indican que se consideran cualquiera de los productos establecidos en los Criterios Generales del Reglamento; las restantes son mucho más precisas, no solamente en relación a qué productos, sino también en términos de sus cantidades mínimas requeridas. Además, aunque con alguna excepción, para las Comisiones Dictaminadoras “productos” parece ser sinónimo de “publicaciones”.

La Comisión Dictaminadora del área de Biología y Química exige tener dos trabajos publicados como primer autor, o entre dos y tres como co-autor, y si bien no realiza especificaciones adicionales sobre el órgano de publicación, señala que se tendrá en cuenta la calidad de la revista. Adicionalmente, indica que para ingresar como *candidato* no puede haber transcurrido más de quince años desde la licenciatura.

Las Comisiones Dictaminadoras asociadas a las áreas de Medicina y Ciencias de la Salud, y de Biotecnología y Ciencias Agropecuarias, son las únicas que consideran los desarrollos tecnológicos como parte de los productos y exigen para ingresar como *candidato*, tener un desarrollo tecnológico documentado; o publicaciones –no especifica cantidad, ni rol en la autoría de las mismas- en revistas indexadas en el JCR en el caso del área de Medicina y Ciencias de la Salud; o un artículo como primer autor en revistas indexadas, ya sea en el JCR o en el Índice de Revistas Mexicanas de Investigación Científica y Tecnológica del CONACYT.

Por su parte, las Comisiones Dictaminadoras de las áreas de Humanidades y Ciencias de la Conducta y de Ciencias Sociales, exigen un artículo como autor único o primer autor en revistas arbitradas, o un capítulo de libro de investigación que haya sido dictaminado y publicado por una editorial de reconocido prestigio científico y académico, o varios<sup>45</sup> artículos en la misma línea de investigación como co-autor en revistas arbitradas.

A excepción de las Comisiones Dictaminadoras de Física, Matemáticas y Ciencias de la Tierra e Ingeniería, ninguna de las restantes comisiones considera la participación del postulante en la

---

<sup>45</sup> La Comisión Dictaminadora de Ciencias Sociales indica dos artículos.

formación de recursos humanos (dictado de cursos y/o asesoría de tesis) como productos requeridos para la categoría *candidatos*.

Nuevamente aquí, al ingreso y para quedar categorizado como *Nivel I* en el sistema, las distintas Comisiones Dictaminadoras de todas las áreas exigen cumplir con los requerimientos señalados para la categoría de *candidato*, y todas ellas también consideran, la contribución de los investigadores en la formación de recursos humanos, en particular a través de la dirección de tesis o el dictado de cursos, ya sean de licenciatura o posgrado. La Comisión Dictaminadora de Humanidades y Ciencias de la Conducta es especialmente precisa en relación a la participación en actividades de formación y agrega, además de lo dicho, la participación en proyectos de servicio social, en actividades de educación continua y en la actualización profesional. Además de esos requisitos, también deben cumplirse otros vinculados a la investigación y el desarrollo tecnológico.

Las variaciones de los criterios internos para ingresar al sistema como *nivel I* de las distintas Comisiones Dictaminadoras aparecen en relación a la especificidad de los productos asociados a las actividades de investigación y/o desarrollo tecnológico. Otra vez aquí, los criterios más generales se observan en las Comisiones Dictaminadoras de las áreas de Física, Matemáticas y Ciencias de la Tierra, y de Ingeniería, que solamente indican que los postulantes deben demostrar que sus aportes han contribuido a la solución de problemas específicos o al alcance de determinados objetivos en la investigación o al desarrollo individual, grupal o institucional (esto último, solo lo señala la Comisión Dictaminadora de Ingeniería, que además agrega que el postulante debe tener una línea de investigación definida).

El requerimiento de tener una línea de investigación definida también es exigido por las Comisiones Dictaminadoras de las áreas de Humanidades y Ciencias de la Conducta y de Ciencias Sociales, que además exigen -en términos de publicaciones de los últimos tres años- un libro de investigación como autor en editorial de reconocido prestigio y con dictamen, o al menos, cinco artículos en revistas científicas arbitradas (nacionales o internacionales) o cinco capítulos de libro, también dictaminados y en editorial de reconocido prestigio. Asimismo, señalan que complementariamente se consideran otras publicaciones tales como libros de investigación coordinados y editados, libros de texto de alta calidad, ediciones críticas, entre otros.

Por su parte la Comisión Dictaminadora de Biología y Química exige la publicación de trabajos de investigación en los últimos tres años y señala que la *calidad* será el criterio predominante en la evaluación, y además que se espera un artículo publicado por año en promedio.

Similarmente o más explícitamente, las Comisiones Dictaminadoras de Medicina y Ciencias de la Salud y de Biotecnologías y Ciencias Agropecuarias exigen que el postulante demuestre continuidad en su producción de los últimos tres y cinco años respectivamente y calidad a través de la publicación de artículos en revistas del JCR; o mediante desarrollos tecnológicos de impacto en el área, protegidos por patentes, registros de autor, secretos industriales, etc.

Además, la Comisión Dictaminadora de Biotecnologías y Ciencias Agropecuarias también considera de calidad, los artículos publicados en revistas del Índice de Revistas Mexicanas de Investigación Científica y Tecnológica del CONACYT, y los libros y capítulos de libro en editoriales de reconocido prestigio, asimismo señala que la contribución del postulante a una línea definida de investigación

será valorada a través de la participación del mismo como primer autor, autor de correspondencia o publicando con sus alumnos tesis.

Al igual que para los casos anteriores, aquí también, al ingreso y para quedar categorizado como *Nivel II* en el sistema, las diversas Comisiones Dictaminadoras de todas las áreas exigen cumplir con los requerimientos señalados para la categoría *nivel I*, y además todas ellas también indican como requisito la contribución de los investigadores en la formación de recursos humanos, con énfasis en la dirección de tesis de posgrado. Simultáneamente, y para todas las Comisiones Dictaminadoras los investigadores que aspiren al *Nivel II* deben por un lado, demostrar consolidación e independencia en su línea de investigación y por otro, contar con liderazgo; las diferencias entre las áreas respecto a estas dos grandes dimensiones se observan según los elementos que las comprueban.

Nuevamente aquí, los criterios menos detallados se observan en las Comisiones Dictaminadoras de las áreas de Física, Matemáticas y Ciencias de la Tierra y de Ingeniería, y los mismos no difieren sustantivamente de lo ya señalado para este nivel.

De acuerdo con los criterios internos del área de Biología y Química, la independencia y la consolidación se aproximan a través de los conceptos de productividad y calidad, que a su vez son medidos a través de publicaciones indexadas en el JCR y, se espera que el investigador tenga acumuladas en el entorno de quince a veinticinco publicaciones. El liderazgo en la Comisión Dictaminadora de esta área se aproxima, según señalan en sus criterios internos, a través de la evidencia que presente el investigador sobre el impacto de sus trabajos en la comunidad científica internacional, específicamente a través de las citas y otros reconocimientos como experto en su área.

Estos lineamientos también están indicados para el caso de los criterios internos de las Comisiones Dictaminadoras de Medicina y Ciencias de la Salud y de Biotecnología y Ciencias Agropecuarias, aunque no especifican la cantidad esperada de artículos publicados, y además agregan a los desarrollos tecnológicos, cuyos comprobantes son el registro de patente, otros registros y cartas de los usuarios. Adicionalmente, el liderazgo y su consolidación -según se indica en los criterios internos de estas áreas-, se aproxima no solamente a través de una alta cantidad de citas a sus trabajos, sino también siendo el primer autor, o autor de correspondencia en al menos un tercio del total de artículos publicados.

Además, se les solicita a los investigadores que son evaluados por la comisión del área de Medicina y Ciencias de la Salud, que demuestren capacidad en la creación de infraestructura para la investigación, mediante la dirección de grupos de investigación y el acceso a fuentes de financiamiento para dicha actividad.

Para las Comisiones Dictaminadoras de las áreas de Humanidades y Ciencias de la Conducta, y de Ciencias Sociales, los criterios son esencialmente los mismos que los señalados para las otras Comisiones Dictaminadoras: amplia cantidad de publicaciones, las que incluyen libros originales de referencia en el campo de especialidad del investigador, amplia cantidad de artículos en revistas de alta calidad, de capítulos de libros y, complementariamente, libros coordinados y editados, ediciones o traducciones críticas de autores y de lenguas.

Los criterios internos de la Comisión Dictaminadora de Humanidades y Ciencias de la Conducta señalan además que el investigador tiene que haber dirigido tesis de posgrado concluidas (ídem en



la de Ciencias Sociales), demostrar la participación en grupos de investigación, participar en tareas de divulgación y demostrar su liderazgo nacional a través de la importancia de su rol en la formación de investigadores o grupos de investigación y mediante las citas a sus trabajos.

Finalmente, aquí también, al ingreso y para quedar categorizado como *Nivel III* en el sistema, las diversas Comisiones Dictaminadoras de todas las áreas exigen cumplir con los requerimientos señalados para la categoría *nivel II* y además, los investigadores que aspiren a este nivel deben demostrar: i) una producción científica (y/o tecnológica) ampliamente reconocida por su impacto nacional e internacional; ii) el reconocimiento como líder nacional e internacional en su disciplina; y iii) un trabajo destacado en la formación y/o fortalecimiento de recursos humanos altamente calificados.

Nuevamente aquí, las Comisiones Dictaminadoras de las áreas de Física, Matemáticas y Ciencias de la Tierra y de Ingeniería, no especifican mucho más de lo ya dicho respecto a los requisitos para este nivel. De acuerdo con los criterios internos de las restantes Comisiones Dictaminadoras por áreas, esas exigencias se traducen en:

- i) una gran cantidad de citas -350 como promedio mínimo-, reconocimientos y/o distinciones en los ámbitos científicos y en la formación de investigadores destacados, tales como los pertenecientes a los niveles I y II del SNI, para el caso del área de Biología y Química;
- ii) excelente productividad y calidad en la investigación, con artículos publicados en revistas del JCR –al menos las dos terceras partes de los artículos del investigador deben estar publicados en revistas con alto factor de impacto de acuerdo con su disciplina, y en el 30% debe ser primer autor o co-autor de correspondencia-, además debe tener una cantidad importante de citas –internacionales-; y/o desarrollos tecnológicos con registros vigentes, impacto demostrado y uso en el área de la salud avalado por cartas de los usuarios. También tienen que tener un rol destacado en la formación de investigadores, en la dirección de tesis de doctorado y en las actividades de docencia. Asimismo, demostrar capacidad en el acceso a fuentes de financiamiento para la investigación y una destacada trayectoria con reconocimientos nacionales e internacionales, para el caso del área de Medicina y Ciencias de la Salud;
- iii) producción de conocimiento publicada en revistas y/o por editoriales de alto prestigio y amplia circulación nacional e internacional, que represente una contribución destacada y sea una referencia en su campo; lo que debe ser demostrado con reconocimientos y/o distinciones académicas, citas a sus trabajos de reconocidos autores internacionales, reseñas y comentarios en revistas de circulación internacional, invitaciones para el dictado de conferencias y/o cursos en el exterior; y mostrar un papel destacado en la dirección de tesis de doctorado y en la integración de grupos de investigación; para el caso de las áreas de Humanidades y Ciencias de la Conducta y Ciencias Sociales; y,
- iv) producción científica ampliamente reconocida por sus citas en el ámbito nacional como internacional; el reconocimiento como líder científico y/o tecnológico, con un desempeño destacado en el desarrollo de líneas de investigación, y/o un impacto por transferencias tecnológicas; y la formación de RRHH, a través de la dirección de tesis de

doctorado y contribución a la formación de grupos; para el caso del área de Biotecnología y Ciencias Agropecuarias.

Finalmente, y de acuerdo con lo señalado en los criterios internos de cada una de las Comisiones Dictaminadoras por área, para permanecer en el sistema y en la misma categoría, los investigadores deben, además de cumplir con los requisitos al ingreso de cada categoría, demostrar continuidad, calidad, alta productividad y, además en el caso de los niveles III, gran liderazgo. En el caso de las Comisiones Dictaminadoras de las áreas de Física, Matemáticas y Ciencias de la Tierra y de Biología y Química, se esperan resultados y productos anuales, lo que representa una presión adicional muy fuerte y difícil, sino imposible de sostener, por lo menos respecto a la producción de conocimiento científicamente relevante, original, válido y confiable.

De todas maneras, y tal como señala Bensusán et al. (2014), en qué medida estos criterios internos son aplicados estrictamente requiere de estudios específicos. La descripción realizada en esta subsección, no alcanza al análisis de las prácticas de evaluación, sino que observa únicamente lo indicado -como requerimientos mínimos de exigencia- en los diversos documentos sobre criterios internos de evaluación, elaborados por las diversas Comisiones Dictaminadoras.

Sin embargo, de acuerdo con los estudios sobre el SNI de México, revisados en la sección IV.3 del capítulo IV, en particular los que analizan el impacto del SNI en la productividad de los investigadores miembros del sistema, muestran el notable aumento de las publicaciones en revistas internacionales incluidas en el JCR. También, es indicativo de la aplicación de estos criterios, los diversos trabajos críticos y de reflexión sobre los efectos del sistema de evaluación y que puntualizan los principales desafíos y/o problemas que enfrenta actualmente el SNI de México que se revisan en el siguiente capítulo.

### VI.3 Conclusión

De acuerdo a lo descrito en la sección (VI.2), se puede sostener que el SNI puede caracterizarse como un sistema de evaluación fuerte.

En primer lugar, el SNI puede definirse como un conjunto organizado de procedimientos para evaluar el desempeño individual de investigación de los investigadores.

En segundo lugar, y desde la perspectiva de su organización y gobernanza, se trata de un sistema cuya evaluación es sistemática y periódica, con amplia formalización de sus procedimientos, con una relativamente alta estandarización de criterios de evaluación comunes entre las áreas de conocimiento y estables en el tiempo; a la vez, también se trata de un sistema que, progresivamente, ha aumentado su transparencia. Dicha transparencia se determina porque las evaluaciones son realizadas en comisiones cuyos integrantes (evaluadores) fueron seleccionados formalmente, con mecanismos de selección públicos y participativos, y además, los resultados de la evaluación también son de conocimiento público, al igual que los criterios de evaluación específicos que dieron lugar a dichos resultados y juicios de las comisiones dictaminadoras.

En tercer lugar, y desde el punto de vista de las implicancias del sistema de evaluación para la asignación de recursos, se puede afirmar la gran importancia que tiene mantener la pertenencia al SNI, no solamente por el impacto directo del estímulo económico en el aumento de los ingresos de los investigadores, sino también e indirectamente por la importancia que tiene la membresía del

SNI para el acceso a fuentes de financiamiento externo para la investigación y la formación de recursos humanos del más alto nivel en los programas de posgrado de calidad.

Asimismo, el SNI puede ser identificado como un sistema de evaluación fuerte porque se trata de una evaluación pública e institucionalizada de la “calidad” de la investigación, que además es desarrollada por la élite científica (los evaluadores son investigadores categorizados en el nivel III, el más alto del sistema), se realiza periódicamente de acuerdo con reglas y procedimientos altamente formalizados, y que responden a criterios estándar de calidad e impacto (indicadores bibliométricos). A la vez, la evaluación está organizada en torno a las disciplinas existentes, y sus resultados, además de ser públicos, afectan directamente y significativamente el ingreso económico de los investigadores al tiempo que, indirectamente, el acceso a fuentes externas de financiamiento para la investigación.

Tal como se planteó en el capítulo IV, la implementación de sistemas de evaluación fuertes como el SNI puede desalentar la producción de conocimiento de largo plazo, de alto riesgo, interdisciplinaria, colectiva, y diversa desde el punto de vista cognitivo; además, puede estimular la estandarización de trayectorias, perfiles u orientación de la investigación.

Dicho de otro modo y de acuerdo con la literatura, los sistemas de evaluación fuertes pueden tener consecuencias sobre ciertas características epistemológicas de la producción de conocimiento. Además dichas consecuencias, y para el caso de México en particular, se pueden ver muy reforzadas debido a algunas características específicas de su sistema de CTI descritas en la sección VI.1. En particular, destacan las asociadas a la importancia del financiamiento público externo para la investigación, basado en la asignación de recursos, a través de la presentación de proyectos de investigación de relativamente corto plazo (de 1 a 3 años) a fondos competitivos, provistos por una única agencia de financiamiento, el CONACYT.

De aquí la relevancia y pertinencia que tiene conceptualizar y medir la flexibilidad epistemológica del sistema de evaluación del SNI que se propone en este trabajo. Es decir, en qué medida el sistema de evaluación permite: i) la obtención de resultados suficientes y relevantes –de calidad-; ii) el reconocimiento de distintos perfiles de investigación; iii) la diversidad disciplinaria; iv) la no estandarización de trayectorias; v) asumir estrategias de investigación innovadoras –novedosas, riesgosas-; vi) involucrarse en investigaciones de largo plazo; y vii) trabajar colectivamente.

Entre otras de las consecuencias probables de los sistemas de evaluación fuertes, identificada en la literatura, está la vinculada con el aumento de la importancia del reconocimiento reputacional, especialmente a nivel internacional. De hecho, y para el caso de México, este elemento forma parte de los criterios internos de evaluación. Los investigadores que aspiren a los niveles más altos del sistema, deben demostrar que gozan de prestigio y reconocimiento internacional, lo que se logra a través de publicaciones en las revistas internacionales con altos factores de impacto y con destacada cantidad de citas internacionales.

Por ello, y de acuerdo con la literatura, también es probable que la implementación de sistemas de evaluación fuertes genere un aumento de la competencia entre los investigadores por publicar en las revistas líderes y para obtener financiamiento en los temas que más se trabajan en la literatura internacional. Esto se refuerza aún más en el caso de México –aunque no es el único país-, porque las medidas de trascendencia e impacto se aproximan por los factores de impacto de las revistas,

las citas recibidas –cuya calidad mejora si provienen de otros investigadores líderes en el campo-, y si provienen del ámbito internacional, la trascendencia y relevancia es mayor. Genera a la vez, una mayor estratificación de investigadores, universidades y revistas; no es casual que en México los investigadores se concentren en muy pocas universidades, y que las revistas internacionales incluidas en el JCR sean prácticamente las únicas válidas para reflejar la calidad de la investigación.

La contextualización realizada a lo largo de este capítulo, también se espera sirva para un mayor entendimiento de las razones detrás de las expresiones de múltiples académicos de México que, a través de diversos trabajos críticos y de reflexión, plantean un conjunto de problemáticas y/o desafíos centrales que deberían ser abordado en el SNI de México. En el siguiente capítulo se revisan estos trabajos.

Finalmente, asociar al SNI a un sistema de evaluación fuerte, permite además una mejor comprensión de los resultados empíricos obtenidos en este trabajo en relación a la satisfacción con el proceso de evaluación y la flexibilidad epistemológica del sistema de evaluación, y también respecto a los observados en los ítems específicos explicados por estas dimensiones conceptuales. Simultáneamente, permite entender mejor por qué para algunos investigadores categorizados en el SNI es importante reformar el sistema de evaluación del SNI en una dirección que reformule sus criterios de evaluación en una dirección más cualitativa, que modifique la composición de las comisiones dictaminadoras y que valore de mejor manera la investigación colectiva, multi/inter/transdisciplinaria, y la orientada a la aplicación. Estos resultados se presentan en el capítulo IX.

## CAPITULO VII: Antecedentes: el debate sobre el SNI en el ámbito nacional

El SNI ha sido objeto de intensos debates en el ámbito nacional debido a sus implicancias potenciales sobre el contenido y dirección de la producción de conocimiento y sus características epistemológicas, la relevancia de las diversas actividades académicas y la asignación de recursos públicos para la investigación. Esto se debe en definitiva, como señalan Didou y Gérard (2010), a que "(...) el SNI representa un mecanismo central de estructuración del campo científico en lo que refiere a sus integrantes, a sus elites y a sus excluidos" (Didou y Gérard, 2010: 119).

El objetivo de este capítulo es establecer los fundamentos, derivados del ámbito nacional, que le dan sentido y permiten conceptualizar -en su contexto- el modelo teórico de ecuaciones estructurales desarrollado en esta tesis, sobre las relaciones entre la evaluación del desempeño de la investigación individual y las formas diferenciales de producción de conocimiento, que se operacionaliza y pone a prueba a partir de la evidencia que provee el caso del SNI de México.

Para alcanzar dicho objetivo, y debido a la falta de una teoría robusta sobre el tema para el caso del SNI mexicano, también se recurrió a la revisión del debate sobre el SNI en el ámbito nacional. De aquí también la relevancia teórica conceptual del estudio que se realiza en esta tesis, ya que es el primero en analizar la relación entre la evaluación del desempeño de investigación de los investigadores y la producción de conocimiento. Específicamente se estudia la satisfacción con el proceso de evaluación del SNI, su flexibilidad epistemológica y la dirección de una eventual reforma al mismo dependiendo de dos modos diferenciales de producción de conocimiento. A la vez, los resultados obtenidos además de originales, válidos y confiables, debido al enfoque teórico-metodológico-empírico desarrollado, son pertinentes por sus implicancias para el diseño e implementación de instrumentos de reconocimiento y valoración de la producción de conocimiento como lo es el SNI.

A partir de esos debates, de algunos de los análisis disponibles sobre el SNI, de sus logros, desafíos y direcciones de cambio, y de las ideas, percepciones, experiencias y visiones de hacedores de políticas públicas, gestores de la ciencia, implementadores de políticas, autoridades, asesores, evaluadores e investigadores, es posible identificar un conjunto de relaciones probables/esperables. Estas sirven como antecedentes de la teoría, representada por el modelo, que se contrasta en esta tesis, cuyo marco teórico conceptual se desarrolló en el capítulo V y los resultados del modelo empírico se presentan en el capítulo X.

En particular se revisaron dos grandes debates. El primero en 2005 con motivo de los veinte años del SNI, que fue co-organizado por la Academia Mexicana de Ciencias (AMC) y el Foro Consultivo Científico y Tecnológico (FCCyT). El segundo en 2012, casi una década después, organizado por la Universidad Autónoma Metropolitana, Unidad Xochimilco (UAM-X). La elección de estos dos debates se basa en dos razones, una práctica y otra sustantiva.

La razón práctica se vincula con el acceso a todas las ponencias y materiales en torno a los cuales se dio la discusión y el debate, ya que los mismos fueron publicados como libros. El primero bajo el título "*Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación*" (FCCyT-AMC, 2005). El segundo titulado "*Sistema Nacional de Investigadores. Retos y perspectivas de la ciencia en México*" (Vega y León, 2012).

La razón sustantiva tiene que ver con la pertinencia de los temas centrales de ambos debates para los objetivos de esta tesis y, además, al estar separados por casi una década, es posible observar la evolución en el tiempo de los temas en debate. También permite una valoración sobre la consistencia, entre las perspectivas de los autores y a lo largo del tiempo, de lo que han denominado como logros, desafíos y transformaciones requeridas del SNI.

El capítulo se estructura en cinco secciones. Le sigue a esta introducción la sección VII.1, en la que se describen, a juicio de algunos autores, los logros del SNI. En la sección VII.2, de acuerdo con la revisión del debate, se presentan algunos desafíos para el SNI. En esta sección, si bien se tratan diversos aspectos vinculados con los desafíos que enfrenta el sistema, el foco se pone principalmente en lo que en el contexto de esta tesis se ha denominado como la flexibilidad epistemológica del sistema de evaluación, en particular respecto a la relación entre la evaluación y la diversidad de la producción de conocimiento, su calidad y su escasa contribución a la resolución de los problemas del desarrollo social y económico de México. En la sección VII.3 se plantean dos dimensiones: la inconsistencia de la política pública y ciertas características específicas del modelo de evaluación del SNI que, en interacción, sirven para entender tanto los logros del sistema como los desafíos que enfrenta. En la sección VII.4 se sistematizan un conjunto de propuestas –que emergen de estos debates– sobre la dirección que debería asumir una eventual reforma al SNI para atender a sus desafíos. Finalmente, la sección VII.5 concluye.

## VII.1 El SNI y sus logros: algunas consecuencias performativas de la implementación de un sistema de evaluación fuerte

De la revisión del debate en el ámbito nacional sobre el SNI en México, se identificaron -de acuerdo a la visión y experiencia de ciertas autoridades, evaluadores e investigadores un conjunto de logros del sistema, que con fines analíticos se sistematizaron en seis categorías, a saber: i) institucionalización de la evaluación, retribución económica según desempeño y prestigio; ii) introducción de estándares internacionales de evaluación; iii) profesionalización de la investigación, una actividad altamente normada; iv) fortalecimiento del sistema científico nacional; v) retención de investigadores en el ámbito académico nacional; y vi) mejoramiento del proceso de evaluación del SNI. A continuación se presentan en las sub-secciones VII.1.1 a VII.1.6.

### VII.1.1 Institucionalización de la evaluación, retribución económica según desempeño y prestigio

Desde la perspectiva de los autores, a partir del SNI se *institucionalizó la evaluación* de desempeño, no sólo de la investigación sino que también de otras actividades académicas, en particular de la actividad de enseñanza, sobre todo a nivel de posgrado. Además, otorgó retribuciones económicas diferenciales según desempeños desiguales. Asimismo, estar categorizado en el sistema es una distinción y provee prestigio.

Adicionalmente, conforme pasó el tiempo, la evaluación de desempeño de investigación de los investigadores a través del SNI, se convirtió paulatinamente en una medida para aproximar la calidad y prestigio de las organizaciones empleadoras de los investigadores categorizados en el sistema, y también de los programas de formación de posgrado.

Si bien antes del SNI en México existía la evaluación externa, el SNI fue el primero en introducir un mecanismo de alcance nacional de reconocimiento a los investigadores y retribución económica

basado en su desempeño de investigación (Malo, 2005). El SNI se ha convertido en una pieza importante de la evaluación de la investigación del país, diseñado para evaluar individualmente el desempeño de los investigadores con pagos extra-salariales por rendimiento (Peña, 2005b).

Pérez Tamayo (2005) destaca que lo valioso del SNI es justamente el mecanismo de premiación extra-salarial vinculado a la productividad de los investigadores, que a su vez es proporcional -y en forma escalonada según niveles de categorización- al rendimiento de los investigadores, y que los juicios para determinar el ingreso de los investigadores al sistema -en cada uno de los niveles de categorización-, su permanencia y/o ascenso, es realizado de forma periódica por comisiones integradas por pares expertos en las disciplinas que califica como *pertinentes*. Como se verá más adelante en este capítulo, otros autores son críticos respecto a la instalación de pagos extra-salariales por productividad.

A 20 años de la creación del SNI, González Valenzuela (2005) entre otros, destaca que el SNI ha sido muy importante para crear una cultura de evaluación, lo cual permitió aumentar el nivel académico y el avance de las ciencias y las humanidades. Así, formar parte del SNI es una distinción para los investigadores (Hernández-Calderón, 2005), y también representa una medida de referencia para dar cuenta de la calidad de las instituciones que los albergan (Hernández-Calderón, 2005), fundamentalmente de las Instituciones de Educación Superior (Peña, 2005b). Una década más tarde, a casi treinta años de la creación del SNI, Rosales (2012) coincide en esta apreciación y señala que el número de investigadores categorizados en el SNI se ha convertido en un indicador para aproximar el prestigio de las organizaciones que los emplean, facilitando con ello el acceso preferencial a recursos económicos públicos para la investigación (Esteinou Madrid, 2012).

En síntesis, el SNI identificó y reconoció a los académicos que se dedican profesionalmente a la investigación (Pescador Osuna, 2012) y los integró en un sistema, agrupando así a la comunidad científica del país (Esteinou Madrid, 2012). Para ello, introdujo una cultura específica de evaluación, que se basa en un modelo de medición de la productividad de los investigadores y su calidad de acuerdo a una valoración estándar de la producción de conocimiento (Esteinou Madrid, 2012). Así el SNI institucionalizó la evaluación por pares del desempeño de los investigadores a nivel nacional, como mecanismo reconocido y aceptado (Bensusán et al., 2014; Pescador Osuna, 2012; Álvarez Gallegos, 2012; Reséndiz Núñez, 2005), otorgó un complemento económico extra-salarial al ingreso de los académicos y distinguió su labor de investigación con un reconocimiento de gran prestigio (Pescador Osuna, 2012).

Finalmente, Peña (2005b), desde un enfoque probablemente más asociado al *new public management*, sostiene que la evaluación del SNI no se ha generalizado del todo, en la medida que no ha tenido el impacto, la utilidad y *“las consecuencias lógicas de todo proceso de evaluación, para tomar decisiones y corregir rumbos”* (Peña, 2005b: 117). Desde un enfoque conceptual claramente diferente, Esteinou Madrid (2012) destaca que el SNI y sus resultados son usados, por un lado, por las Instituciones de Educación Superior como indicador de prestigio y calidad para respaldar la obtención de financiamiento público y, por otro, por el Gobierno para posicionar al país en los índices de desarrollo internacional que exige la OCDE, pero nunca como un instrumento útil para orientar estratégicamente las políticas de producción de conocimiento hacia los objetivos del desarrollo nacional.

### VII.1.2 Introducción de estándares internacionales de evaluación

Un reto importante para los sistemas de evaluación del desempeño de la investigación, es justamente determinar cuáles son –o deberían ser- los criterios de evaluación que permiten discriminar entre los investigadores y seleccionar a los mejores, atendiendo a los objetivos que se plantea dicho sistema.

La experiencia del SNI en México ha sido radical en ese sentido por dos razones. La primera porque la percepción que tienen autoridades, evaluadores y algunos académicos del ámbito nacional es que los criterios de evaluación del SNI, no solamente discriminan entre los investigadores y seleccionan a los mejores, sino que se asume que los que no son seleccionados por este sistema de evaluación es porque sencillamente no son investigadores. La institucionalización del sistema de evaluación ha sido tal que, *estar o no estar* categorizado en el SNI implica –además- *ser o no ser investigador*.

La segunda razón de la radicalidad de la experiencia del SNI se asocia con la relativa desvinculación entre los fundamentos y objetivos del SNI y sus criterios de evaluación e indicadores, utilizados para determinar el desempeño de la investigación y su calidad (ver sub-secciones VI.2.1 y VI.2.3 del capítulo VI). Debido a que descansa en la introducción de criterios y estándares internacionales de evaluación que, como se verá a lo largo de este capítulo y también se discutió en el capítulo IV, pueden no ser los mejores –si son los únicos- para alcanzar los objetivos que se plantea el SNI.

Ruiz Herrera (2005), que ha formado parte de una de las comisiones dictaminadoras en la primera convocatoria del SNI y que veinte años después volvió a actuar como evaluador del sistema, reflexiona sobre la sistematicidad y mantenimiento de los mismos criterios de evaluación a lo largo del tiempo, y concluye que son esos los criterios que determinan la calidad de la investigación. Desde su perspectiva, y en coincidencia con otros autores, el SNI ha sido clave para definir conceptualmente lo que significa ser investigador. Más aún, sostiene que si el académico no forma parte del SNI, no es un investigador (Ruiz Herrera, 2005).

Las palabras de uno de los fundadores y creadores del SNI, que en su mayoría pertenecen a las ciencias exactas y naturales, son esclarecedoras sobre los fundamentos que dan sentido a los criterios de evaluación implementados en el SNI, y de alguna manera, también permiten entender su relativa desvinculación de los objetivos centrales del sistema. A casi treinta años de la creación del sistema, Flores Valdés (2012) señala que la experiencia del SNI muestra que las reglas y los criterios de evaluación son muy importantes para orientar la investigación, y que al inicio –él y otros, los fundadores y creadores del sistema- observaron “*mucha farsa en la producción científica*”, entonces dijeron: “*(...) se necesita que la gente publique para ser miembros del SNI, (...) en revistas de a de veras, o sea en las revistas que tienen unas reglas de selección muy grandes*”, y la gente, finalmente, lo hizo (Flores Valdés, 2012:14-15).

En definitiva, esta forma de plantear la definición de los criterios de evaluación, trasluce –al menos- que los mismos fueron definidos atendiendo más a la detección de *falsos profetas* que al alcance integral de los objetivos planteados en el Acuerdo de creación del SNI.

Siete años antes de las expresiones de Flores Valdés (2012), Peña (2005b) destacaba que el mérito más importante del SNI fue distinguir la *producción primaria esencial* –cantidad y calidad de publicaciones- de la *producción secundaria* -docencia, divulgación, vinculación, construcción



institucional, etc.. Esta última no sustituye a la primaria porque *“un investigador que no publica es como si no trabajara”* (Peña Díaz, 2005:26).

De hecho en ese momento, el trabajo de Ruiz Herrera (2005) -entre otros- ya mostraba, según datos de CONACYT, el aumento en el número de publicaciones anuales, en la cantidad de artículos por investigador y en el número de citas recibidas como indicadores de productividad, calidad e impacto de la producción de conocimiento nacional, y adjudicaba esos resultados al SNI por su estímulo a la *productividad científica de calidad*.

Más aún, Peña (2005b) destaca como positivo que los criterios de evaluación del SNI tienen validez universal<sup>46</sup>, y no están sujetos a la utilidad inmediata del conocimiento sino a su *calidad*. Agrega, que “[los criterios de evaluación] *no se someten a las necesidades de una región, un estado o una nación, y menos a lo que no sea investigación, sino simple aplicación de conocimiento existente para resolver un problema. Explorar el país o mejorar su estado de salud, deberían ser tarea de instituciones gubernamentales y no de los investigadores, si no es como asesores de alto nivel”* (Peña, 2005b: 117).

En una línea similar a Peña (2005b), Mulás del Pozo (2005) sostiene que *“Los individuos dedicados al desarrollo tecnológico y estudios de orientación aplicada, si bien se pueden llamar investigadores, deben ser recompensados por las personas físicas o morales (clientes) que encargan y costean los mismos, dependiendo del grado de satisfacción que obtengan del resultado. El modelo de SNI no es la forma de hacer progresar las actividades de desarrollo tecnológico. El premio por realizar con excelencia una actividad, debe estar alineado con la misma para que florezca”* (Mulás del Pozo, 2005:127).

En efecto, diversos autores coinciden en el acierto del SNI al introducir criterios de evaluación homogéneos<sup>47</sup>, definidos e implementados por pares (Membrillo Hernández, 2005), que permiten la generación de medidas equivalentes de productividad en cada rama de la ciencia (Reséndiz Núñez, 2005), basados en indicadores y estándares internacionales para evaluar el desempeño de la investigación y su calidad (Asomoza Palacio, 2005; Reséndiz Núñez, 2005).

Serra Puche (2005) también destaca que luego de veinte años del SNI, los criterios para valorar la calidad de los productos de investigación son cada vez más rigurosos; las dictaminaciones, evaluaciones externas y la mirada sobre el impacto de las publicaciones, son permanentes y forman parte del quehacer cotidiano. De ese modo, y desde la perspectiva de la autora, dichos criterios están legitimados por la comunidad científica en cualquier ámbito.

Sin embargo, Hernández-Calderón (2005) si bien coincide respecto a la contribución del SNI en el establecimiento de esos criterios, y destaca que los mismos han sido útiles para la superación de

---

<sup>46</sup> La literatura internacional especializada cuestiona, entre otras cosas, justamente la *validez universal* de los indicadores bibliométricos como medidas para aproximar una medida de la calidad e impacto de la producción de conocimiento (ver capítulo IV).

<sup>47</sup> Uno de los temas que genera más tensión en el debate internacional, y que la literatura especializada ha cuestionado fuertemente, es justamente el uso de criterios e indicadores homogéneos. Entre otras cosas porque esos criterios e indicadores no permiten apreciar la diversidad de productos derivados de la investigación en los diversos campos científicos, ni valorar adecuadamente su calidad, incluso en el ámbito de la ciencia académica. Resulta elocuente la expresión que utiliza Hicks (2013) *“One size doesn't fit all”* cuando analiza el caso de las ciencias sociales y humanidades.

muchos investigadores, a la vez considera que también son cuestionables. En particular por lo que dice Membrillo Hernández (2005) cuando plantea que “[e]l SNI debe reconocer a la labor científica pero en una forma integral. Un científico no es solo publicaciones” (Membrillo Hernández, 2005: 107).

En una línea similar, pero casi una década después, Rosales (2012) destaca que el SNI estableció la diferencia entre *ser y no ser* reconocido como investigador nacional, a través de “mecanismos objetivos” para determinar que el ingreso, permanencia y ascenso -dentro del sistema- se asocie con ciertos niveles de productividad (Álvarez Gallegos, 2012). Argumenta que si bien el sistema ha sido aceptado, también ha sido ampliamente cuestionado por sus dificultades para reconocer la calidad de la producción de conocimiento realizada por los investigadores (Rosales, 2012).

Así, el SNI propició una cultura de la eficiencia y la productividad, privilegiando la publicación de artículos científicos en revistas internacionales arbitradas con alto factor de impacto (Pescador Osuna, 2012) como medida para aproximar la calidad, es decir, introduciendo un modelo para medirla periódicamente, transformando con ello el trabajo académico (Esteinou Madrid, 2012).

De Ibarrola (2012) es clara en la descripción de los ítems que más valora el modelo de evaluación del SNI y en la especificación de algunas de sus consecuencias. En particular destaca que el SNI pondera mejor los reconocimientos internacionales a la productividad, tales como: doctorados obtenidos en el extranjero, artículos publicados en revistas de circulación internacional, citas recibidas, pero realizadas -preferentemente- por investigadores reconocidos internacionalmente en las distintas áreas del conocimiento, y las redes internacionales de los investigadores mexicanos. De acuerdo con la autora, el SNI representa el mecanismo más poderoso para jerarquizar la profesión académica, y ubica a los investigadores en el nivel más alto. Permite la diferenciación de ingresos entre los académicos al interior de las instituciones según su desempeño individual, tiene dinámica propia respecto a la legitimación y prestigio, la evaluación es realizada directamente por *pares* y descansa indirectamente en los *pares internacionales*, y genera tensión económica y de prestigio entre docentes e investigadores (De Ibarrola, 2012).

### VII.1.3 Profesionalización de la investigación: una actividad altamente normada

Diversos autores coinciden en que el SNI contribuyó significativamente a la profesionalización de la actividad de investigación y del sistema científico mexicano (Drucker Colín, 2005b; Membrillo Hernández, 2005), a través de la creación de un sistema que provee reconocimiento, prestigio y estímulos económicos (Esteinou Madrid, 2012). De ese modo, fomentó y estimuló el trabajo de investigación (Pescador Osuna, 2012).

El SNI permitió que los investigadores entendieran cómo se hace ciencia, qué hay que hacer, qué se espera de ellos para ser considerados como tales (Drucker Colín, 2005b), y, de acuerdo con Reséndiz Núñez (2005), el SNI es un faro que orienta la acción de individuos e instituciones para obtener reconocimiento científico. Además, los jóvenes que inician su carrera de investigación ya saben *ex ante* qué se espera de ellos (Reséndiz Núñez, 2005).

En coincidencia, Porter (2012) indica que el SNI ha guiado no solamente el trabajo académico de los miembros del sistema sino también de los que aspiran a ingresar. Según este autor, todos “*tienden a acordar, reportar y, en suma, trabajar, con los ojos puestos en el SNI y sus requisitos*” (Porter, 2012:86). El SNI introdujo la figura de *investigador nacional* con cuatro categorías (*candidato, nivel*

I, nivel II y nivel III), con requisitos académicos diferenciales para cada una y, gradualmente, se transformó en parte importante de los méritos que norman la carrera académica e intelectual de los investigadores (Esteinou Madrid, 2012).

En el campo de las humanidades y las ciencias sociales, Serra Puche (2005) es contundente: los sistemas de evaluación han modificado la organización de la investigación en estos campos. Es decir, dichos sistemas han promovido una comunidad académica más homogénea, con publicaciones de calidad, generando nuevas pautas culturales entre los investigadores, más universales sobre el quehacer de la investigación.

#### VII.1.4 Fortalecimiento del sistema científico nacional

A veinte años de la creación del SNI, Drucker Colín (2005b) califica el sistema científico nacional como de excelencia y competitivo y, destaca el papel significativo del SNI en su fortalecimiento, elevando la cantidad de artículos en editoriales de prestigio y en revistas indexadas internacionalmente y, con ello, aumentando la cantidad de citas. En general, estas ideas se comparten en las ciencias físicas, matemáticas, biología, química, ciencias de la salud y en menor medida en las ingenierías.

Para el caso de las ciencias sociales y humanidades, Serra Puche (2005), sostiene que el SNI ha tenido un impacto muy positivo en el desarrollo de los procesos de investigación de esas áreas. Esta autora destaca como impactos positivos del SNI que: i) los objetivos de investigación son más claros, viables y significativos para el desarrollo teórico-disciplinar y de la sociedad; y, ii) los procesos de investigación se caracterizan por tener duraciones más limitadas, convergentes y más especializados en ciertas disciplinas y temas. Lo que permite la obtención de resultados más rápidamente y que se expresa en el aumento del número de publicaciones.

Interesa señalar que lo que expresa Serra Puche (2005) como “positivo” es ampliamente cuestionado por la literatura especializada internacional y por diversos estudios realizados en el ámbito nacional, tanto para el caso específico de las ciencias sociales y humanidades, como en general para el desarrollo de la ciencia (ver capítulo IV).

En sintonía con Drucker Colín (2005b), Hernández-Calderón (2005) identifica el aumento de la producción científica –entendida como producción de artículos científicos– como un resultado positivo derivado del SNI, y González Valenzuela (2005) señala la contribución del SNI al colocarla al nivel –o muy cercanamente– de la producción de conocimiento de los países más avanzados.

González Brambila (2005) en su tesis doctoral estudia los determinantes de la productividad científica de los investigadores del SNI, o que han sido parte del sistema en algún momento, en el período comprendido entre los años 1991 y 2001 (14328 investigadores). La medida de productividad que utiliza, la aproxima a través del número de publicaciones que están contenidas en el *Science Citation Index* (SCI) del *Institute for Scientific Information* (ISI). Para el estudio de los determinantes de la productividad utiliza una sub-muestra de 7793 investigadores SNI o que han sido parte del sistema y que cuentan con al menos una publicación en el SCI en el período señalado. Entre sus resultados, la autora destaca que:

- i) se confirma la relación cuadrática entre la edad y la productividad de los investigadores, y sugiere que el SNI ha creado incentivos que inducen a que los investigadores publiquen hasta el final de sus carreras laborales<sup>48</sup>;
- ii) existen diferencias significativas de productividad según el área de conocimiento, tanto respecto al número máximo de publicaciones como en el ciclo productivo, lo que sugiere que los incentivos para publicar en el ISI varían entre los investigadores dependiendo del área de conocimiento: mientras que los investigadores de Ciencias Exactas, Biología, Química y Ciencias de la Salud publican más, por más años y desde que son más jóvenes, los de Ciencias Sociales, Humanidades, Ciencias Agrícolas, Biotecnología e Ingeniería eligen temas de investigación con intereses más locales, cuyos resultados tienden a difundirse en otros medios –menos globales, y más asociados a las audiencias locales/regionales-;
- iii) las mujeres investigadoras son algo menos productivas que sus pares hombres, y hay variaciones según el área de conocimiento;
- iv) el país donde los investigadores obtuvieron el doctorado no presenta diferencia significativa;
- v) el nivel de categorización de los investigadores es importante y el salto en la productividad se observa cuando los investigadores ascienden de candidato a nivel I, y cuando se observa *calidad* (medida por el número de citas en el SCI), los nivel III son los más citados; y,
- vi) los resultados también indican que los investigadores que se graduaron después de 1990 son algo más productivos que las generaciones de los ochenta y setenta, pero los investigadores que se graduaron antes de 1970 son el grupo más productivo.

Finalmente, González Brambila (2005) reconoce que las publicaciones no son el único producto de investigación relevante, y que para obtener un estudio más completo de la productividad de los investigadores también hay que analizar otros productos, tales como: libros y capítulos, desarrollos tecnológicos, etc.. El estudio realizado por Rivera et al (2011), además de coincidir en esa apreciación –como ya fue señalado en el capítulo anterior-, muestra la importancia de otros productos derivados de la investigación en el caso de las ciencias agropecuarias, tales como el desarrollo de nuevas recomendaciones y nuevas técnicas para el sector productivo nacional.

Casi una década después, Fortes Besprosvani (2012) reitera lo extraordinario que ha sido el SNI como instrumento para el impulso de la investigación de alto nivel en México y el aumento de la productividad (Vega y León, 2012). Más específicamente, Esteinou Madrid (2012) señala que el SNI

---

<sup>48</sup> Probablemente el SNI induce a que se publique hasta el final porque perder la membresía implica también una pérdida significativa de ingreso y prestigio; además, tal como señala González Brambila (2005), estos estímulos económicos sumados a los otros disponibles en las instituciones de adscripción alcanzan al 60% del ingreso total de los investigadores. A la vez, al ser extra-salariales no forman parte de las prestaciones sociales como la jubilación. En ese marco, el SNI induce y su efecto se refuerza por la alta precariedad salarial de los investigadores, que los induce a no jubilarse, porque hacerlo no solo significaría renunciar al prestigio y todos los estímulos, sino que también implicaría sobrevivir con menos de la tercera parte de lo que viven sin jubilarse.

incrementó la producción de artículos especializados y su difusión en revistas internacionales con alto factor de impacto.

En relación a la contribución del SNI a la formación de recursos humanos, Drucker Colín (2005) indica que si bien hay que continuar los esfuerzos, considera que el SNI ayudó a la formación de alto nivel. Similarmente, pero para el caso de las ciencias sociales y humanidades, Serra Puche (2005) remarca la importancia del sistema y su efecto positivo respecto a la conformación de grupos de investigación y el *aprendizaje compartido* debido a la incorporación de becarios, tesistas, ayudantes de investigación, estudiantes de (pre)grado, posgrado y servicio social a las actividades de investigación, destacando la complementariedad entre la formación escolarizada y una más académica y profesional. Adicionalmente, Ruiz Herrera (2005) y Malo Álvarez (2005) señalan que el SNI ha contribuido al fortalecimiento de los posgrados de excelencia y de las instituciones (universidades, centros de investigación, etc.).

Vega y León (2012) destaca el aporte del SNI al crecimiento de científicos de primer nivel. En la misma dirección, Esteinou Madrid (2012) señala el aporte del SNI a la formación de una nueva generación de profesores e investigadores en las diversas áreas del conocimiento, el fortalecimiento del plantel académico de las universidades dedicados a la investigación y su contribución a la existencia de posgrados de mayor calidad. Adicionalmente, Álvarez Gallegos (2012) señala que el SNI promovió el incremento del número de investigadores en diferentes instituciones del país, a través de mecanismos específicos para contribuir a la descentralización de la ciencia en México (Esteinou Madrid, 2012).

#### VII.1.5 Retención de investigadores en el ámbito académico nacional

Diversos autores coinciden en que el SNI fue creado para mitigar, en el ámbito académico, las consecuencias de la enorme crisis social y económica que enfrentaba México a inicios de la década de los ochenta. En ese marco, el Estado, ante la imposibilidad de atender las demandas salariales del sector académico y la necesidad de retener a los investigadores en el ámbito académico nacional, creó el SNI como un instrumento diseñado para otorgar incrementos económicos en los ingresos de los investigadores, discriminados según productividad, extra-salariales y temporales.

De acuerdo con Asomoza Palacio (2005), el SNI fue creado para retener a los investigadores en tiempos de crisis. Desde la perspectiva de Ruiz Herrera (2005) y Martínez Palomo (2005) ha sido una excelente estrategia para lograrlo a través de los aumentos discriminados de sus ingresos mensuales. De ese modo, el SNI contribuyó a mitigar el deterioro de los ingresos de los investigadores, y con ello, tanto a la repatriación de científicos (Álvarez Gallegos, 2012) como a retener a los investigadores en el país, evitando parcialmente la fuga de cerebros (Esteinou Madrid, 2012).

A medida que pasó el tiempo, por un lado, el mecanismo de pagos económicos según desempeño se difundió a otras actividades académicas, en particular a las de enseñanza a través del PROMEP; y por otro, los mismos se convirtieron en componente esencial de los ingresos económicos de los investigadores. Para el 75% de los investigadores categorizados en el SNI, los estímulos económicos extra-salariales que perciben multiplican por dos o más sus ingresos salariales anuales (Bensusán, Ahumada e Inclán, 2013). De aquí la importancia del SNI para retener a los investigadores en el ámbito académico nacional. Sin embargo, esto también tiene otras consecuencias sobre la planta

académica nacional, el mercado de trabajo de los investigadores y la concentración de recursos económicos para la investigación.

Desde el punto de vista económico los investigadores están mejor con SNI que sin él (Membrillo Hernández, 2005). En coincidencia González Valenzuela (2005) resalta la importancia del estímulo económico que provee el sistema, como complemento fundamental tanto para investigadores como para las instituciones de educación superior. Adicionalmente, para Hernández-Calderón (2005) el SNI ha logrado la instalación de pagos económicos por productividad y, de acuerdo con Vega y León (2012), si bien se pensó como un mecanismo para premiar la productividad individual, a lo largo de los casi treinta años desde su creación, se ha convertido en un mecanismo para que los investigadores mantengan un ingreso decoroso.

De Ibarrola (2012), que coincide con los análisis anteriores, también es crítica y reflexiona sobre el impacto de los pagos extra-salariales temporales por méritos, nacidos con el SNI y difundidos a la docencia a través del PROMEP, lo que para ella ha determinado enormes diferencias y desigualdades en los ingresos de los académicos. De acuerdo con esta autora, dichos pagos han sustituido los aumentos salariales, lo que genera una enorme presión respecto al envejecimiento del plantel académico, y la no liberación de puestos de trabajo por jubilación, limitando con ello el crecimiento de la investigación nacional y la renovación de los académicos en los puestos de trabajo. Como resultado, el ingreso al mercado de trabajo de la investigación, y al propio SNI, es cada vez más rígido y la competencia cada vez es mayor. A la vez y debido a la falta de oportunidades laborales, México ha perdido algunas generaciones de investigadores. Esteinou Madrid (2012) coincide con el impacto de los estímulos económicos extra-salariales y sus consecuencias sobre el envejecimiento y la dificultad de renovación de la planta académica. Ruiz Gutiérrez (2012) señala además la escasa incorporación de jóvenes y mujeres a la investigación a través del SNI. Para un estudio específico sobre el envejecimiento académico y las perspectivas de jubilación de los miembros del SNI, ver Bensusán, Ahumada e Inclán (2013).

Además, de acuerdo con De Ibarrola (2012), al tiempo que el SNI crea condiciones para la investigación, también genera un importante *Efecto Mateo*, ya que los investigadores categorizados en él tienen acceso preferencial en la asignación de recursos a través de otros fondos competitivos para la investigación (Rosales, 2012) y otros beneficios (ver Didou y Gerárd, 2010). Esto dificulta aún más el ingreso de los investigadores no categorizados en el sistema a las actividades de investigación. De acuerdo con Esteinou Madrid (2012), el SNI no fomenta el avance de la ciencia sino “*la competencia por recursos económicos para sobrevivir*” (Esteinou Madrid, 2012: 154).

#### VII.1.6 Mejoramiento del proceso de evaluación

Desde el acuerdo de creación y hasta la actualidad, la normatividad del SNI y su reglamento de operación y funcionamiento han sido modificados en diversas ocasiones como resultado de los cuestionamientos que ha recibido, procurando una mayor *satisfacción con el proceso de evaluación* –en términos de este trabajo- y su mejora continua (ver capítulo VI, sub-secciones VI.2.1 y VI.2.2).

Entre los elementos que a lo largo de la historia del SNI han sido críticos, a pesar de los esfuerzos liderados por el CONACYT, están los que se vinculan con las comisiones dictaminadoras del sistema, en particular respecto a la representatividad institucional, disciplinar, regional y de género de su integración; la adecuación de los criterios específicos para seleccionar a los evaluadores; y la transparencia del proceso de evaluación y de sus criterios de evaluación (Bensusán et al., 2014).

Distintos autores han coincidido en que en el SNI se han implementado diversos cambios con el fin de mejorarlo en reiteradas oportunidades. En particular Álvarez Gallegos (2012) –entre otros– reconoce que la ampliación del número de comisiones dictaminadoras según áreas de conocimiento del SNI, facilitó la evaluación especializada y pertinente. Si bien es importante continuar los esfuerzos en esta dirección, incluyendo la diversidad cognitiva –científica y tecnológica– contenida en el país, el SNI fue avanzando en la definición de los campos científicos abarcados en su proceso de evaluación (Esteinou Madrid, 2012).

La publicación de los criterios internos de evaluación de las diversas comisiones dictaminadoras y de los nombres de sus miembros, ha dotado al sistema de una mayor transparencia de su proceso de evaluación que, junto con el recurso de reconsideración de los resultados de evaluación para atender las disconformidades y el recambio de evaluadores, contribuye a reducir resultados de evaluación inapropiados o erróneos (Álvarez Gallegos, 2012; Esteinou Madrid, 2012).

A la vez, también se reconoce que los dictámenes han servido para orientar a los investigadores a fortalecer ciertos aspectos de su actividad para lograr la permanencia y ascenso en el sistema (Álvarez Gallegos, 2012; Esteinou Madrid, 2012).

## VII.2 El SNI y sus desafíos: otras consecuencias constitutivas de la implementación de un sistema de evaluación fuerte

### VII.2.1 El SNI y su flexibilidad epistemológica: la diversidad institucional, de actividades y de la orientación de la producción de conocimiento cuestionada

El SNI, con el propósito de reconocer y premiar económicamente la labor de investigación y el desempeño de los investigadores, ha utilizado ciertos indicadores para operacionalizar sus criterios de evaluación. Simultáneamente, si bien tales indicadores logran su propósito, también parecen inducir ciertos *efectos constitutivos* sobre: i) la diversidad de las organizaciones dedicadas –entre otras cosas– a la investigación; ii) la integración y equilibrio de sus actividades académicas; y, iii) la pluralidad de enfoques y orientación predominante de la producción de conocimiento.

Distintos autores coinciden en que el SNI, a pesar de los esfuerzos realizados, favorece la concentración de la *ciencia* (Ruiz Herrera, 2005) y de investigadores y de los posgrados de calidad (Ruiz Gutiérrez, 2005) en ciertas instituciones y regiones del país, con la inequidad que ello implica (Drucker Colín, 2005) en términos –por ejemplo– del acceso a fuentes de financiamiento público para la investigación, y su *Efecto Mateo* asociado.

De acuerdo con Asomoza Palacio (2005), si bien la concentración sigue siendo alta, la brecha regional muestra una tendencia a la baja desde 1994. Mientras que en ese momento el 56% de los investigadores categorizados en el SNI se concentraban en el Distrito Federal (en la actualidad Ciudad de México), hacia 2003 ese porcentaje se redujo al 47%. La UNAM, UAM, CINVESTAV y el IPN son las organizaciones que concentraron a la mayor cantidad de investigadores del SNI en el período comprendido entre 1991 y 2003. (Asomoza Palacio, 2005).

Más recientemente, Vega y León (2012) vuelve a sugerir la concentración de los investigadores en pocas instituciones y el centralismo de la producción de conocimiento nacional como consecuencia del SNI. Asimismo, Álvarez Gallegos (2012), en base a datos del SNI del año 2011, muestra la tendencia concentrada e inequitativa de la distribución de investigadores categorizados en el

sistema según estados, instituciones y género. Además muestra que dichas inequidades se amplían conforme aumenta el nivel de categorización de los investigadores. Los investigadores de mayor rango dentro del sistema se concentran en la UNAM, en la Ciudad de México y son mayoritariamente hombres. (Álvarez Gallegos, 2012)

A la vez, una crítica recurrente al SNI se vincula con su probable impacto sobre la reducción de la autonomía y capacidades estratégicas de las organizaciones dedicadas -entre otras cosas- a la producción de conocimiento, que se refleja en la escasa capacidad que tienen para incidir no solamente en el trabajo de investigación y su orientación, sino también para equilibrar las restantes actividades académicas centrales en sus objetivos y misión.

De acuerdo con Malo Álvarez (2005) el SNI fue criticado, entre otras cosas, porque en cierto sentido redujo el impacto potencial de las políticas de investigación y la capacidad de dirección de las instituciones. Distintos autores coinciden en la percepción de que el SNI ha debilitado los mecanismos internos de las instituciones para orientar el trabajo de sus investigadores debido al importante peso del estímulo económico externo que otorga, en relación al salario de los investigadores (Reséndiz Núñez, 2005; Ruiz Herrera, 2005). Esto se ve agravado cuando dichas instituciones no se dedican exclusivamente a la investigación (Reséndiz Núñez, 2005).

En particular, está el caso de las Instituciones de Educación Superior que -como se puede observar en el capítulo VI- concentran al 71% de los investigadores categorizados en el SNI. Como señala Reséndiz Núñez (2005), las Instituciones de Educación Superior además de investigar tienen entre sus objetivos y actividades centrales, la enseñanza -de grado y posgrado-, la divulgación del conocimiento, la construcción institucional y, también, en algunas Instituciones de Educación Superior, las actividades de vinculación y/o articulación con el medio o con otros actores no académicos. En general, estas instituciones buscan una interacción fuerte y armónica entre sus actividades de investigación, docencia, divulgación y vinculación. Sin embargo, debido al SNI resulta difícil equilibrar esas actividades. (Reséndiz Núñez, 2005)

Probablemente un mecanismo que está detrás de la explicación de ese fenómeno se vincula con las respuestas estratégicas de los investigadores para adaptarse a las presiones que les impone el SNI. Por ejemplo, Ruiz Herrera (2005) sostiene que el SNI castiga -indirectamente- las actividades de enseñanza porque los investigadores se adaptan a lo que el sistema premia. Es decir, los investigadores saben que si no publican, salen del sistema; pero si no realizan actividades de enseñanza, la sanción no es tal.

Claramente, como se señaló en la sub-sección VII.1.2 a partir de Peña (2005b), si el mérito más importante del SNI fue distinguir entre la producción *primaria* y la *secundaria*, resulta verosímil o esperable que uno de los impactos derivados del sistema de evaluación -dado el contexto más general de la CTI en México- sea justamente el descuido de las actividades académicas consideradas en el sistema como *secundarias* (docencia, divulgación, vinculación, construcción institucional, etc.). Como se señaló anteriormente, estas últimas, como tales, no son reconocidas ni premiadas de igual manera que la *producción primaria* (publicación de artículos científicos en revistas indexadas, preferentemente internacionales y de alto factor de impacto).

De hecho y a casi treinta años de la creación del SNI, distintos autores señalan como un problema persistente del sistema que, si bien en la evaluación que se realiza en el SNI se consideran diversas



actividades –productos- académicas, las mejor calificadas son la producción de artículos publicados en revistas indexadas y la formación de estudiantes de posgrado, principalmente de nivel doctorado. Esto induce a que los investigadores del SNI dediquen poco tiempo a las restantes actividades, entre ellas, a la enseñanza a nivel de licenciatura (Rosales, 2012). Similarmente, Esteinou Madrid (2012) coincide en los desequilibrios sobre las actividades académicas y señala el desbalance a favor de aquellas –actividades y productos- que el SNI reconoce y premia económicamente. Desde la perspectiva de Pescador Osuna (2012), el SNI no ha valorado ni favorecido suficientemente las actividades de docencia, y desde una mirada más general, De Ibarrola (2012) considera que el SNI reconoce y premia la creación de conocimiento nuevo y original y subordina, e incluso margina, las otras actividades académicas orientadas a la enseñanza, divulgación del conocimiento, innovación y vinculación con actores externos.

Mientras la producción de artículos, capítulos de libro o libros gana espacio en el tiempo y dedicación de los académicos, las restantes actividades lo pierden. Si bien no se puede negar la importancia fundamental de la publicación de los resultados de investigación y su incidencia en las audiencias internacionales, es importante generar equilibrios en la evaluación, porque de lo contrario, si lo único que importa es la cantidad de publicaciones de prestigio, la tendencia será a concentrar los esfuerzos en ello y descuidar las restantes actividades académicas, que son centrales, incluso para la investigación.

Para dar cuenta de la importancia de la vinculación como una actividad importante para la investigación, Reséndiz Núñez (2005) refiere al caso de los centros públicos de investigación e institutos nacionales de investigación sectoriales, cuya misión central es producir conocimiento científico y tecnológico orientado a la resolución de los principales problemas nacionales de México. De acuerdo con el autor, en ese contexto los académicos deben identificar dichos problemas, estudiarlos a través de la investigación sistemática, explorar soluciones posibles, desarrollarlas e implementarlas. Los resultados o las soluciones encontradas por los académicos que trabajan en estas organizaciones, requieren investigación científica de calidad y también la interacción con otros para el desarrollo de soluciones tecnológicas efectivas (Reséndiz Núñez, 2005). Esto se podría asociar como el tipo ideal modo 2 de producción de conocimiento.

Sin embargo, este autor señala que en el caso de México, dichas organizaciones han tenido un sesgo a favor de la *investigación básica* –más asociada al tipo ideal modo 1 de producción de conocimiento- en detrimento de sus otras actividades, y que de ese modo, el sistema de CTI del país se torna cada vez más *unimodal*, la diversidad institucional y sus orientaciones son cada vez más reducidas, lo que impacta negativamente en los propósitos integrales de todo el sistema (Reséndiz Núñez, 2005).

De acuerdo con Mateos Gómez (2005), durante sus veinte años de existencia, el SNI ha estado enfocado al estímulo de la *ciencia básica*. La percepción dicotómica existente en el país entre *ciencia básica* y *aplicada* ha frenado el desarrollo de la producción científica en su conjunto, y en México, se ha promovido siempre a la primera (Ruiz Gutiérrez, 2012).

Adicionalmente, Sarukhán Kermez (2005) reconoce que los problemas más importantes que enfrentan la sociedad y la investigación son cada vez más complejos, caracterizados por su no linealidad, y cuyo abordaje requiere de enfoques multidisciplinarios. Sin embargo, señala que ninguno de los sistemas de evaluación a la investigación (de individuos o de proyectos) en México

valora y estimula la investigación multidisciplinar. Adicionalmente, Fortes Besprosvani (2012) sostiene que el SNI está actuando como un agente inhibidor de la movilidad de los investigadores entre la academia, la industria y otros sectores de la economía.

Finalmente, el SNI ha desvalorizado a las ciencias sociales y a las humanidades (Pescador Osuna, 2012), sus criterios de evaluación tienden a favorecer el trabajo teórico sobre el experimental y, con cierta miopía, castigan la formación de *verdaderos* equipos de trabajo<sup>49</sup>, lo que afecta negativamente a la investigación porque el trabajo científico requiere cada vez más de colaboraciones entre grupos e investigadores (Hernández-Calderón, 2005). Ante esto, Asomoza Palacio (2005) coincide en la necesidad de buscar un sistema de evaluación más equilibrado, que no castigue el trabajo colaborativo ni el no disciplinar.

### VII.2.2 El SNI y su flexibilidad epistemológica: la calidad de la producción de conocimiento cuestionada

Nuevamente aquí, el SNI, con el propósito de reconocer y premiar económicamente la labor de investigación, su calidad y el desempeño de los investigadores, ha implementado criterios de evaluación centrados fundamentalmente en indicadores bibliométricos (ver capítulo VI, sub-sección VI.2.3), tales como:

- i) la cantidad de artículos publicados en revistas arbitradas e indexadas en ciertas bases de datos –internacionales- o la cantidad de libros y capítulos de libros publicados en editoriales de prestigio en un determinado período de tiempo, como medidas centrales para aproximar la *productividad* de los investigadores del SNI o de los que aspiran a ingresar al sistema;
- ii) la cantidad de citas que reciben dichas publicaciones, como medida para aproximar la *calidad y trascendencia* de esa producción que, además, mejoran si esas citas fueron referenciadas por investigadores líderes y reconocidos de la comunidad científica internacional; y,
- iii) el factor de impacto de las revistas en las que los investigadores publican sus artículos, como otra medida para aproximar la *calidad* de los mismos y, también, como una medida para aproximar el *prestigio y la calidad* del órgano de publicación (de la revista) –la hipótesis subyacente es que a mayor factor de impacto, mayor es el prestigio y la calidad de la revista-.

De acuerdo con lo planteado en la revisión de la literatura especializada, estos indicadores bibliométricos parecen inducir más a un tipo de producción internacionalizada y acelerada de artículos científicos indexados en WoS o SCOPUS, que a una producción de conocimiento original, relevante y de calidad (ver capítulo IV). Los dos debates sobre el SNI, que se sistematizan en este capítulo, también revelan este efecto constitutivo de los indicadores.

Distintos autores, separados en el tiempo por casi una década, coinciden en que el SNI ha cumplido parcialmente sus objetivos. Mientras que -como se señaló anteriormente- el sistema ha fomentado

---

<sup>49</sup> El autor refiere a que si un investigador joven se asocia con otro de mayor trayectoria no es considerado como un mérito, sino que por el contrario es un demérito, en la medida en que el joven no alcanzó *independencia en su línea de investigación*, lo que se ve peor aún si el investigador de mayor trayectoria fue además su asesor de tesis.

el aumento de la productividad, medida por cantidad de publicaciones, Hernández Calderón (2005) sostiene que es discutible que haya inducido efectivamente a la excelencia académica y a la consolidación de la comunidad científica nacional.

En palabras de Drucker Colín (2005), el SNI –a veinte años de su creación- “(...) *está inhibiendo la calidad y la creatividad de sus investigadores*” (Drucker Colín, 2005: 54). Para este autor, los requisitos del SNI no consideran los “*grandes proyectos estratégicos*”, que en general se pueden caracterizar por ser más ambiciosos en términos de los problemas que abordan y, por lo mismo, son más riesgosos respecto a la obtención de resultados de investigación. A la vez, de acuerdo con el autor, son esos los proyectos que pueden dar lugar a una producción de conocimiento relevante y pertinente, es decir, de gran importancia tanto para el progreso de la ciencia como del país.

Por el contrario, los requisitos del SNI exigen que los investigadores acumulen cierta cantidad de *artículos* dependiendo del nivel de categorización. En ese marco, y debido al prestigio y la importancia relativa del estímulo económico del SNI en el ingreso de los investigadores y para el acceso a otros recursos públicos para la investigación, éstos cumplirán esa exigencia para mantenerse o ascender en el sistema, negando así la posibilidad de desarrollar estos proyectos estratégicos, en los que en general se deben invertir varios años de investigación antes de obtener resultados publicables -y/o publicados-, lo que no resulta en una estrategia viable con el régimen de productividad exigido por el SNI (Drucker Colín, 2005). A su vez, Martínez Palomo (2005) coincide con Drucker Colín (2005) respecto a que el SNI deja poco espacio para la creatividad porque los investigadores están preocupados por cumplir con las exigencias del sistema de publicar porque de eso depende el acceso a recursos para su investigación.

Similarmente, Ramírez Amador (2012) sostiene que el SNI ha fomentado un esquema de *productividad acelerada*, induciendo a los investigadores a la elección de temas de investigación cuyos resultados tienen mayores posibilidades y probabilidades de publicación, y al abandono de proyectos de investigación de largo plazo, más complejos, y colectivos. En definitiva, sostiene, ha inducido al deterioro de la calidad de las investigaciones. De acuerdo con Esteinou Madrid (2012) el modelo de evaluación del SNI estimula más la “*puntitis*” orientada por la cantidad de artículos, libros y capítulos publicados –*productividad*- que por su *calidad*.

Estos desafíos que enfrenta el SNI no son nuevos y, a pesar de los cambios introducidos en el sistema, han persistido a lo largo de los años. Además, estos desafíos no son exclusivos del caso mexicano, los sistemas de evaluación del desempeño de la investigación del Reino Unido, Australia y Holanda (por citar sólo algunos casos internacionales) parecen producir efectos constitutivos similares (ver capítulo IV).

A la vez, y quizás también como reflejo de lo anterior, el esquema de *productividad acelerada* derivado del modelo de evaluación, también parece inducir ciertas respuestas estratégicas de los investigadores para enfrentar las presiones del sistema que, en el contexto mexicano, son duramente cuestionadas desde el punto de vista ético. A continuación se describen. Antes, importa señalar que esas estrategias tampoco son exclusivas de los investigadores mexicanos. Por el contrario, investigadores de diversas partes del mundo –Reino Unido, Australia, Holanda, entre otros- las implementan y están ampliamente documentadas en la literatura internacional especializada (ver capítulo IV).

En particular, algunos autores sostienen que el modelo de evaluación del SNI ha inducido a los investigadores a: i) la *publicación prematura y fragmentada* de sus contribuciones, reduciendo con ello su impacto potencial (Membrillo Hernández, 2005); ii) la producción de *artículos en serie*, en detrimento de proyectos más ambiciosos (Ruiz Herrera, 2005); y, iii) los *“refritos”* para referir a trabajos ya publicados que vuelven a publicarse con cambios relativamente menores (Hernández-Calderón, 2005).

Además, también ha inducido a un tipo de co-autoría por *“conveniencia”* (Hernández Calderón, 2005), *“deshonesta”*, porque no resulta de procesos colectivos de generación de conocimiento (Ruiz Herrera, 2005), o *“arreglada”*, más que sustantiva derivada del trabajo colaborativo y/o multidisciplinar (Membrillo Hernández, 2005).

Finalmente, Pérez Tamayo (2005: 112) sostiene que *“sin el SNI hoy no habría ciencia en México, pero no sé si con el SNI la ciencia en México es mejor”*.

### VII.2.3 El SNI y la escasa contribución del conocimiento científico y tecnológico al desarrollo social y económico de México

En las sub-secciones VII.2.1 y VII.2.2 se plantearon un conjunto de efectos constitutivos del SNI, probablemente derivados de la aplicación de ciertos criterios de evaluación –fundamentalmente bibliométricos-, sobre las características epistemológicas de la producción de conocimiento, su calidad, originalidad, diversidad y orientación predominante. También ha dificultado la orientación de la investigación hacia una producción de conocimiento que puede conceptualmente asociarse más a lo que Gibbons et al. (1994) definió y caracterizó teóricamente como una *nueva forma de producción de conocimiento* y que denominó como *Modo 2*.

Diversos autores coinciden en que el SNI, entre otras cosas, ha dificultado los procesos de investigación en colaboración, tanto entre investigadores –nacionales- como entre éstos y otros actores no académicos, no disciplinares o multi/trans o interdisciplinarios, y orientados por la búsqueda de soluciones para contribuir a la resolución de problemas específicos dependientes del contexto, ya sean sociales y/o productivos. A la vez, tampoco valora la producción de conocimiento basada en el desarrollo de *grandes proyectos estratégicos*, de largo plazo y más riesgosa, pero por lo mismo, probablemente más relevante y pertinente.

Además de esos efectos constitutivos, diversos autores señalan de forma directa la escasa contribución del SNI para orientar la producción de conocimiento nacional hacia la atención de los principales problemas económicos y sociales del país, y con ello, destacan la disociación –una vez más- entre los objetivos que se plantea el sistema y sus resultados.

De hecho, los resultados de la evaluación al SNI realizada por la Auditoría Superior de la Federación (ASF) en 2011 destacan –entre otras cosas- que el sistema no logra cumplir todos sus objetivos, entre ellos los asociados a la contribución del desarrollo científico, tecnológico y la innovación de calidad y en vinculación con diversos sectores, al incremento de la productividad y el mayor bienestar social. Algunos de los datos que provee dicha auditoría son que desde la creación del SNI hasta el año 2009, el 99.1% de los productos realizados por los investigadores del SNI son publicaciones y apenas el 0.9% son patentes, y que no se registraron desarrollos tecnológicos, ni transferencias tecnológicas, ni innovaciones (Citado en Ruiz, 2012: 46 y 47). Vale la pena destacar que la auditoría responde a un análisis de la relación entre resultados y objetivos planteados en el

programa, sin valorar que tan correctamente planteados están dichos objetivos. Esta auditoría condujo a replantearse los objetivos mismos del sistema, y precisar el alcance que puede tener un programa de esta naturaleza.

Tal como sostiene De Ibarrola (2012), los indicadores utilizados por el SNI para calificar la producción de conocimiento nacional son indirectos, reduccionistas y redundan fundamentalmente en la cantidad de artículos publicados en las revistas incluidas en el ISI-WoS y en las citas recibidas. De hecho sólo se reconoce como conocimiento científico aquel certificado por las publicaciones internacionales o las patentes, así el SNI no tiene en cuenta que los indicadores utilizados no reconocen –o tienen dificultades para reconocer- las especificidades de algunos campos de conocimiento, las innovaciones y los cambios. Asimismo, desconoce otras formas de producción de conocimiento y creación. Del mismo modo, el SNI no ha sido capaz de analizar las contribuciones sustantivas de sus miembros (De Ibarrola, 2012).

De ese modo y contrariamente a lo que se indica en sus objetivos, el SNI profundiza la ruptura entre la producción de conocimiento y la resolución de los grandes problemas de la agenda nacional, aumenta la desvinculación entre la academia y los diversos sectores de la producción de bienes y servicios, y de la sociedad en general, lo que conduce a una muy escasa contribución del conocimiento científico y tecnológico -generado y financiado con recursos públicos- al desarrollo nacional (Esteinou Madrid, 2012). Para este autor, el SNI ha instalado una imagen desvirtuada de lo que es ser *investigador nacional*, a saber: es investigador el que publica artículos, el que es citado y el que recibe una compensación económica por ello; y no necesariamente el que produce conocimientos relevantes y pertinentes para contribuir a resolver los desafíos que plantea la ciencia contemporánea y/o la nación (Esteinou Madrid, 2012).

Similarmente, para Fortes Besprosvani (2012), el éxito dentro del SNI depende de un mecanismo que tiende a consolidar el modelo académico tradicional que favorece la investigación básica pero que aísla al investigador de una producción de conocimiento que contribuya al desarrollo social y económico del país. De acuerdo con Pescador Osuna (2012), el SNI no ha estimulado la investigación orientada al desarrollo tecnológico y ha inducido a los investigadores a trabajar en *temas de moda*, y a no atender temas de investigación, emergentes y de urgencia para México.

### VII.3 Dos dimensiones que en interacción sirven para entender las consecuencias de la implementación de un sistema de evaluación fuerte como el SNI

De la revisión del debate nacional sobre el SNI, se identificaron algunas posturas que sirven como piezas para entender, lo que en el contexto de este trabajo y de este capítulo en particular se ha referido genéricamente como las consecuencias –probablemente- derivadas de la implementación de un sistema de evaluación fuerte como el SNI (analizadas y descriptas en las secciones VII.1 y VII.2 de este capítulo). Estas se ven reforzadas por el contexto más amplio del sistema de CTI y la inconsistencia de la política pública nacional.

Una primera postura sostiene que: i) la investigación orientada a la resolución de un problema específico dependiente del contexto, no representa una forma de producción de conocimiento relevante, pertinente, original y legítima, sino simplemente aplicación de conocimiento existente; y a la vez que, ii) la atención de los problemas del desarrollo deberían ser competencia exclusiva del Estado y no de los investigadores.

Esta es una postura relativamente conservadora respecto al rol que le asigna al conocimiento científico y tecnológico y que, además, parece reflejar un desconocimiento importante de los determinantes del crecimiento económico de los países altamente industrializados. Desconecta la excelencia científica de las necesidades de conocimiento científico y tecnológico de un país, región o estado para enfrentar los desafíos que le impone su desarrollo social y económico. Es una postura que descontextualiza la pertinencia de la producción de conocimiento, cuya relevancia depende de si el tema de investigación forma parte de la agenda internacional de producción de conocimiento.

Esta postura, a la vez, reduce la producción de conocimiento de calidad a aquella que se difunde en el ámbito internacional, a través de la publicación de artículos científicos en revistas indexadas en cierta base de datos, cuya trascendencia e impacto dependen exclusivamente de las citas recibidas y de políticas editoriales orientadas a aumentar el factor de impacto de la revista en cuestión. También es la postura que profundiza el encapsulamiento de la producción de conocimiento en el ámbito de la academia, más aún, de la comunidad académica internacional. Es asimismo, la que si bien parece ceder el espacio de la retórica y de lo discursivo<sup>50</sup>, no cede en los hechos. Es decir, en el diseño específico del modelo de evaluación, de los criterios e indicadores que utiliza para valorar la producción de conocimiento de excelencia y, a la vez, para orientar qué ciencia se hace, para qué y para quiénes.

Desde una perspectiva conceptual diferente y contrapuesta a la visión anterior, otras posturas consideran que los escasos resultados del SNI respecto a la contribución del conocimiento científico y tecnológico al desarrollo nacional son resultado de la incongruencia entre la política pública en CTI y otras políticas públicas. Estas posturas, que están más alineadas con el enfoque conceptual adoptado en este trabajo, reconocen que la contribución efectiva del conocimiento científico y tecnológico al incremento de la productividad de un país y al bienestar social de su gente depende, por un lado, de un conjunto integrado, articulado y sistémico de acciones, políticas públicas e instrumentos específicos en diversos ámbitos, que incluyen al de la CTI pero también al económico, productivo, social, entre otros (ver Dutrénit y Sutz, 2014; Bortagaray y Gras, 2014). Por otro lado, depende también de las características específicas del sistema de evaluación del desempeño de los investigadores (Whitley, 2007). En particular, de sus criterios de evaluación y de los indicadores utilizados para valorar la producción de conocimiento realizada, ya que son éstos los que norman la actividad de investigación, guían la acción de individuos e instituciones y con ello, también orientan la dirección y contenido de la producción de conocimiento.

De acuerdo con Balankin (2005) la transición de un país hacia el desarrollo social requiere que sus capacidades de creación, difusión y aplicación del conocimiento científico y tecnológico estén orientadas a atender sus necesidades sociales, y para ello es fundamental –de acuerdo con la experiencia de los países desarrollados- una política pública consistente y grandes inversiones, tanto públicas como privadas. Para este autor, los pobres resultados de la ciencia mexicana para enfrentar los principales problemas del desarrollo económico y social del país<sup>51</sup>, son consecuencia directa de la falta de interés de los gobiernos y de las empresas, que en la práctica -y más allá de lo retórico-

---

<sup>50</sup> Plasmado en los planes de desarrollo nacional, en los documentos sobre política de CTI e incluso en los fundamentos y en los objetivos que sustentan la creación e implementación del SNI.

<sup>51</sup> Entre los que menciona la pobreza, el atraso en la infraestructura, la baja competitividad de las empresas mexicanas, etc..

han demostrado hacia la ciencia. Lo que se expresa –sostiene- en reducidas asignaciones de recursos para la educación, y peores aún, para la investigación y desarrollo tecnológico. (Balankin, 2005).

En una línea argumentativa similar, Pradilla Cobos (2012) destaca que el SNI por sí solo no puede resolver los problemas –persistentes- asociados a la centralización de la investigación, el limitado desarrollo económico nacional o la baja competitividad. Para el autor, encontrar soluciones a esos problemas requiere de políticas públicas específicas y articuladas, donde el SNI es un componente más.

Diversos autores coinciden en que los problemas del SNI no son consecuencia de un desempeño inapropiado de sus investigadores, sino de políticas públicas de CTI y económicas erráticas, y que para superarlos se necesita una visión de Estado, con políticas de CTI articuladas con otras políticas económicas (Ruiz Gutiérrez, 2012).

Como se señaló, el SNI introdujo ciertas normas para el desarrollo y consolidación del trabajo científico. Sin embargo, el conocimiento y el trabajo científico en general se dispersa y fragmenta al no estar integrado, con lineamientos claros, en el marco más general de la política pública, tanto la vinculada a la de ciencia y tecnología como a la política educativa (Valenti, 2012). En sintonía, Rosales (2012) explica la históricamente baja inversión -pública y privada- en ciencia y tecnología por la ausencia de políticas, instrumentos y mecanismos consistentes.

De acuerdo con Valenti (2012) es recién a finales de la década de 1990 que se establece un vínculo entre lo educativo y lo científico, a través de la introducción de la figura de los *cuerpos académicos*. Sin embargo, la ausencia de una política de Estado conduce a incongruencias como la que se pueden observar en relación con la inversión pública orientada a la formación de recursos humanos calificados (*Becas*), ya que no existe contraparte del sector productivo que los emplee (Ruiz, 2012).

A finales de 2011, el Dr. Enrique Villa Rivera -titular del CONACYT- señalaba el escaso rendimiento que tendrán los esfuerzos realizados para ampliar la cantidad de becas para formación en todas las áreas del conocimiento, si -además y conjuntamente- no se crean las oportunidades laborales capaces de absorber a esa gran cantidad de personas altamente calificadas (citado en Rosales, 2012: 96). De ahí –quizás- que muchos de los mexicanos decidan quedarse en el país en el que realizan sus estudios de posgrado (Rosales, 2012).<sup>52</sup>

Por su parte, Ramírez Amador (2012) coincide en los problemas derivados de la escasa inversión en ciencia y tecnología y la falta de políticas de Estado y sus inconsistencias. Adicionalmente, señala un problema que se refuerza debido al SNI. En particular, el vinculado con el hecho de que los estímulos económicos que provee el sistema no estén integrados al salario de los investigadores, lo que limita su jubilación y representa una presión adicional sobre el envejecimiento de los planteles académicos y con ello la no liberación de plazas y la consecuente falta de oportunidades laborales para absorber a los jóvenes altamente calificados. De ese modo también, los jóvenes más talentosos, formados en México o en otros países, financiados con recursos públicos, se ven forzados a buscar oportunidades laborales en el exterior o dedicarse a actividades que nada tienen que ver con la ciencia, reforzando así el problema de la fuga de cerebros, tan común en México y, con ello, perdiendo la inversión

---

<sup>52</sup> El programa de “Cátedras CONACYT para Jóvenes Investigadores”, introducido en el 2014, atiende esta problemática.

pública realizada en su formación, lo que representa una enorme pérdida para el país (Ramírez Amador, 2012).

Similarmente, pero casi una década antes, Pérez Tamayo (2005) y Martínez Assad (2005) alertaron sobre este mismo problema. Desde su perspectiva, el SNI fue el mecanismo que encontró el Estado mexicano para resolver –temporalmente- un problema social y económico que –aunque ya pasó la enorme crisis de inicios de los ochenta- veinte años después, sigue vigente, y es el vinculado al “*de los sueldos miserables de los trabajadores académicos e intelectuales.*” (Pérez Tamayo, 2005: 112)

Como se señaló, en ese marco de crisis, el Estado no podía incrementar salarios, entonces “*inventó*” los estímulos extra-salariales por productividad que, como tales, no forman parte de las prestaciones sociales, por lo que desaparecen cuando los académicos deciden jubilarse. Si bien en su momento, eso significó un “*salvavidas*”, rápidamente se convirtió en parte “*indispensable del presupuesto básico familiar de los investigadores, en lugar de ser un premio adicional a sus esfuerzos*” (Pérez Tamayo, 2005: 112; Martínez Assad, 2005).

Flores Valdés (2012), uno de los fundadores del SNI -entre otros roles en el ámbito de la CTI-, reconoce que el sistema no surge como resultado de una política científica del gobierno mexicano, aunque sí haya sido enmarcado dentro del Programa Nacional de Educación Superior (PRONAES) de la época. Surge más bien de la conjunción casual en tiempo y espacio de personas y voluntades políticas de encontrar una salida económica viable para retener en el país a los mejores investigadores.<sup>53</sup>

De acuerdo con Esteinou Madrid (2012), la primera limitación del SNI está en que fue concebido como un “*proyecto parche*” del Estado mexicano para destrabar el conflicto por aumentos salariales de los académicos y a la vez, evitar la fuga de cerebros; pero nunca como un instrumento orientado hacia el avance de la CTI en el país en el marco de una planificación consistente de la política pública. Esa concepción, desde el origen del SNI, sigue influyendo de forma sustantiva la manera de entender el propósito de la producción científica, su evaluación y su vinculación con el proyecto de desarrollo social y económico del país, de ahí que el sistema no ha contribuido a generar directrices para poner a la producción de conocimiento al servicio de los grandes problemas de la agenda nacional. (Esteinou Madrid, 2012)

Adicionalmente a todas estas incongruencias entre lo científico-tecnológico, educativo y económico-productivo, Valenti (2012) –entre otros- sostiene que en el país ha habido poca reflexión en torno a las formas diferenciales de producción de conocimiento, y que el debate en el ámbito nacional ha sido relativamente estéril, en la medida en que se ha enfocado en una falsa dicotomía –que todavía persiste-: *investigación básica vs investigación aplicada*; sin dejar lugar, otra vez, para diseños e implementaciones de políticas estratégicas para la generación, uso, apropiación, difusión y divulgación del conocimiento.

De todas maneras, aunque el debate haya sido estéril, sí ha tenido un favorito, a saber y parafraseando a Reséndiz Núñez (2005): un sistema de CTI *unimodal*, sesgado a favor de la investigación básica. De acuerdo con este autor, si lo que se busca es que la ciencia repercuta

---

<sup>53</sup> Para una descripción detallada del contexto que da origen al SNI y un análisis sobre el impacto del estímulo económico del SNI en las condiciones materiales de vida de los investigadores y sus limitaciones, ver Pradilla Cobos (2012).



sinérgicamente en el desarrollo social y económico de México, es preciso generar un ambiente que estimule el desarrollo de un verdadero *sistema de CTI*, con diversidad de organizaciones que desempeñen sus actividades de forma articulada y sistémica. (Reséndiz Núñez, 2005). En coincidencia, Cantú (2005) apunta estas incongruencias o inconsistencias de la política pública en CTI y del SNI en particular.

Por todo lo dicho a lo largo de este capítulo, se puede argumentar que el SNI, como instrumento de política pública de CTI, responde más a lo que en la literatura se conoce como el *modelo lineal* de la ciencia y no tanto al *modelo interactivo* (Rothwell y Zegveld, 1985; Rothwell, 1994). Mucho menos parece responder al enfoque de los Sistemas Nacionales de Innovación (Freeman, 1987), ya sea en articulación estrecha con la estructura productiva (Lundvall, 1992), o más recientemente, también preocupados por la contribución del conocimiento científico y tecnológico al desarrollo con inclusión social –*Inclusive Innovation Systems*- (Arocena y Sutz, 2012; Dutrénit y Sutz, 2014).

Si bien el Plan Nacional de Desarrollo de México y la política científica y tecnológica del país (PECITI, 2014-2018), en su retórica parecen alinearse con los enfoques interactivos o, más aún, de los sistemas de innovación, su diseño e implementación continúan anclados en el modelo lineal de segunda mitad del SXX –décadas de los 50 y 60- a través de la asignación de recursos para la CTI con instrumentos de fomento fundamentalmente concentrados del lado de la oferta (ver sub-sección VI.1.1 del capítulo VI) y esperando que sus resultados se derramen por sí solos o a través de mecanismos de mercado –*con manos invisibles*- al resto de la sociedad.

El estudio realizado por Dutrénit et al. (2010) es especialmente relevante para comprender esta inconsistencia de la política pública de México que, entre otras cosas, muestra el importante peso que tienen en la asignación de recursos públicos para la CTI, los programas orientados al fortalecimiento de los recursos humanos (SNI y Becas) y de la ciencia básica.

Adicionalmente, es posible observar otra inconsistencia –o la misma- al interior del SNI. Entre los fundamentos para la creación del SNI se indica la importancia del conocimiento científico y tecnológico para atender los principales problemas del desarrollo social y económico de México; en los objetivos se plantea el fomento a la investigación fundamental y aplicada u orientada a la resolución de problemas; y, entre los criterios de evaluación también se consideran los desarrollos tecnológicos. Sin embargo, algunos resultados y la percepción generalizada de los académicos discutida anteriormente destacan la escasa contribución del conocimiento científico y tecnológico al desarrollo nacional.

Una explicación probable, quizás se vincule con la relativa sobrevaloración de las publicaciones, en tanto “producción primaria”. Otros productos, más ajustados a aquella producción de conocimiento orientada a la búsqueda de soluciones a problemas específicos son –en la mejor hipótesis- menos valorados. Además, importa resaltar que, para demostrar la calidad e impacto de este último tipo de productos, hay que presentar el registro de patente concedida y/u otros registros de propiedad intelectual (ver sub-sección VI.2.3 del capítulo VI).

¿No sería más apropiado, en el caso de México, demostrar cómo es que esa solución o desarrollo tecnológico efectivamente contribuye a la resolución de un problema específico? Si uno de los objetivos de la política pública es poner el conocimiento científico y tecnológico al servicio del

desarrollo social y económico del país, ¿no sería más adecuado diseñar indicadores que aproximen la implementación y uso de las soluciones, más que quienes son sus dueños?

No hay que olvidar que el registro y la obtención de una patente depende de la interacción de un conjunto complejo de factores estructurales, tales como: la regulación, los mercados específicos para la tecnología en cuestión, intereses, costos y beneficios económicos esperados, entre otros (Cimoli y Primi, 2008). Factores, todos ellos, que trascienden largamente la calidad del desempeño del investigador como productor de conocimiento -en este caso- orientado a la resolución de un problema específico. El estudio de Aboites y Díaz (2012) da luz en este sentido. Así, solicitar patentes concedidas como criterio de evaluación para demostrar la calidad de la producción tecnológica desarrollada es, quizás, desmesurado.

Como ya se señaló, tal vez sea por esto que Martínez Palomo (2005) nota -varios años antes que la Auditoría Superior de la Federación- que si bien el SNI ha aumentado la productividad medida por artículos publicados, dicha productividad no ha aumentado, si se mide por el número de patentes concedidas (nacionales y/o internacionales).

Sin embargo, esto no quiere decir que los investigadores mexicanos no forman parte de importantes invenciones, en tecnologías de punta, protegidas con patentes concedidas por la USPTO (*United States Patent and Trademark Office*), es decir, por la oficina de propiedad intelectual y patentes más importante del mundo. De hecho, el estudio de Aboites y Díaz (2012) muestra -entre otras cosas- que en el período comprendido entre 1977 y 2010, la USPTO concedió un total de 2840 patentes en las cuales participó al menos un investigador mexicano. Además, de esas patentes, sólo el 27.3% corresponden a empresas e instituciones de investigación mexicanas, el restante 72.7% son de empresas globales que emplean a investigadores mexicanos para sus desarrollos tecnológicos.

A la vez, el estudio también muestra que, a partir de los años siguientes a la firma del Tratado de Libre Comercio de América del Norte (TLCAN), la brecha de patentamiento -con al menos un investigador mexicano- entre los agentes nacionales y los globales en la USPTO se amplía enormemente, lo que sugiere que las empresas o instituciones de I+D extranjeras (globales) emplean con mayor frecuencia a los investigadores mexicanos para generar nuevos conocimientos y desarrollos tecnológicos con potencial económico, que las empresas y las instituciones de I+D nacionales. (Aboites y Díaz, 2012). Visto desde otro ángulo, ese resultado sugiere que el TLCAN refuerza las ventajas económicas y las capacidades de generación de conocimiento y tecnología con alto valor económico de las empresas o instituciones de I+D globales.

Aboites y Díaz (2012) indican que más de la mitad de estos investigadores fueron miembros del SNI en algún momento del período comprendido entre 1985 y 2010. Sin embargo, los autores señalan que los mecanismos del SNI no fueron suficientes para retener a estos investigadores en México. Esos investigadores trabajan en empresas e institutos de I+D extranjeros, en particular de Estados Unidos, y sus argumentos más comunes para permanecer en ese país son que: i) los salarios allí son más altos; y, ii) en México no existen las condiciones de infraestructura, instrumentos, materiales, masa crítica, *ambiente* para el desarrollo de esas actividades de investigación científica y desarrollo tecnológico (Aboites y Díaz, 2012).

Esteinou Madrid (2012) -entre otros- señala que el desarrollo del país requiere de la contribución del conocimiento científico y tecnológico. Sin embargo, y desde su perspectiva, el SNI no ha sido

capaz de canalizar –más allá de la ausencia de políticas de Estado planificadas e integradas a los objetivos de desarrollo nacional- la producción de conocimiento realizada por sus miembros en esa dirección. Por el contrario, este autor sostiene que el acervo de nuevos conocimientos “*se deja espontáneamente a expensas de la dinámica de aprovechamiento de los intereses del mercado (...)*”, así la inversión pública en CTI termina beneficiando más a las grandes empresas transnacionales que a empresas de servicio público o comunitarias del país (Esteinou Madrid, 2012: 152-153).

Los desequilibrios del SNI, identificados por De Ibarrola (2012), sintetizan algunas de las tensiones planteadas a lo largo de este capítulo, a saber: i) el rol asignado al conocimiento científico y tecnológico, y sin embargo, una muy débil contribución del mismo al desarrollo del país y a la solución de los problemas nacionales; ii) una validación del conocimiento a través de la cuantificación de los acervos ya legitimados -preferentemente internacionalmente- en vez de una validación del conocimiento por la creación y su originalidad; iii) a favor de una producción de conocimiento internacionalizada en detrimento del reconocimiento de aquella producción preocupada por la especificidad nacional, en particular en las Humanidades y en las Ciencias Sociales; y, iv) que reconoce investigadores *nacionales* pero desconoce y deslegitima otras formas de producción y transferencia de conocimiento.

#### VII.4 El SNI y la dirección del cambio: algunas propuestas para atender sus desafíos

Sin desconocer que el SNI ha hecho una contribución al sistema científico de México, ya sea por el mantenimiento y continuidad de las actividades académicas, como por lo que significa en términos de complemento en los ingresos económicos de los investigadores. También se han realizado esfuerzos y cambios desde su creación para su mejora continua. No obstante, diversos autores coinciden en la necesidad de transformar al sistema de tal manera que permita afrontar los desafíos que impone el avance del conocimiento y el desarrollo social y económico del país. Se facilitaría, de ese modo, el tránsito hacia un sistema donde la CTI se convierta en el motor del desarrollo social y económico de México.

Como se señaló en la sección anterior, poner a la CTI al servicio del desarrollo social y económico de México, requiere de un conjunto integrado de políticas públicas e instrumentos consistentes, donde el SNI puede desempeñar un papel muy importante.

En esa dirección, desde la perspectiva más general de la *planeación estructural y la vinculación con la sociedad*, Esteinou Madrid (2012) propone diseñar una nueva política de Estado integral de CTI cuyos ejes centrales sean, por un lado, el impulso plural y equitativo de la CTI para enfrentar los desafíos que impone la *sociedad del conocimiento*, y por otro lado, que se vincule estratégicamente la producción de conocimiento con la resolución de los grandes problemas de la agenda nacional del país. A la vez, y con esa idea en mente, realiza un conjunto de propuestas para el *ámbito de la evaluación a los investigadores* que coinciden en mucho con las realizadas por otros autores.

A continuación se presentan el conjunto de transformaciones al SNI propuestas por los autores, que surgen de esta revisión del debate sobre el sistema.

En 2005, cuando se cumplían veinte años de la creación del SNI, algunos autores ya consideraban necesario un cambio en su modelo de evaluación; cerca de una década después esa necesidad siguió vigente. En su momento, Membrillo Hernández (2005) planteaba la necesidad de revisar el modelo

de evaluación del SNI para estimular más la calidad que la cantidad de artículos. Para Pescador Osuna (2012) el sistema debe dejar de valorar más la cantidad que la calidad.

De acuerdo con Ramírez Amador (2012), el desafío para la ciencia mexicana está también en la forma en que se juzga el desempeño de los investigadores, tanto en lo referido al ingreso al SNI, como a la permanencia y ascenso en el sistema. Las comisiones dictaminadoras cuantifican el desempeño de los investigadores y califican mejor a aquellos que publican más artículos en revistas con mayor factor de impacto, que reciben mayor cantidad de citas o índice h. Para la autora, los criterios de cuantificación prevalecen sin considerar criterios cualitativos que en ocasiones no pueden ser juzgados por pares que pertenecen a diferentes campos científicos.

Mientras que el Dr. Lalette San Román –en ese momento coordinador general del FCCyT- considera que el análisis bibliométrico provee instrumentos pertinentes para la evaluación transversal de la producción de conocimiento realizada por los investigadores del SNI; el Dr. de la Peña destaca el peligro de utilizar dichos indicadores en soledad y enfatiza que *“los números no son superiores a los juicios”* y que se requieren evaluaciones cualitativas más que numéricas (citados en Ramírez Amador, 2012: 121-122).

En coincidencia, Porter (2012) sostiene que es preciso que el sistema supere su tendencia a la evaluación cuantitativa certificadora y de paso a una valoración cualitativa de la producción de conocimiento<sup>54</sup>, porque de lo contrario, *“el mérito seguirá siendo proclive a la mentira, la simulación y las trampas (...).”* (Porter, 2012:89).

Similarmente, Drucker Colín (2005b) indica que es preciso cambiar la estrategia y los criterios de evaluación del SNI, porque la producción de conocimiento mundial se hace en grandes grupos de investigación, en áreas estratégicas importantes, donde México tiene mucho para aportar en la resolución de esos problemas, y ser exitoso. Según este autor, la estrategia tiene que orientarse a planear, buscar áreas de oportunidad para los investigadores del país, desarrollar proyectos estratégicos *“de gran envergadura”*, ambiciosos, más riesgosos y que pueden no generar -en el corto plazo- resultados y publicaciones inmediatas. Para poder hacer eso hay que cambiar los criterios de evaluación del SNI, porque si se sigue obligando a los *“jóvenes y a los no tan jóvenes a seguir siendo simplemente emisores de uno o dos papers por año para poder mantenerse dentro del SNI, no vamos a ir a ningún lado en los próximos años; de hecho en los próximos veinte [años] vamos a perder nuestra capacidad, porque en el mundo la investigación hoy día se hace de una manera muy diferente a la que la estamos haciendo”* (Drucker Colín, 2005b: 85).

En sintonía con Drucker Colín, si el objetivo es contribuir decididamente a la resolución de los grandes problemas nacionales, Terrones Maldonado (2005) sostiene que el SNI tiene que cambiar en una dirección más dinámica, que considere el trabajo multidisciplinar. En la misma dirección, Membrillo Hernández (2005) argumenta que debe favorecer el desarrollo de proyectos de

---

<sup>54</sup> Es decir, que el SNI otorgue menos importancia a la cantidad de artículos publicados y valore más su contribución intelectual e impacto, considerando el contexto institucional, disciplinar y la orientación de la investigación que produce ese conocimiento; además, que las revisiones de los expedientes sean menos rutinarias y se orienten más a una valoración de las trayectorias académicas contextualizadas y de largo plazo de los investigadores. (Porter, 2012)

investigación de *alta envergadura y de riesgo elevado*, porque es eso lo que –eventualmente– permitirá la realización de contribuciones científicamente relevantes, significativas y de alto impacto (Membrillo Hernández, 2005).

Para Álvarez Gallegos (2012) el SNI debe valorar especialmente los trabajos de calidad que contribuyen a la resolución de los problemas relevantes y estratégicos nacionales, cuya resolución tiene alto impacto, propiciando el trabajo en grupo, multidisciplinario, interinstitucional e intergeneracional. Además, considera necesario apostar al largo plazo y fortalecer la vinculación. González Valenzuela (2005) ya señalaba la necesidad de repensar los períodos temporales de la producción de conocimiento científica y humanística, ajustándolos a las diversas modalidades de investigación. De acuerdo con Porter (2012), es necesario que la evaluación transite hacia una valoración de la producción de conocimiento que reconozca que se trata de un proceso esencialmente colectivo.

Adicionalmente, algunos autores han planteado la necesidad de que en el SNI se realice una evaluación equilibrada del desempeño de los investigadores, atendiendo a las especificidades de las diversas áreas de conocimiento (Esteinou Madrid, 2012). Para Porter (2012) es necesario revisar la división histórica de las áreas de conocimiento y sus campos de actuación, superando todo tipo de discriminación (*ciencias duras vs ciencias blandas; investigación básica vs investigación aplicada*), incluyéndolas a todas, respetando su orientación y reconociendo la importancia y calidad de todas ellas. De lo contrario –sostiene– continuará prevaleciendo un tipo de conocimiento que no refleja la diversidad ni la pluralidad de la producción de conocimiento nacional, sino sólo una parte arbitraria de la misma.

En una línea argumentativa similar, Ramírez Amador (2012) señala que si bien la publicación de artículos científicos es ineludible, no es razonable que no se tomen en cuenta las características específicas que diferencian a los diversos campos científicos y sus modalidades heterogéneas de producción de conocimiento.

En ese sentido y desde la perspectiva de Pescador Osuna (2012) es necesario -sin bajar la calidad- ampliar la visión del trabajo académico y de investigación que tienen las comisiones dictaminadoras, para que incluya cierto tipo de investigación que requiere el país para su desarrollo social y económico, y que reconozca el impacto de la investigación, no solamente a través de indicadores cuantitativos sino también a través de la valoración de la investigación en sí misma, por sus propios méritos o naturaleza. Esteinou Madrid (2012) coincide cuando señala que la evaluación debería apreciar las contribuciones científicas y tecnológicas de los investigadores por sus méritos intrínsecos y valorar su impacto según la asimilación de los resultados de investigación por los diversos sectores – el académico, el de la producción, el de la política pública, el de la sociedad civil, etc.-. En ese marco, las citas a los trabajos representan un indicador complementario para aproximar la trascendencia de la producción de conocimiento pero nunca el único para evaluar el impacto de la producción científica y tecnológica.

También, mientras que Pescador Osuna (2012) propone ampliar los medios válidos de difusión de la producción de conocimiento; Esteinou Madrid (2012) señala que el SNI debe valorar todos los reconocimientos, distinciones, premios y citas, ya sean nacionales o internacionales, y reconocer todas las modalidades de formación de grupos de investigación.

Desde el 2005 se planteaba la necesidad de que el SNI valore adecuadamente y estimule el desarrollo tecnológico y la innovación (Mateos Gómez, 2005), más recientemente Álvarez Gallegos (2012) sugiere asignar, mediante los criterios de evaluación, mayor peso a los desarrollos tecnológicos ponderando diversos elementos para valorar su calidad e impacto. Además y atendiendo a las especificidades de las actividades asociadas al desarrollo tecnológico, propuso la creación de una comisión dictaminadora adicional para evaluar dichas actividades y los productos derivados de la actividad *tecnológica y de innovación*.

Fue recién en 2017 que se crea la Comisión Dictaminadora Transversal de Tecnología, pero a diferencia de las restantes, no selecciona sino que actúa como asesora de las siete comisiones dictaminadoras. Es decir, no tiene capacidad de selección. La decisión sobre los seleccionados para –eventualmente- ingresar, permanecer o ascender en el SNI sigue en la órbita de las restantes siete comisiones dictaminadoras.

Diversos autores marcaron la necesidad de que el SNI valore de forma sustantiva la trayectoria académica de los investigadores considerando todas las actividades que desempeñan (Esteinou Madrid, 2012).

Desde 2005 se plantea la necesidad de una mayor valoración –o valoración real- de las actividades docentes de enseñanza (Membrillo Hernández, 2005), y buscar un mayor equilibrio respecto a las de investigación, en particular dada la sub-valoración existente de la docencia en licenciatura (González Valenzuela, 2005). En 2012, Álvarez Gallegos (2012) se planteaba la necesidad de ponderar adecuadamente las actividades educativas, de docencia y formación de recursos humanos, valorando: el dictado de cursos (grado y posgrado), la tutoría y asesorías de posgrados en línea, la elaboración de materiales, las direcciones de tesis incluyendo las de licenciatura, los diseños y las reestructuraciones de los programas de posgrados, la publicación de libros de texto, entre otros productos. En términos más generales, Pescador Osuna (2012) proponía equilibrar el peso relativo de la actividad científica a favor de la actividad de enseñanza y ampliar los criterios para valorar la formación de recursos humanos.

Debido a la importancia de las actividades de divulgación del conocimiento científico dentro de la actividad científica y por su impacto en el desarrollo de la ciencia en México y en el nivel cultural de su gente, diversos autores destacaron la necesidad de que la evaluación reconozca esta actividad (Ruiz Gutiérrez, 2012; Membrillo Hernández, 2005; González Valenzuela, 2005).

Finalmente, algunos autores también señalaron la necesidad de facilitar el ingreso al SNI de los jóvenes investigadores (Membrillo Hernández, 2005) y también su permanencia (Álvarez Gallegos, 2012). Para facilitar el ingreso de los más jóvenes, Álvarez Gallegos (2012) propone generar programas de retiro-jubilatorios decorosos, y para ayudar a su permanencia, sugiere la extensión de los períodos entre evaluaciones para los *candidatos*. Sin embargo, en el último reglamento del SNI de 2017, otros criterios predominaron y los períodos de vigencia para los *candidatos* se redujeron (ver sub-sección VI.2.2 del capítulo VI).

## VII.5 Conclusión

El debate nacional sobre el SNI refleja la percepción que tienen los actores sobre sus efectos performativos y constitutivos sobre el contenido, dirección y características epistemológicas de la producción de conocimiento.

El debate nacional sobre el SNI es consistente con el debate internacional respecto a la inconformidad que genera -entre algunos actores- la amplia difusión de indicadores bibliométricos como criterios centrales de evaluación del SNI y también respecto a las propuestas que realizan para mejorar las prácticas de evaluación del SNI. Ambos debates, el nacional e internacional, convergen en la necesidad de propender a una evaluación más cualitativa, que valore todos los productos derivados de la investigación y su impacto, y que trascienda la contabilización de artículos publicados en revistas internacionales indexadas en bases de datos internacionales, sus citas y los factores de impacto de dichas revistas como medidas para aproximar la calidad e impacto de la producción de conocimiento. Es decir, que reconozca la diversidad y las diversas prácticas de producción de conocimiento de las distintas áreas de conocimiento, y valore los resultados de la investigación de forma integral y atendiendo a los contextos en que se produce conocimiento.

El debate nacional sobre el SNI también es consistente con los hallazgos de investigaciones pasadas respecto a los sesgos que introduce la evaluación cuantitativa a favor de un estilo único de producción de conocimiento internacionalizado y en contra de la producción de conocimiento contextualizada, orientada por la búsqueda de soluciones a problemas específicos dependientes del contexto, ya sean sociales, productivos o de política pública. También es consistente respecto a los sesgos que introduce el SNI, en contra de la investigación de largo plazo, inter/multi/transdisciplinar y en colaboración o interacción con otros, en particular con actores no académicos. A la vez, y de acuerdo con la percepción de algunos de los actores que participan del debate nacional y en sintonía con los hallazgos de investigaciones pasadas, el SNI y la presión que ejerce sobre los investigadores, debido a la cantidad de publicaciones anuales que exige, induce a una cierta reducción de la calidad, relevancia y originalidad de la producción de conocimiento nacional.

Así, el marco teórico conceptual (capítulo V) de esta tesis también queda justificado y fundamentado de acuerdo con el debate nacional sobre el SNI, ya que es consistente con lo expuesto a lo largo de las sub-secciones VII.2.1, VII.2.2 y VII.2.3, y la sección VII.4 de este capítulo. De aquí también que sea razonable esperar que la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI varíe significativamente dependiendo de sus formas diferenciales de producción de conocimiento que predominantemente orientan la investigación que realizan (*propensión modo 1* y *propensión modo 2*). En particular se espera, que los investigadores con una mayor *propensión modo 2 de producción de conocimiento* perciban una *flexibilidad epistemológica* menor que los investigadores con una mayor *propensión modo 1 de producción de conocimiento*. A la vez, también es razonable esperar una relación positiva y significativa entre *flexibilidad epistemológica* percibida del SNI y la *satisfacción con su proceso de evaluación*; a mayor *flexibilidad*, mayor también es la *satisfacción con el proceso de evaluación*. Simultáneamente, es razonable esperar que la *satisfacción con su proceso de evaluación* varíe significativamente y en direcciones opuestas dependiendo también de las formas diferenciales de producción de conocimiento de los investigadores categorizados en el SNI. Se espera que a mayor *propensión modo 2 de producción de conocimiento*, menor sea la *satisfacción con el proceso de evaluación*; y que a mayor *propensión modo 1 de producción de conocimiento*, mayor también sea la *satisfacción con el proceso de evaluación*, ya que el sistema reconoce y premia mejor los resultados y productos derivados de la investigación fundamental o básica. Finalmente, también resulta razonable esperar que cuanto mayor sea la *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI, menor sea su interés en reformarlo hacia una

evaluación más cualitativa y que valore mejor la investigación colaborativa, inter/multi/transdisciplinar y aplicada u orientada a la resolución de problemas dependientes del contexto (*dirección de la reforma del sistema de evaluación*). En el capítulo X se ponen a prueba estas hipótesis y se presentan los resultados obtenidos.

La relevancia de las hipótesis que se plantean se vincula al hecho de que si bien muchos autores -a lo largo de este capítulo- han señalado directa o indirectamente estas relaciones, lo han hecho desde sus experiencias, perspectivas o análisis puntuales. En ese sentido esta tesis pone a prueba una teoría de medición para garantizar la validez de constructo de las variables latentes involucradas -constructos o dimensiones conceptuales- y, además, una teoría estructural que concentra todas las relaciones hipotetizadas entre esas variables, en un marco teórico conceptual (modelo) plausible de ser refutado empíricamente -o por el contrario, no refutado-, con confiabilidad, validez y generalizable a toda la población del SNI. Es decir, en esta tesis se propone una teoría sobre las relaciones entre la evaluación del desempeño de la investigación y las formas diferenciales de producción de conocimiento de Gibbons et al. (1994), basada en la evidencia empírica del SNI. Hasta donde se sabe y de acuerdo a la revisión de la literatura realizada en los capítulos II, III y IV ningún estudio (ni del ámbito nacional ni del internacional) ha explorado estas relaciones.

En el capítulo X se retoman las definiciones conceptuales de los constructos, se operacionalizan todas las variables latentes -los constructos-, se establecen sus relaciones, y se presentan los resultados obtenidos. En el siguiente capítulo (VIII) se desarrollan todos los métodos para asegurar la validez, confiabilidad y generalización de los resultados de esta investigación.

Finalmente, tal como se planteó en la sección VII.3, si bien el contexto más general del sistema de CTI de México es útil para entender el reforzamiento de los efectos del SNI que perciben los actores que participan del debate nacional sobre ese sistema de evaluación, sus logros y desafíos; ese contexto y los efectos del SNI se ven reforzados por la interacción de dos dimensiones, que sintéticamente se pueden resumir en: i) la inconsistencia en el diseño y linealidad de la política pública en CTI de México y su desarticulación de otras políticas públicas, y ii) la inconsistencia entre los fundamentos y objetivos del SNI y sus criterios de evaluación. Así, mientras que la retórica de la política pública en CTI de México establece entre sus objetivos el papel singular de la CTI como motor de su desarrollo social y económico, sus instrumentos, en particular el SNI, dados los criterios de evaluación que emplea, parece contradecirlos.



## CAPITULO VIII: Diseño y estrategia de investigación: métodos, técnicas y fuente de información

El objetivo de este capítulo es múltiple; se trata no solamente de describir el conjunto de métodos y procedimientos utilizados en este trabajo, sino fundamentar las razones de su elección. A lo largo de este capítulo el lector encontrará respuestas, que se espera sean relevantes, para responder a las preguntas sobre qué, cómo y por qué se utiliza el enfoque seleccionado.

En primer lugar, resulta imprescindible retomar aquí los objetivos -general y específicos-, la pregunta central de investigación y las hipótesis que orientan este trabajo. A la vez, es preciso indicar también qué compromisos se asumen -en el marco de este trabajo- respecto a los criterios de la inferencia científica.

Como se señaló, el objetivo general de investigación es explorar la naturaleza de las relaciones inmersas entre un sistema de evaluación fuerte del desempeño individual de investigación y los dos modos diferenciales de producción de conocimiento -Modo 1 y Modo 2- caracterizados por Gibbons et al. (1994) en el contexto del SNI de México.

Los objetivos específicos de investigación son explorar la naturaleza de las relaciones (significancia estadística, magnitud, dirección y su implicancia conceptual) entre las cinco dimensiones involucradas en el análisis, a saber:

- i) *la flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI y su *propensión modo 1* de producción de conocimiento;
- ii) *la flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI y su *propensión modo 2* de producción de conocimiento;
- iii) *la satisfacción con el proceso de evaluación* del SNI que tienen los investigadores categorizados en ese sistema y su *propensión modo 1* de producción de conocimiento;
- iv) *la satisfacción con el proceso de evaluación* del SNI que tienen los investigadores categorizados en ese sistema y su *propensión modo 2* de producción de conocimiento;
- v) *la satisfacción con el proceso de evaluación* del SNI que tienen los investigadores categorizados en ese sistema y *la flexibilidad epistemológica del sistema de evaluación* que perciben;
- vi) *la dirección de la reforma del sistema de evaluación* del SNI según el interés de los investigadores categorizados en ese sistema y *la satisfacción con el proceso de evaluación* del SNI que dichos investigadores tienen.

La principal pregunta de esta investigación es sobre la naturaleza de dichas relaciones, a saber: ¿qué relaciones hay entre un sistema de evaluación fuerte del desempeño individual de investigación como el SNI y los Modos 1 y 2 de producción de conocimiento que predominantemente orientan la investigación que realizan los investigadores categorizados en ese sistema de evaluación y cuál es la naturaleza de esas relaciones? Si dichas relaciones existen: ¿cuán importantes son, qué magnitud tienen?, ¿cuál es la dirección de tales relaciones, se trata de relaciones positivas o negativas?

De acuerdo con la teoría que vincula a la evaluación del desempeño de la investigación y la producción de conocimiento, las contribuciones de Gibbons et al. (1994) y los hallazgos de investigaciones pasadas respecto a las consecuencias de la evaluación sobre la dirección, contenido

y características epistemológicas de la producción de conocimiento documentadas en la literatura, las hipótesis que orientan esta investigación son:

- ✓ H1: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *flexibilidad epistemológica del sistema de evaluación*.
- ✓ H2: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.
- ✓ H3: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *flexibilidad epistemológica del sistema de evaluación*.
- ✓ H4: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *satisfacción con el proceso de evaluación*.
- ✓ H5: la *flexibilidad epistemológica del sistema de evaluación* se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.
- ✓ H6: la *satisfacción con el proceso de evaluación* se relaciona significativamente y negativamente con la *dirección de la reforma del sistema de evaluación*.

Para alcanzar los objetivos de investigación planteados, contrastar las hipótesis orientadoras con la realidad y en definitiva, sugerir una explicación teórica válida, confiable y representativa como respuesta a la pregunta sobre la naturaleza de las relaciones entre las diversas dimensiones de la evaluación de la investigación y la producción de conocimiento, se optó por el modelado con ecuaciones estructurales (SEM por sus siglas en inglés) (Kline, 2011; Bollen, 1989).

En ese sentido entonces, lo que se hizo fue construir y validar un modelo teórico con ecuaciones estructurales para explicar las relaciones entre diversas dimensiones de la evaluación de la investigación y, de éstas, con dos formas diferenciales de producción de conocimiento.

En principio, la razón central que fundamenta la elección del modelado estructural se asocia a dos temas relacionados. El primero se vincula con el tipo de fuente de información disponible para este trabajo; y el segundo se asocia con las características específicas del problema de investigación planteado.

De acuerdo con Kish (2004) existen tres grandes áreas del diseño estadístico de la investigación, a saber: i) diseños experimentales; ii) encuestas por muestreo; y iii) estudios observacionales. Según este autor, la elección de alguna de esas tres áreas del diseño depende de lo deseable y de lo posible; y a la vez, implica asumir ciertos compromisos con alguno de los tres principios de la inferencia científica: i) aleatorización; ii) representatividad; y, iii) realismo. (Kish, 2004)

Debido, justamente, a lo posible<sup>55</sup> -gracias al acceso a la encuesta "*Percepciones sobre la evaluación académica*" (Bensusán et al., 2014; FCCyT, 2014), esta investigación se enmarca en el área del diseño de las *encuestas por muestreo* y el compromiso que se asume, en términos de la inferencia, es con el principio de *representatividad*. En la sección VIII.3.3 se detallan las características específicas de la encuesta y la muestra. A la vez, tal como destaca Kish (2004), si bien no se pueden asumir

---

<sup>55</sup> Cabe señalar que este trabajo, si bien ha mantenido las preocupaciones originales de investigación, ha modificado su abordaje en términos del diseño estadístico y, consistentemente, ha cambiado los objetivos, preguntas e hipótesis de investigación en tres oportunidades, siendo –finalmente- *la tercera la vencida*. Las razones de dichas modificaciones fueron: primero la falta de acceso a ciertas fuentes de información existentes en CONACYT, y luego, la imposibilidad material de construir los datos necesarios.

compromisos con la aleatorización -propio de los diseños experimentales-, ni con el realismo – propio de los estudios observacionales-, se implementan un conjunto de estrategias para: i) controlar los efectos de ciertas variables perturbadoras, y ii) asegurar validez de constructo.

El problema de investigación que se aborda se caracteriza por ser complejo y multidimensional. Es decir, propone una explicación teórica de la relación entre la evaluación y la producción de conocimiento a partir de un conjunto de interrelaciones de dependencia simultáneas entre cinco dimensiones conceptuales (flexibilidad epistemológica del sistema de evaluación, satisfacción con el proceso de evaluación, dirección de la reforma del sistema de evaluación, propensión modo 1 y propensión modo 2 de producción de conocimiento).

La explicación teórica que se propone se refleja a través de las hipótesis de investigación planteadas. En ellas se plantean las múltiples interrelaciones de dependencia entre las cinco dimensiones conceptuales de análisis. A la vez, las cinco dimensiones conceptuales de análisis, que encarnan los objetivos e hipótesis de investigación, representan constructos inobservables directamente por su complejidad. El desafío entonces es múltiple, se trata de seleccionar un procedimiento que permita contrastar las hipótesis simultáneamente, y además, que permita la medición de los constructos latentes -inobservables directamente- involucrados en dichas hipótesis.

Los *constructos latentes* refieren a conceptos abstractos, que se definen conceptualmente pero que pueden ser medidos indirectamente a través de variables manifiestas (que contienen un error de medición). Por lo tanto, los constructos o *variables latentes* deben ser aproximados/medidas a través de variables manifiestas (*indicadores, ítems*), y como toda aproximación, contiene un *error de medida* que es considerado en el análisis (Bollen, 1989; Hair et al., 2014)

De acuerdo con la literatura, el modelado estructural es el único procedimiento estadístico capaz de contrastar todas las hipótesis simultáneamente, incluir *variables latentes* (las relaciones que se predicen en las hipótesis son relaciones entre las variables latentes), y no solamente considerar el error teórico de la predicción, sino también el error de medición durante el proceso de estimación.

Adicionalmente, el modelado con ecuaciones estructurales es un procedimiento que incorpora la estructura de covarianzas y procedimientos de máxima verosimilitud. Los modelos de ecuaciones estructurales también se denominan *análisis de la estructura de la covarianza o modelado de la estructura de la covarianza* entre otros nombres, y de acuerdo con Kline (2011), esos términos son utilizados para clasificar el conjunto de técnicas o procedimientos estadísticos involucrados en esta forma de modelado.

Al igual que cualquier técnica estadística de análisis multivariado, el objetivo principal de los modelos de ecuaciones estructurales es ampliar la capacidad explicativa y la eficiencia estadística (Hair et al., 2014).

De acuerdo con estos autores el modelado de ecuaciones estructurales tiene tres características que lo diferencian de otras técnicas de análisis multivariado, a saber: i) permite la estimación simultánea de múltiples relaciones de dependencia; ii) tiene la habilidad de representar conceptos inobservables en esas relaciones de dependencia y considerar su error de medición en el proceso de estimación; y, iii) provee la capacidad de definir un modelo teórico para explicar y contrastar la totalidad de las relaciones teóricas en él involucradas. Para ello, el modelado con ecuaciones

estructurales puede dividirse en dos partes, la parte de la medición –el modelo de medida- y la parte estructural –el modelo estructural- (Kline, 2011; Hair et al., 2014).

Es por lo expuesto hasta aquí que en esta investigación se seleccionó el modelado con ecuaciones estructurales como procedimiento estadístico y analítico para explicar la relación entre las diversas dimensiones de la evaluación de la investigación, y de éstas, con las formas diferenciales que predominantemente orientan la producción de conocimiento realizada por los investigadores.

Una vez establecido qué se hace y por qué, a continuación, en la siguiente sección, a través de la descripción de la estrategia de investigación y sus etapas, se da cuenta del cómo se hace.

### VIII.1 Estrategia de Investigación

Con el propósito de describir y fundamentar el conjunto de métodos y procedimientos que se utilizan en este trabajo, a continuación se exponen las etapas de esta investigación y el conjunto de técnicas multivariadas utilizadas en cada una de ellas. La intención es facilitarle al lector tanto una mirada global como la racionalidad que fundamenta su aplicación.

La estrategia de investigación seguida en este trabajo se divide en tres grandes etapas complementarias y consecutivas:

1. La etapa descriptiva y exploratoria.
2. La etapa de la medición y validación.
3. La etapa de la explicación de la relación entre la evaluación de la investigación y la producción de conocimiento.

En cada una de estas etapas se emplean distintas técnicas de análisis multivariado de acuerdo con el objetivo central de investigación de cada una. Así, en la primera etapa se realiza un Análisis Factorial Exploratorio (AFE) (Hair et al., 2014), en la segunda un Análisis Factorial Confirmatorio (AFC) (Hair et al., 2014; Brown, 2015), y en la tercera el modelo teórico estructural (Kline, 2011; Bollen, 1989; Brown, 2015; Hair et al., 2014). A continuación se presentan brevemente.

El objetivo central de la primera etapa es describir y explorar las asociaciones entre las variables observadas seleccionadas para el análisis, y entre éstas y las dimensiones conceptuales que, eventualmente, pueden explicarlas. Para alcanzar dicho objetivo, se realiza un AFE. Esta técnica de análisis multivariado permite explorar empíricamente la estructura de correlación entre todas las variables observadas de interés para identificar las dimensiones conceptuales latentes –factores- capaces de explicarla.

De acuerdo con los objetivos de investigación y la revisión de la literatura, con el AFE se espera identificar una estructura multifactorial de cinco factores, cuya interpretación pueda asimilarse a las dimensiones conceptuales definidas en este trabajo como *propensión modo 1*, *propensión modo 2*, *satisfacción con el proceso de evaluación*, *flexibilidad epistemológica del sistema de evaluación* y *dirección de la reforma del sistema de evaluación*.

La racionalidad que justifica esta primera etapa es establecer la base empírica para la construcción de una teoría de medida válida y confiable. Si no hubiera una estructura de correlación significativa entre las variables observadas de interés, el objetivo central de investigación de esta tesis carecería de sentido. Si la solución del AFE coincide total o parcialmente con los resultados esperados, se

procede con la segunda etapa para confirmarlos, o bien, para poner a prueba y validar una teoría de medida, representada por un modelo para medir la relación entre la evaluación de la investigación y la producción de conocimiento.

Teniendo presente, por un lado los resultados empíricos obtenidos en la primera etapa, y por otro, la revisión de la literatura, se establecen -en la segunda etapa de investigación- un conjunto de hipótesis sobre las relaciones entre las variables observadas de medida seleccionadas para el análisis y las dimensiones conceptuales que, eventualmente, las explican. En este sentido, el objetivo de esta segunda etapa es poner a prueba y validar la teoría de medición que sustenta dichas hipótesis. Dicha teoría se representa a través de un modelo de medida para la relación entre la evaluación del desempeño individual de investigación de los investigadores y la producción de conocimiento.

Para alcanzar dicho objetivo, se implementa un AFC. Esta técnica de análisis multivariado permite contrastar con la realidad y validar el modelo de medida multifactorial de cinco factores construido teóricamente y con base empírica, y así medir la relación entre la evaluación de la investigación y las formas diferenciales de producción de conocimiento. Contrastar con la realidad y validar el modelo de medida implica analizar, por un lado, su bondad de ajuste a los datos, y por otro, evaluar su validez de constructo. Si bien no es posible un compromiso con el realismo como el que se tendría si se tratara de un estudio observacional, el análisis de la validez de los constructos, es decir, la preocupación por la calidad de la medición de los conceptos, y con ello, la evaluación sobre si efectivamente se está midiendo lo que se busca medir, intenta mitigar la falta de realismo de los diseños de las encuestas por muestreo.

El fundamento racional detrás de esta segunda etapa es que no se pueden obtener resultados robustos ni extraer conclusiones rigurosas, si se parte de medidas pobres -poco confiables y/o sin validez-. Por lo tanto, de este análisis se espera alcanzar un modelo de medición válido y confiable. Si este resultado se logra, se procede con la tercera etapa de la investigación. Si no se logra, es preciso regresar al inicio y revisar la teoría de medición que sustenta este trabajo.

Una vez asegurado un modelo de medida válido y confiable, el objetivo de la tercera etapa de investigación es poner a prueba y validar la teoría estructural propuesta para explicar la relación entre la evaluación del desempeño individual de investigación y las formas diferenciales de producción de conocimiento. Nuevamente, la teoría estructural se representa formalmente a través del modelo teórico estructural, que encarna las hipótesis de investigación orientadoras de este trabajo. Como se desprende y a diferencia de la etapa anterior, aquí las hipótesis se hacen sobre las relaciones teóricas estructurales entre las dimensiones conceptuales o variables latentes. En contraste, en la segunda etapa, el análisis se concentró en validar las medidas, es decir, las relaciones entre las variables de medida observadas y dichas dimensiones.

Contrastar y validar la teoría estructural que sustenta este trabajo implica, por una parte, analizar la bondad de ajuste para evaluar la aceptación o no del modelo teórico estructural propuesto, y por otra, estudiar la naturaleza y significado conceptual de los parámetros estructurales estimados. Los parámetros estructurales representan las hipótesis que se realizaron sobre las relaciones de dependencia entre las dimensiones conceptuales o variables latentes. Si los resultados empíricos sostienen el modelo teórico estructural, se habrá encontrado evidencia que valida la teoría estructural que se propone, y por lo tanto, nuevos elementos de juicio que contribuyen al

entendimiento de las relaciones teóricas entre la evaluación de la investigación y las formas diferenciales de producción de conocimiento.

A continuación se describen en profundidad cada una de las etapas de investigación y las técnicas de análisis multivariado asociadas a cada una de ellas.

## VIII.2 Etapa 1: Exploración y descripción de las relaciones entre las variables observadas, el análisis factorial exploratorio

El AFE se implementa para identificar el conjunto de dimensiones conceptuales latentes, su interpretación permite la comprensión y descripción de dichos datos en un número menor de dimensiones conceptuales que las variables originales.

La idea central en el resumen de los datos es la *estructura*. De acuerdo con Hair et al. (2014), la estructura de correlación permite observar las variables en diversos niveles de generalización. Desde el nivel más detallado, donde las variables observadas importan por lo que representan individualmente; hasta el más generalizado, donde las variables individuales son agrupadas y su relevancia está dada por lo que representan colectivamente en la expresión de un concepto. Para identificar dicha estructura de correlación es esencial seleccionar, para el análisis, sólo aquellas variables observadas que se espera den cuenta de las dimensiones conceptuales que se espera encontrar. En la sub-sección X.2.1 del capítulo X se presenta la tabla 21 con todas las variables observadas seleccionadas para el análisis.

### VIII.2.1 Diseño del AFE

Existen tres aspectos clave que hay que considerar desde el punto de vista del diseño del AFE.

El primero se vincula con el hecho de que debido a que el interés está en encontrar la estructura latente entre las variables, la implementación del AFE requiere, para ello, del *cálculo de la matriz de correlación entre las variables*.

El segundo tiene que ver con el *tipo y número de variables a incluir*, así como con sus *propiedades métricas*. La importancia de considerar el tipo de variables está dada porque determina si es posible o no el cálculo de la matriz de correlación requerida. Además, atender a las propiedades métricas de las variables es importante porque afecta la medida de correlación que se calcula entre ellas (Pearson, policórica, tetracórica, poliserial, etc.). En el contexto de este trabajo y como se observa en la tabla 21, todas las variables observadas seleccionadas para el análisis son categóricas: ordinales y dicotómicas. Esto implica que la matriz de correlación que se calcula es la matriz de correlación policórica. Adicionalmente, el número de variables a incluir en el análisis también se debe tener en cuenta porque de ello depende la fortaleza del AFE para encontrar la estructura de correlación entre las variables. La literatura recomienda incluir en el análisis cinco o más variables por factor esperado (Hair et al., 2014). En este trabajo se incluyeron 33 variables observadas para identificar –idealmente– las cinco dimensiones conceptuales que encarnan los objetivos de investigación de esta tesis.

El tercer aspecto tiene que ver con el *tamaño de la muestra* mínimo requerido, tanto en términos absolutos como relativos al número de variables incluidas en el análisis. La importancia de contar con un tamaño de muestra adecuado radica en la probabilidad de obtener soluciones generalizables y no específicas a la muestra. Si bien en la literatura existe debate en torno al tamaño muestral

requerido, existen algunas reglas generales que sirven de orientación. De acuerdo con Hair et al. (2014) el AFE no debería realizarse con muestras de menos de 50 observaciones y es deseable que contengan 100 o más. Además señala que un tamaño de muestra más aceptable es aquel que provee 10 observaciones por variable incluida, algunos autores son más exigentes y postulan que deben ser 20 observaciones por cada variable. De acuerdo con estas reglas generales, el tamaño de la muestra de este trabajo es lo suficientemente grande como para garantizar soluciones generalizables. En términos absolutos se cuenta con  $N = 8053$  observaciones, lo que da lugar a un total de 244 observaciones por variable incluida en el análisis.

### VIII.2.2 Sentido de implementar el AFE

El supuesto básico y que da sentido a la implementación del AFE tiene que ver con que existe alguna estructura subyacente a las variables seleccionadas para el análisis. Sin embargo, la presencia de variables correlacionadas, cuya interpretación da lugar a la definición de los factores, no necesariamente garantiza relevancia conceptual, incluso aunque se cumplan empíricamente todos los requisitos estadísticos (porcentaje de varianza explicada, bondad de ajuste, etc.). La relevancia conceptual de la estructura factorial identificada depende del análisis crítico e informado que se realice.

De acuerdo con Hair et al. (2014) hay diversos enfoques para establecer la existencia de alguna estructura subyacente a las variables seleccionadas y de ese modo justificar la implementación del AFE. Entre los enfoques que mencionan y describen estos autores están: i) la inspección visual de la matriz de correlación; ii) el análisis de las correlaciones parciales; iii) la prueba de esfericidad de Bartlett; iv) la medida de adecuación muestral; v) el valor del determinante de la matriz de correlación; y vi) el índice Kaiser-Meyer-Olkin.

Aquí sólo se describen dos de estos enfoques, la *inspección visual de la matriz de correlación* y el *valor del determinante de dicha matriz*, ya que estos fueron los que se utilizaron para justificar el sentido de implementar el AFE en el contexto de este trabajo.

En primer lugar se calculan todas las correlaciones entre todas las variables seleccionadas para el análisis, dando lugar así a la matriz de correlación policórica. Se trata de correlaciones policóricas porque entre las variables seleccionadas algunas son ordinales y otras son dicotómicas. En segundo lugar, se inspecciona visualmente dicha matriz para asegurar que existe una cantidad importante de correlaciones considerables que justifiquen el AFE. De acuerdo con Hair et al. (2014) una correlación “considerable” es aquella que presenta un valor absoluto de 0.3 o más. Adicionalmente, es preciso observar cierto patrón de correlación entre las variables de medida. En particular, se espera que las variables que componen cada uno de los factores tengan correlaciones altas entre sí, y además que presenten correlaciones más bajas con las variables que componen los restantes factores. Este tipo de patrón proporciona evidencia que sugiere o refleja que sí existe una estructura de correlación que subyace a los grupos de variables, y por lo tanto tiene sentido implementar el AFE para justamente identificar dicha estructura.

Otra medida de correlación entre las variables puede ser obtenida a través del cálculo del *Determinante de la Matriz de Correlación*. El criterio es que si el valor del determinante es pequeño y cercano a cero, entonces hay evidencia que sugiere que existe una estructura de correlación fuerte que subyace a las variables y por lo tanto la implementación del AFE puede ser pertinente. En la

sección X.2.1 del capítulo X, se presentan en la tabla 22 la matriz de correlación policórica y el valor de su Determinante, y se concluye que sí tiene sentido la implementación del AFE.

### VIII.2.3 Implementación del AFE: obtención de los factores y evaluación del ajuste global

Una vez especificadas todas las variables a ser incluidas en el AFE (ver tabla 21), calculada la matriz de correlación y su determinante, y al haber concluido que sí tiene sentido la implementación del AFE, corresponde la estimación. Para estimar o implementar el AFE es preciso tomar algunas decisiones sobre el método de extracción de los factores y la cantidad de factores que efectivamente se van a retener, y de ese modo determinar las dimensiones del análisis. A continuación se describen y explicita la estrategia seguida para ello.

En la literatura es posible identificar dos métodos para la extracción de los factores, a saber: i) análisis de factores comunes (*common factor analysis*), y ii) análisis de componentes principales (*principal components analysis*). La elección de cada uno de ellos sobre el otro depende del objetivo central que se quiera alcanzar con el AFE (Hair et al., 2014).

En el contexto de este trabajo, el método de extracción factorial seleccionado es el análisis de factores comunes, porque el objetivo es identificar las dimensiones conceptuales latentes que reflejan lo que las variables tienen en común o comparten<sup>56</sup>. El otro elemento a tener en cuenta es el método de estimación de los parámetros. En particular, debido al tipo de variables que se incluyen en el análisis y sus propiedades métricas, se seleccionó la estimación por mínimos cuadrados ponderados con media y varianza ajustada (WLSMV). En la sub-sección VIII.3.3.4 se discute este aspecto.

El siguiente paso es determinar el número de factores –dimensiones conceptuales latentes- a extraer o retener para representar la estructura de correlación subyacente a los datos. Para ello, a continuación se describen los criterios utilizados y la estrategia seguida.

La determinación de cuántos factores se deben retener depende de fundamentos conceptuales y la evidencia empírica. Los fundamentos conceptuales refieren a la cantidad de factores que se debería tener de acuerdo con la teoría. En el contexto de este trabajo y de acuerdo con la teoría se espera identificar –idealmente- una estructura multifactorial de cinco factores. La evidencia empírica refiere a la cantidad de factores que es razonable esperar de acuerdo con el análisis preliminar de la matriz de correlación (ver resultados en sub-sección X.2.1 del capítulo X, tabla 22) y otros criterios empíricos que se detallan a continuación.

Los criterios empíricos que se utilizaron en este trabajo para determinar la dimensionalidad del problema de investigación que se aborda, combinan los propuestos por Kaiser (1960) y Hair et al. (2014) y son:

- i) *el criterio de la raíz latente*, su racionalidad radica en que cada factor debe acumular como mínimo la varianza de al menos una variable, por lo tanto los factores a retener –

---

<sup>56</sup> Cuando una variable se correlaciona con otra variable, se dice que ambas variables comparten varianza, y la cantidad de varianza compartida entre ambas variables es el cuadrado de la correlación. En el análisis factorial, se agrupan variables de acuerdo a su correlación, formando así los factores, las variables que conforman el factor presentan altas correlaciones entre ellas. Cuanto más correlacionada está una variable con otras, mayor es la varianza común –*comunalidad*- o compartida. (Hair et al., 2014)



- los significativos- son todos aquellos que tienen raíces latentes o valores propios mayores a 1 (ver resultados en sub-sección X.2.2 del capítulo X, tabla 23);
- ii) *el criterio de la varianza*, implica alcanzar un porcentaje específico de la varianza total acumulada extraída por los factores, trata de asegurar que los factores expliquen al menos cierta cantidad de varianza; en el contexto de la investigación empírica en ciencias sociales, una solución factorial satisfactoria es aquella en la que los factores a retener acumulan –explican- al menos el 60% de la varianza total (ver resultados en sub-sección X.2.2 del capítulo X, tabla 23); y
  - iii) la inspección visual del *gráfico de sedimentación (scree plot)*, este gráfico dibuja la curva derivada de la unión de los puntos determinados por los valores propios de cada factor, cuando cambia la pendiente de dicha curva, se establece esa inflexión como punto de corte e indica el número máximo de factores a retener (ver resultados en sub-sección X.2.2 del capítulo X, gráfico 18).

Además y complementariamente, se analizaron las medidas de bondad de ajuste derivadas de estimar estratégicamente el AFE con diversas estructuras factoriales, desde la unifactorial de 1 sólo factor hasta la multifactorial de 7 factores. Las medidas de bondad de ajuste se detallan en la sub-sección VIII.3.4.1, aquí sólo se señala que las medidas utilizadas fueron el estadístico Ji-Cuadrado de bondad de ajuste, el RMSEA, el CFI y TLI; y el criterio general empleado para la selección de la estructura factorial más apropiada fue respetar los valores de referencia de corte establecidos en la literatura y que indican si el ajuste de los modelos factoriales estimados (de 1 a 7 factores) son aceptables o no -ver resultados en sub-sección X.2.2 del capítulo X, tabla 24-.

La combinación y análisis de todos estos criterios empíricos, da como resultado una idea inicial del número de factores a retener. Específicamente y en el contexto de este trabajo, atendiendo a lo que se espera de acuerdo a la teoría y los hallazgos de investigaciones pasadas (ver capítulo V) y al análisis combinado de todos estos criterios empíricos, se exploran y discuten tres estructuras multifactoriales (la solución multifactorial de 4 factores, la de 5 y la de 6; ver resultados en sub-secciones X.2.2.1 a X.2.2.3 del capítulo X).

La determinación final del número de factores o identificación de la dimensionalidad del problema de investigación depende de la interpretación conceptual de los factores y su discusión. A continuación se describe el proceso que se sigue en este trabajo.

#### VIII.2.4 Interpretación conceptual de los factores e identificación de la estructura factorial más apropiada para los objetivos de este trabajo

Se utilizan tres procesos clave en esta etapa para la interpretación y posterior identificación de la estructura factorial (Mulaik, 2010; Hair et al. 2014), a saber: i) importancia de las cargas factoriales estimadas en la conceptualización del factor; ii) interpretación conceptual de los factores; y iii) la rotación factorial.

Cuando se implementa el AFE, su solución provee –entre otras cosas- las cargas factoriales de cada variable observada en cada uno de los factores. Las cargas factoriales son las correlaciones entre cada variable observada y el factor y por lo tanto, indican el grado de asociación entre la variable y dicho factor. En ese sentido, cargas factoriales altas significan que la variable contribuye a formar

dicho factor<sup>57</sup>. Así, a través de las cargas factoriales, se interpreta el rol de cada variable en la definición conceptual de los factores.

De acuerdo con Hair et al. (2014), cuanto mayor sea el valor absoluto de la carga factorial, mayor es la importancia de dicha carga en la interpretación factorial. De acuerdo con estos autores, cargas factoriales mayores que |0.3| se utilizan para la interpretación de la estructura factorial mínima; de |0.5| o más son aceptables; y de |0.7| o más son indicadoras de que se ha definido una estructura factorial aceptable. Además, si el número de observaciones es grande, los valores umbrales de las cargas factoriales también son probablemente significativos desde el punto de vista estadístico, con un nivel de confianza del 95% (Hair et al., 2014). De todas maneras, y a efectos de asegurar la significancia estadística de las cargas factoriales para la interpretación de los factores, entre los resultados de la estimación del AFE se encuentran los asociados a las pruebas específicas para la significancia estadística de cada una de las cargas factoriales estimadas (ver resultados en tablas 25, 26 y 27 de las sub-secciones X.2.2.1, XX.2.2.2 y XX.2.2.3 del capítulo X).

Una vez analizadas las cargas factoriales (magnitud y significancia estadística) de la solución sin rotar, se procede con la interpretación conceptual de cada uno de los factores. Sin embargo, y en la mayoría de los casos, el proceso de interpretación resulta difícil por la complejidad de la solución obtenida. En general, la estructura factorial no emerge con claridad debido, por ejemplo, a la existencia de cargas factoriales cruzadas. Es decir, una variable observada puede tener cargas factoriales importantes y significativas en más de un factor, dificultando la interpretación factorial.

Una forma de simplificar la interpretación es a través de la rotación factorial (Hair et al., 2014; Mulaik, 2010). De hecho, la rotación de los factores tiene por objetivo simplificar la estructura factorial. La rotación factorial no cambia los resultados, sino su perspectiva. Existen diversos métodos de rotación factorial, ortogonales u oblicuos. Los métodos de rotación factorial ortogonales mantienen a los factores perpendiculares entre sí, es decir, independientes. Por el contrario, los métodos de rotación factorial oblicuos relajan la restricción de perpendicularidad entre los factores, es decir, ya no son tratados como si fueran independientes, y su solución provee información sobre el grado de correlación existente entre ellos.

Debido a la complejidad de las soluciones factoriales obtenidas (para estructuras multifactoriales estimadas de 4, 5 y 6 factores), fue necesario realizar una rotación de los factores para simplificar dichas soluciones y así poder interpretar –conceptualmente- cada uno de los factores. Si bien las rotaciones ortogonales son las que se utilizan con mayor frecuencia, en el contexto de este trabajo y debido a que no hay elementos de juicio para sostener independencia entre los factores –sino todo lo contrario- se optó por utilizar un método oblicuo de rotación, en particular, el método geomín (Muthén y Muthén, 1998-2015; Yates, 1987; Browne, 2001).

En las sub-secciones X.2.2.1, XX.2.2.2 y XX.2.2.3 del capítulo X, se discute la relevancia conceptual de las soluciones factoriales encontradas y se fundamenta la necesidad de poner a prueba una teoría de medida multifactorial de cinco factores para la relación entre evaluación y producción de conocimiento. Para ello se implementa el AFC, que a continuación se describe en profundidad.

---

<sup>57</sup> O dicho de otro modo, que ese factor explica una proporción elevada de la varianza de esa variable. Por ejemplo: si la carga factorial de una variable es 0.7 en un factor, es posible sostener que ese factor explica el 49% de la varianza de esa variable ( $0.7 \times 0.7 = 0.49$ ).

### VIII.3 Etapa 2: Medición y validación, el análisis factorial confirmatorio

Luego de la implementación del AFE, cuyos resultados proveen la evidencia empírica necesaria para justificar la construcción y desarrollo de una *teoría de medida válida y confiable*, el objetivo de esta sección es describir el conjunto de procedimientos y técnicas de análisis multivariado que se realizan para construir, desarrollar y validar dicha teoría.

Para ello, lo primero que hay que establecer es que dicha teoría se formaliza a través de un modelo, a saber: el *modelo de medida*. A la vez, dicho modelo se representa gráficamente mediante un *diagrama de senderos*, que contiene un error de medición. La formalización de la teoría y su representación gráfica tienen por objetivos, en primer lugar, mostrar cómo las variables observadas colectivamente representan ciertos *constructos* o conceptos; y en segundo lugar, contrastar y validar dicha teoría con la realidad.

En este contexto, el AFC, como técnica de análisis multivariado, permite evaluar el grado con el cual el modelo de medición refleja ciertos conceptos o constructos, que son inobservables directamente por su complejidad y multidimensionalidad. Sin embargo, son estos conceptos teóricos –también llamados constructos– los que explican la variación de las respuestas observadas de los individuos sujetas a cierto error de medición. Es decir, el AFC permite probar analíticamente una teoría conceptualmente fundamentada (Hair et al., 2014; Bollen, 1989). De acuerdo con Brown (2015), el AFC es una herramienta fundamental para la validación de la teoría de medida en general y de los constructos en particular.

El AFC supone una teoría de medida *reflectiva*, no *formativa* (Hair et al., 2014). Es decir, la idea central es que los constructos latentes *causan* a las variables manifiestas y los errores de medición, son una medida de la incapacidad de dichos constructos para explicar la totalidad de la variabilidad observada. En este sentido, y en el contexto de este trabajo, el enfoque de medición que se utiliza es a través de la formalización y representación de un modelo reflectivo y, por lo tanto, los constructos latentes que se definen también son de naturaleza reflectiva.

Además, en este trabajo se combinan los resultados del AFC con diversas pruebas sobre la validez de constructo, lo que permite una mejor comprensión de la calidad de las medidas que se utilizan para entender la relación entre la evaluación del desempeño individual de investigación de los investigadores con los modos diferenciales de producción de conocimiento.

La preocupación por la calidad de la medición –en términos de su validez y precisión–, determinada por el análisis que supone esta técnica de análisis multivariado y la posibilidad de contabilizar la magnitud de los errores de medida, es una ventaja y una característica diferencial y sobresaliente del AFC, que lo posiciona por encima de otras técnicas de análisis multivariado –por ejemplo: regresión múltiple–. Esto se debe a que permite un análisis más riguroso y más confiable sobre la validez conceptual y empírica del modelo teórico de medición que se está poniendo a prueba (Brown, 2015; Bollen, 1989).

Los conceptos que se buscan medir y relacionar en esta investigación (propensión modo 1, propensión modo 2, satisfacción con el proceso de evaluación, flexibilidad epistemológica del sistema de evaluación y dirección de la reforma del sistema de evaluación), podrían haber sido contruidos arbitrariamente a partir, por ejemplo, de la agregación o alguna combinación de las variables que se considera *a priori* asociadas a ellos. Sin embargo, en este trabajo se optó por el AFC

justamente porque mejora sustantivamente la calidad de la investigación, es decir, la validez y confiabilidad de los resultados de investigación.

A diferencia del AFE, y como se desprende de lo dicho, el AFC tiene una orientación predominantemente teórica. Es decir, se orienta por las hipótesis que se hacen sobre la relación entre las variables de medida y los constructos o *variables latentes*. De alguna manera, el proceso que se sigue con el AFC es inverso al que se sigue con el AFE. Mientras que en el AFE los factores cobran sentido conceptual según la importancia relativa de las cargas factoriales de cada variable observada en cada factor (conceptualización empírica), en el AFC se especifica *ex-ante* y siguiendo a la teoría, los conceptos/constructos/variables latentes que se busca medir y su relación con las variables observadas de medida: las hipótesis (conceptualización teórica) que se contrastan con la realidad. Es decir, se definen conceptualmente las variables latentes, se operacionalizan con las variables de medida observadas y se contrastan las hipótesis con la realidad *–los datos–*. Además, se hacen un conjunto de pruebas para asegurar su validez y confiabilidad.

Una vez asegurada la validez y confiabilidad se calculan las puntuaciones factoriales estandarizadas de los individuos. Estas puntuaciones factoriales se calculan como una medida compuesta que combina los valores que asume cada una de las variables originales (respuestas de los individuos en cada uno de los ítems) y los pondera por el peso o carga factorial del ítem en la variable latente.

En síntesis, la teoría de medida es útil para especificar *–a través de las hipótesis–* cómo un conjunto determinado de variables de medida representan ciertos constructos o variables latentes. Las relaciones clave de análisis en esta segunda etapa de la investigación son, por un lado, entre los constructos y las variables de medida, que se observan a través de las cargas factoriales estimadas; y por otro, entre los constructos, que se observan a través de las correlaciones estimadas entre ellos. De acuerdo con Hair et al. (2014), con estas estimaciones es posible realizar un análisis empírico de la teoría de medida que se somete a prueba. A continuación se detallan la secuencia de actividades que se realizan y los criterios que se utilizan para construir, desarrollar y validar dicha teoría.

### VIII.3.1 La teoría de medida, sus hipótesis y la consecuente definición de los constructos o variables latentes

El primer paso para implementar el AFC se vincula con el desarrollo de la teoría de medida que se quiere poner a prueba. Desarrollar dicha teoría de medida implica como actividad inicial definir conceptualmente cada uno de los constructos o variables latentes (ver capítulo V). Una vez que las variables latentes están definidas conceptualmente, la segunda actividad es establecer un conjunto de hipótesis sobre la relación entre cada variable latente y el conjunto de variables observadas de medida, cuya variación se hipotetiza que puede ser explicada por dicha variable latente.

En el contexto de este trabajo se definen conceptualmente cinco variables latentes *–satisfacción con el proceso de evaluación, flexibilidad epistemológica del sistema de evaluación, dirección de la reforma del sistema de evaluación, propensión modo 1 y propensión modo 2 (en la la tabla 28 se retoman sus definiciones conceptuales)–*, y se establecen cinco hipótesis que vinculan las dimensiones conceptuales definidas con determinados sub-conjuntos de variables observadas de medida (ver sub-sección X.3.1 del capítulo X para una descripción detallada de las hipótesis realizadas y la racionalidad detrás de cada una de ellas).

La definición conceptual de las variables latentes y el establecimiento de las hipótesis, implican un avance significativo hacia la operacionalización de las dimensiones conceptuales centrales de la investigación. Así, la tercera actividad específica es justamente la operacionalización de las variables latentes definidas conceptualmente. Esto implica seleccionar el conjunto de variables manifiestas (*indicadoras/ítems*) que se utilizan –justamente– para medir cada una de dichas dimensiones conceptuales. Además, para lograr la medición, es necesario definir la escala de medida de los ítems. De acuerdo con Hair et al. (2014), los dos enfoques más comunes para esto son: i) tomar las escalas definidas en investigaciones pasadas; o, ii) desarrollar una nueva escala.

En el proceso de la operacionalización de las variables latentes -además de seleccionar las variables manifiestas y definir la escala de medición-, hay que considerar ciertos temas operacionales (*identificación y dimensionalidad*) y de validez de constructo (Hair et al., 2014).

Desde el punto de vista operacional es importante incluir al menos tres variables manifiestas como indicadoras por cada constructo para su identificación. Para asegurar la unidimensionalidad, cada variable manifiesta puede ser explicada por un solo constructo. A la vez y desde la perspectiva de la validez de constructo, es en este proceso inicial de construcción y desarrollo de la teoría de medida que se realiza la primera prueba de validez, a saber: la *prueba de validez aparente* o de *contenido* (Hair et al., 2014). Esta prueba, que es de carácter cualitativo y está relacionada con la validez de los constructos, consiste en analizar el grado de coincidencia entre las definiciones conceptuales de los constructos y las variables manifiestas que los operacionalizan. Se realiza por medio de la evaluación de expertos externos y/o por investigaciones pasadas, y el marco teórico juega un papel preponderante.

En la tabla 28 de la sub-sección X.3.1 del capítulo X se presenta la operacionalización de cada una de las variables latentes definidas conceptualmente en este trabajo y sus escalas de medición. Vale señalar que la mayoría de las escalas de los ítems son tomadas de las definidas en el cuestionario realizado en el marco de la investigación titulada “*Reflexiones sobre la evaluación a los investigadores: una mirada desde diferentes perspectivas*” coordinada por el FCCyT y documentada en Bensusán et al. (2014). Otras escalas fueron re-definidas debido a sus distribuciones o para mantener la consistencia entre la medición y la definición conceptual del constructo (en la tabla 24 se detallan todas estas transformaciones). Todos los constructos son operacionalizados con más de tres ítems para asegurar su identificabilidad y, simultáneamente, la variabilidad de cada variable manifiesta es explicada únicamente por una variable latente. Además, la primera prueba de validación de constructo, la *prueba de validación de contenido* de este trabajo ha sido realizada en las diversas instancias de revisión y evaluación previstas en los coloquios del doctorado y llevada adelante por los integrantes del tribunal de esta tesis.

### VIII.3.2 Desde la teoría hacia el modelo: consideraciones para la especificación del modelo de medida

Una vez que las variables latentes fueron definidas tanto en lo conceptual como en lo operacional, y que la escala de los ítems fue establecida, se puede continuar hacia el siguiente paso: especificar el modelo de medición. El modelo de medida es la formalización matemática de la teoría de medición que se propone, que a su vez se representa gráficamente a través de un diagrama de senderos.

Sin embargo, antes de describir los elementos a considerar para la especificación del modelo, es preciso introducir el *enfoque de variables de respuesta latente* (Muthén y Asparouhov, 2002).

Con el AFC y la especificación del modelo de medición lo que se busca evaluar –en última instancia– es el grado con el que dicho modelo reproduce la matriz de varianzas y covarianzas muestral. Dicho de otro modo: ¿en qué medida la matriz de varianzas y covarianzas estimada ( $\Sigma_k$ ) difiere de la matriz de varianzas y covarianzas observada ( $S$ )?

Sin embargo, debido a las propiedades métricas de las variables de medida observadas, que son de naturaleza discreta ordinales y dicotómicas, no es posible –por definición– cumplir con los supuestos de la teoría normal (ej.: normalidad multivariada de las variables observadas), que se requiere para poder utilizar el estimador de máxima verosimilitud y su respectiva función de ajuste, con el que en general se estiman los modelos de ecuaciones estructurales cuando se tienen variables continuas normalmente distribuidas, incluido el AFC como un caso particular de dichos modelos (Finney y Distefano, 2006).

Más aún, de acuerdo con Brown (2015), los procedimientos involucrados en la estimación del AFC cuando se utilizan indicadores dicotómicos u ordinales difieren de los procedimientos de estimación de los modelos AFC en el marco de la teoría normal con variables continuas. En particular, la matriz de varianzas y covarianza muestral no puede ser calculada directamente a partir de las variables observadas originales.

Es por esta razón que antes de plantear los elementos para la formalización matemática del modelo de medida, conviene desarrollar brevemente el enfoque de *variables de respuesta latente* (Muthén y Asparouhov, 2002). Este enfoque provee un marco conceptual integrado para la estimación de modelos de ecuaciones estructurales que involucran variables observadas categóricas (dicotómicas u ordinales) o alguna otra combinación de variables categóricas y continuas (Muthén, 1998-2004; Finney y Distefano, 2006; Kline, 2011; Brown, 2015).

De acuerdo con Kline (2011), entre otros, el enfoque de *variables de respuesta latente* supone que cada indicador categórico observado ( $y_j$ ) se asocia con una variable de respuesta latente subyacente ( $y_j^*$ ), que contiene valores teóricos (*umbrales*,  $\tau$ ) en un continuo normalmente distribuido de las características requeridas para responder una determinada categoría ( $c$ ) de la variable ordinal o dicotómica observada ( $y_j$ ). Donde el número total de umbrales  $\tau$  es igual al número de categorías de la variable observada menos uno ( $c - 1$ ). Además, los umbrales forman parte de la estructura media del modelo de medida en el contexto del AFC con indicadores categóricos (Brown, 2015). Por ejemplo, si la variable observada ( $y$ ) es ordinal de escala Likert con cinco categorías, se requieren cuatro valores umbrales ( $\tau$ ) para dividir ( $y^*$ ) en cinco categorías ordinales:

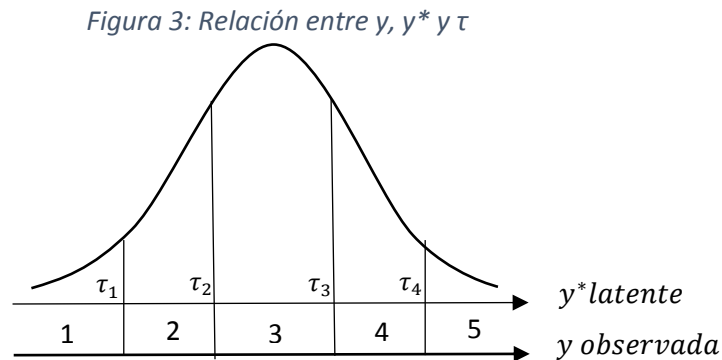
$$y = \begin{cases} 1 & \text{si } y^* \leq \tau_1 \\ 2 & \text{si } \tau_1 < y^* \leq \tau_2 \\ 3 & \text{si } \tau_2 < y^* \leq \tau_3 \\ 4 & \text{si } \tau_3 < y^* \leq \tau_4 \\ 5 & \text{si } y^* > \tau_4 \end{cases}$$

Si la variable observada ( $y$ ) es dicotómica, se requiere un solo valor umbral ( $\tau$ ) para dividir ( $y^*$ ) en dos categorías:

$$y = \begin{cases} 0 & \text{si } y^* \leq \tau_1 \\ 1 & \text{si } y^* > \tau_1 \end{cases}$$

Si la variable observada ( $y$ ) es continua y normalmente distribuida, entonces  $y = y^*$ .

Finney y Distefano (2006) representan la relación entre  $y$  (variable observada de medida con cinco categorías),  $y^*$  (variable de respuesta latente con cuatro umbrales) y  $\tau$  (umbrales) como se muestra en la figura 3.



Así y de acuerdo con Kline (2011), cada variable de respuesta latente es representada como un indicador continuo del factor o constructo hipotético.

La relación clave ya no es entre  $y$  observada y el constructo -que no es lineal-, sino entre  $y^*$  y el constructo -que sí es lineal- (Muthén, 1998-2004). De este modo, la matriz de datos que se analiza bajo este enfoque es *la matriz de correlación asintótica de las variables de respuesta latente*, y no la matriz de varianzas y covarianzas de las variables de medida observadas originales, que se analizaría si las variables fueran continuas y normalmente distribuidas.

El supuesto detrás es que las asociaciones bivariadas entre las variables observadas pueden ser estimadas porque existe un proceso normal y continuo subyacente a cada variable observada original (Kline, 2011). Más aún, y de acuerdo con Brown (2015), dado que se utiliza la matriz de correlación de  $y^*$  como  $S$ , las varianzas de los indicadores observados (*varianzas únicas*, que no pueden ser explicadas por la estructura del modelo de medida teórico propuesto, es decir, los *errores de medida*) no son considerados como parámetros libres a estimar. Esto representa otra diferencia con los modelos AFC en el marco de la teoría normal con variables continuas. Sin embargo, dichos errores de medida ( $\theta$ ) se pueden calcular como residuos -varianzas residuales- (Brown, 2015). Para ello, se fija la varianza de  $y^*$  en 1 para todos los indicadores, lo que está en sintonía con la idea de que la matriz de correlación ( $S$ ) coincide con una matriz de varianzas y covarianzas de  $y^*$ , con varianza igual a 1 (la matriz de correlación no es otra cosa más que la matriz de varianzas y covarianzas estandarizadas). Entonces, el error de medida o varianza residual ( $\theta$ ) en la solución completamente estandarizada va a ser igual a la varianza de  $y^*$  (que es igual a 1 dada la estandarización) menos la carga factorial estandarizada al cuadrado -que refleja la proporción de la varianza de  $y^*$  que es explicada por la variable latente- multiplicada por la varianza de la variable latente ( $\phi$ , que en la solución completamente estandarizada  $\phi = 1$ ). A saber:

$$\theta = 1 - \lambda^2 \phi$$

ó simplemente

$$\theta = 1 - \lambda^2$$

Intuitivamente, las varianzas residuales ( $\theta$ ) expresan la proporción de la varianza de  $y^*$  que no es explicada por la variable latente (Brown, 2015). En el marco conceptual del M-PLUS –paquete estadístico que se utiliza para la estimación en este trabajo, este enfoque se denomina como *parametrización delta* (Brown, 2015). Para una discusión detallada de las parametrizaciones *delta* y *theta* ver Muthén y Asparouhov (2002).

Dado que en los párrafos precedentes se han explicitado las implicancias que tienen las propiedades métricas de las variables observadas de medida sobre la especificación del modelo, ahora sí es posible indicar los elementos que deben considerarse para dicha especificación.

Así la especificación completa del modelo de medida supone, primero especificar las relaciones de medida entre  $y_j^*$  -la variable de respuesta latente asociada a cada variable manifiesta observada- y las variables latentes. Estas relaciones se observan a través de los parámetros de medida, que no son otra cosa que las cargas factoriales o los coeficientes de pendiente de las regresiones probit que se estiman y, representan el grado en el que cada ítem ( $y_j^*$ ) está relacionado con su constructo. Segundo, especificar las relaciones de correlación entre los constructos; y tercero especificar los errores de medición o varianzas residuales, que reflejan la proporción de la varianza de  $y_j^*$  que no es explicada por la variable latente.

Además de lo señalado, también es importante asegurar -para la especificación del modelo de medida- la *unidimensionalidad* de los constructos o variables latentes y su *identificación*. De acuerdo con Hair et al. (2014), la unidimensionalidad de cada constructo significa que la variabilidad de las variables manifiestas (mejor dicho de  $y_j^*$ ), que son indicadoras de cada uno de los constructos, puede ser explicada únicamente por ese constructo o variable latente. Es decir, en la especificación del modelo de medida, todas las cargas factoriales cruzadas se restringen y se fijan igual a cero. Cuando un ítem ( $y_j^*$ ) es causado por más de una variable latente, la unidimensionalidad no se cumple y eso afecta negativamente a la validez discriminante y consecuentemente, la validez de constructo. Es decir, se afecta de manera negativa las propiedades de medida de los constructos o variables latentes. A la vez, también es importante –para no afectar la validez de constructo- restringir la covariación de los errores de medición.

De acuerdo con Hair et al. (2014), los modelos de medida que restringen las cargas factoriales cruzadas y la covariación entre los términos de error, fijándolas en cero, se denominan *modelos de medida co-genéricos*, y cuentan con suficientes restricciones como para garantizar buenas propiedades de medición.

El modelo de medición propuesto en este trabajo sigue estas pautas de especificación (se fijan en cero las cargas factoriales cruzadas y los errores de medida se especifican independientes), por lo que si se satisfacen empíricamente todos estos requerimientos, es esperable que dicho modelo presente validez de constructo y sea consistente con una buena práctica de medición.

La *identificación* estadística de cada uno de los constructos depende de la existencia de información suficiente para identificar una solución matemática para el conjunto de ecuaciones determinadas en el modelo (Hair et al., 2014; Bollen 1989). De acuerdo con estos autores, los constructos y los modelos pueden caracterizarse por su grado de identificación como: i) *sub-identificados*: existen



más parámetros a estimar que información disponible, por lo tanto no es posible encontrar una única solución matemática; ii) *identificados*: existe la misma cantidad de parámetros a estimar que información disponible; en este caso los grados de libertad del modelo son cero, se utiliza toda la información para estimar el modelo y su ajuste es adecuado; se trata del *modelo saturado* y su utilidad es nula, si lo que se busca es poner a prueba una teoría; y, iii) *sobre-identificados*: existen menos parámetros a estimar que información disponible, los grados de libertad del modelo son positivos y sí es posible encontrar una solución matemática. Alcanzar modelos sobre-identificados forma parte de los objetivos de implementar el AFC (Hair et al., 2014; Bollen, 1989).

El modelo de medida propuesto en este trabajo está *sobre-identificado*. De acuerdo con Muthén y Asparouhov (2002: 5), el número máximo de parámetros identificables del modelo -o información disponible- está dado por la siguiente expresión:

$$\left(\frac{p(p-1)}{2}\right) + r$$

Dónde  $p$  es el número de variables del modelo y  $r$  es el número de umbrales.

En la sub-sección X.3.2 del capítulo X se especifica formalmente el modelo de medida propuesto en este trabajo y se representa gráficamente a través de un diagrama de senderos (figura 4).

### VIII.3.3 Diseño del AFC

A esta altura del proceso de investigación ya se ha especificado el modelo de medición y procede entonces su estimación, análisis y validación. Sin embargo, antes de estimar el modelo de medición es importante describir ciertos aspectos del diseño del AFC y de la investigación en general<sup>58</sup>, ya que de ellos depende la calidad de los resultados de investigación que se obtengan.

Los temas del diseño que se tratan en esta sub-sección se vinculan con las características específicas de la fuente de información, es decir, el tipo de datos a analizar; el impacto y tamaño de la muestra, y su representatividad; los datos faltantes, su impacto y tratamiento; y la técnica de estimación y el paquete estadístico utilizado para ello.

#### VIII.3.3.1 Fuente de información: tipo de datos a analizar

Como se señaló, los datos que se analizan en este trabajo fueron recolectados a través de la encuesta titulada "*Percepciones sobre la evaluación académica*" (FCCyT, 2014) auspiciada por el Foro Consultivo Científico y Tecnológico (FCCyT) durante el período comprendido entre los años 2013 y 2014. Dicha encuesta fue diseñada por el grupo de investigación integrado por Graciela Bensusán (coord.), Natalia Gras, Daniel Inclán, Erwin Rodríguez, Giovanna Valenti (coord.) y Gonzalo Varela, en el marco del proyecto de investigación que dio lugar al reporte titulado "*Reflexiones sobre la evaluación a los investigadores: una mirada desde diferentes perspectivas*" (citado en esta tesis como Bensusán et al., 2014).

El objetivo de dicha encuesta fue conocer las percepciones de los investigadores sobre diversas dimensiones de los distintos procesos de evaluación académica -SNI, PROMEP e institución de

---

<sup>58</sup> Los temas sobre diseño que se tratan en esta sub-sección VIII.3.3, son válidos para las tres etapas de la investigación (AFE, AFC y SEM). Cuando se describió la etapa 1 de investigación (VIII.2), en la sub-sección VIII.2.1 se señalaron los aspectos del diseño a tener en cuenta para la implementación del AFE, aquí se retoman y discuten en profundidad.

adscripción- a los que están sometidos, su población objetivo fueron todos los investigadores vivos y con correo electrónico categorizados en el SNI, los vigentes y los no vigentes al año 2013.

La encuesta fue auto-administrada a través del cuestionario en línea. Los investigadores, en calidad de informantes, fueron informados de los objetivos de la encuesta e invitados a contestar el cuestionario por correo electrónico, que además incluía el acceso directo a la encuesta, el usuario y la clave de ingreso. El tiempo promedio para responder toda la encuesta fue estimado en 20 minutos aproximadamente.

El período de levantamiento de la encuesta fue del 1 de noviembre de 2013 al 31 de enero de 2014 para el caso de los investigadores categorizados en el SNI vivos, vigentes y con correo electrónico; y del 21 de enero de 2014 al 14 de abril del mismo año para el caso de los investigadores SNI no vigentes, vivos y con correo electrónico.

El cuestionario incluyó 41 preguntas cerradas<sup>59</sup>, 1 pregunta abierta, y un espacio para que los investigadores pudieran hacer sus observaciones. De las preguntas cerradas, 20 corresponden a la percepción de los investigadores sobre la evaluación que se realiza en el SNI, 18 corresponden a la percepción de los investigadores sobre otros mecanismos de evaluación (PROMEP e institución de adscripción) y las restantes 3 estuvieron dirigidas a los investigadores que se desempeñaron como evaluadores.

En el contexto de esta tesis, la población de interés son los investigadores categorizados en el SNI vigentes, vivos y con correo electrónico (en la siguiente sub-sección se detallan los aspectos vinculados a la muestra y su representatividad), y el objeto de estudio se centra en la percepción de los investigadores sobre las diversas dimensiones del proceso de evaluación del desempeño individual de investigación que se realiza en el SNI. En particular y con respecto a la evaluación que se realiza en el SNI, la encuesta indagó sobre:

- ✓ la relevancia de la información solicitada para la evaluación;
- ✓ la adecuación de los períodos de evaluación;
- ✓ la pertinencia de los mecanismos de designación de evaluadores, requisitos para ser evaluador e integración de comisiones dictaminadoras;
- ✓ la adecuación de los criterios de evaluación;
- ✓ los efectos generados por el SNI sobre la estrategia y desempeño como investigadores;
- ✓ los ámbitos y motivaciones de la producción de conocimiento; y
- ✓ la necesidad de reformar el mecanismo de evaluación y en qué direcciones.

Los datos recolectados a través de dicha encuesta y que se utilizan en este trabajo son de sección cruzada, es decir, refieren a las percepciones de los individuos –en este caso investigadores- en un determinado momento del tiempo. A la vez, el análisis que se realiza en esta investigación es de un solo grupo, es decir, para el conjunto de la población de los investigadores categorizados en el SNI de México.

---

<sup>59</sup> El cuestionario puede consultarse en el ANEXO V del documento *Reflexiones sobre la evaluación a los investigadores: una mirada desde diferentes perspectivas* (Bensusán et al., 2014), disponible en [www.foroconsultivo.org.mx/libros\\_editados/evaluacion\\_de\\_la\\_evaluacion\\_subgrupos\\_individuos.pdf](http://www.foroconsultivo.org.mx/libros_editados/evaluacion_de_la_evaluacion_subgrupos_individuos.pdf). Fecha de consulta 29/03/2017.

Desde el punto de vista de las propiedades métricas de los datos, es decir, el tipo de datos a analizar, importa señalar que se trata de datos categóricos (ordinales y dicotómicos). En la tabla 28 se detallan las 33 variables observadas de medida que se utilizan en el análisis, se describe su contenido, sus propiedades métricas y el conjunto de transformaciones que se le realizaron a cada una de ellas.

### VIII.3.3.2 Tamaño de la muestra: impacto y representatividad

El marco de muestreo de los investigadores del SNI vigentes fue construido a partir del directorio de investigadores del SNI del año 2013 de CONACYT<sup>60</sup>, identificando una población total de 19634 investigadores categorizados en el SNI vivos, vigentes y con correo electrónico al 2013.

De las 19634 invitaciones a los investigadores SNI vigentes, vivos y con correo electrónico, se obtuvo una tasa de respuesta del 41%. Es decir, 8055 investigadores SNI vigentes, vivos y con correo electrónico contestaron al menos una pregunta del cuestionario. Dicha tasa de respuesta da lugar a una muestra simple representativa del total de la población por nivel de categorización, área de conocimiento y género de los investigadores vigentes en el SNI (Bensusán et al., 2014). Importa señalar que esta amplia tasa de respuesta a la encuesta se asocia con su independencia en la medida que fue auspiciada/coordinada por el FCCyT, un organismo autónomo que por la Ley de Ciencia y Tecnología debe representar a las comunidades de CTI. En la tabla 6, que se presenta a continuación, se puede observar la similitud de las distribuciones poblacionales y muestrales según género, nivel de categorización y área de conocimiento.

Tabla 6: Distribución poblacional y muestral según género, nivel de categorización y área de conocimiento

Variable	Categorías	Distribución poblacional <sup>1</sup>		Distribución muestral	
		N	%	N	%
Género	Femenino	6702	34.1	2881	35.8
	Masculino	12932	65.9	5174	64.2
Nivel de categorización	Candidato	3695	18.8	1611	20.0
	Nivel I	10732	54.7	4504	55.9
	Nivel II	3548	18.1	1328	16.5
	Nivel III	1658	8.4	611	7.6
Área de conocimiento	Física, Matemáticas y Ciencias de la Tierra	3180	16.2	1226	15.2
	Biología y Química	3348	17.1	1594	19.8
	Medicina y Ciencias de la Salud	2023	10.3	837	10.4
	Humanidades y Ciencias de la Conducta	2891	14.7	1203	14.9
	Ciencias Sociales	2978	15.2	1142	14.2
	Biotecnología y Ciencias Agropecuarias	2315	11.8	1015	12.6
	Ingenierías	2899	14.8	1038	12.9
Institución	Centros Públicos de Investigación	2918	14.9	1141	14.2
	Empresas	70	0.4	15	0.2
	Gobierno	390	2.0	165	2.0

<sup>60</sup> La población total de investigadores SNI vigentes al año 2013, según el Directorio del CONACYT, era de 19747 investigadores. De ellos, 92 habían fallecido y 21 no contaban con correo electrónico al momento de la encuesta. (Bensusán et al., 2014).

	<i>Institutos de Salud</i>	1285	6.5	561	7.0
	<i>Otros</i>	671	3.4	220	2.7
	<i>Universidades Extranjeras</i>	521	2.7	142	1.8
	<i>Universidades Privadas</i>	718	3.7	266	3.3
	<i>Universidades Públicas</i>	13061	66.5	5545	68.8
<b>TOTAL</b>		<b>19634</b>	<b>100</b>	<b>8055</b>	<b>100</b>

Notas: 1. Población total de los investigadores vivos, vigentes y con correo electrónico categorizados en el SNI al año 2013.

Fuente: Elaboración propia en base a datos disponibles en Bensusán et al. (2014).

Además, de acuerdo con lo señalado por Bensusán et al. (2014), si la muestra obtenida fuera una muestra probabilística con un diseño muestral aleatorio simple y sin reemplazo, se obtendría – dados los tamaños de muestra alcanzados y una confianza del 95%- un error de muestreo que no supera  $\pm 5\%$  para estimaciones globales, por género, nivel de categorización y área de conocimiento. Las estimaciones para empresas no serían confiables y, para el resto de las categorías de la variable institución, el error de muestreo alcanza un máximo de  $\pm 7.5\%$ .

Por lo dicho hasta aquí, se concluye que la muestra no presenta sesgos y es representativa –por género, nivel de categorización y área de conocimiento- del conjunto de la población de los investigadores vigentes, vivos y con correo electrónico categorizados en el SNI mexicano (en adelante investigadores SNI).

Sin embargo, el mejor método y más sofisticado para el establecimiento del tamaño de muestra mínimo requerido para asegurar precisión y estabilidad de las soluciones de los modelos de ecuaciones estructurales –incluidos los modelos AFC- es a través de un estudio de simulación de Monte Carlo específico, como el propuesto por Muthén y Muthén (2002). En este trabajo, por simplicidad y dada la gran disponibilidad de información, se optó por utilizar las reglas generales disponibles en la literatura para evaluar si el tamaño de muestra disponible es suficiente para garantizar precisión y estabilidad de la solución del modelo de medida estimado.

De acuerdo con Hair et al. (2014) bajo ciertas *condiciones ideales*, tales como: i) normalidad multivariada de los datos; ii) sin datos faltantes; iii) modelos identificados o sobre-identificados con cinco o menos variables latentes; y iv) con cargas factoriales de 0.75 o más; el tamaño de muestra mínimo requerido para garantizar soluciones precisas y estables es de 100 a 200 observaciones. Conforme el estudio se aleja de esas condiciones ideales, se requiere mayor cantidad de observaciones. Muthén y Muthén (2002) muestran que el tamaño de muestra mínimo requerido para garantizar precisión y estabilidad de la solución del modelo de medida (AFC), cuando no se cumple la normalidad multivariada de los datos y existen datos faltantes, es de 315 para un *power* de 0.81. Adicionalmente, cuando los datos no cumplen con el supuesto de normalidad multivariada, Hair et al. (2014) proponen utilizar la regla general de 15 observaciones por parámetro libre a estimar en el modelo.

Esta investigación se aleja de esas condiciones ideales porque no se cumple el supuesto de normalidad multivariada de los datos y por la existencia de datos faltantes. Las otras condiciones ideales se cumplen de forma relativamente satisfactoria: i) el modelo que se propone está sobre-identificado; ii) tiene cinco variables latentes; y iii) las cargas factoriales estandarizadas estimadas alcanzan un valor promedio de 0.7 (ver sub-sección X.3.3.2 del capítulo X). Por lo tanto,

considerando el peor escenario: aquel que utiliza en el análisis sólo los casos completos, en este caso  $N=2283$ , se obtiene un total de 26 observaciones por parámetro libre a estimar en el modelo, con lo que se excede ampliamente la regla de 15 observaciones por parámetro libre a estimar. Si se considera, como se verá a continuación, el tratamiento de los datos faltantes que se realiza en este trabajo, el número de observaciones por parámetro libre a estimar aumenta a 75. En cualquiera de los dos escenarios, el tamaño de la muestra de la encuesta disponible para este trabajo es lo suficientemente grande para garantizar soluciones representativas, estables y precisas del modelo de medición.

#### *VIII.3.3.3 Datos faltantes: impacto y tratamiento*

En la tabla 21 se presenta el porcentaje de datos faltantes según cada variable observada utilizada en el análisis. Como se desprende de dicha tabla, mientras que 16 de 33 variables observadas tienen una proporción muy baja de datos faltantes (entre 2.5% y 8.1%), las restantes 17 variables tienen entre 14% y 37% de datos faltantes, lo que representa una proporción demasiado alta de información faltante. A la vez y de acuerdo con el análisis realizado, fue posible identificar un conjunto de 1292 patrones de datos faltantes. Es decir, en la base de datos se encontraron 1292 configuraciones diferentes de valores observados y faltantes. De acuerdo con Enders (2010) los patrones de datos faltantes son sólo una descripción de dónde están los “agujeros” en la base de datos y no proveen información sobre su naturaleza. Para aproximarse a la naturaleza de los datos faltantes, es preciso describir las posibles relaciones entre las variables observadas de medida y la probabilidad de datos faltantes, es decir, describir los *mecanismos de los datos faltantes* (Enders, 2010).

Enders (2010) con base en el trabajo seminal de Rubin (1976), describe tres mecanismos de los datos faltantes para explicar la relación entre los datos faltantes y los datos disponibles, a saber: i) datos faltantes completamente al azar (MCAR por sus siglas en inglés); ii) datos faltantes al azar (MAR por sus siglas en inglés); y iii) datos faltantes no al azar (MNAR por sus siglas en inglés). De acuerdo con Enders (2010), los datos faltantes son MCAR cuando la probabilidad de datos faltantes en una variable observada no puede ser explicada por ninguna otra variable observada del modelo ni por ella misma. Bajo este mecanismo, la probabilidad de datos faltantes es completamente al azar. Sin embargo, en la práctica este mecanismo es muy poco frecuente y sólo se puede alcanzar cuando los datos faltantes son planeados (Vargas, 2016). Los datos faltantes siguen un mecanismo MAR, cuando la probabilidad de datos faltantes en una variable observada puede ser explicada solo por otra(s) variable(s) observada(s) incluida(s) en el modelo. Los datos faltantes siguen un mecanismo MNAR, cuando la probabilidad de datos faltantes en una variable observada solo puede ser explicada por esa variable observada. Asparouhov y Muthén (2010) describen un mecanismo adicional que denominan datos faltantes al azar respecto a  $X$  (MARX). Aquí la probabilidad de datos faltantes en una variable observada puede ser explicada por una o más covariables ( $X$ ). Es decir, por la información contenida en variables observadas que no forman parte del modelo (exógenas) pero que son utilizadas como covariables (edad, género, nivel en el SNI, área de conocimiento, tipo de institución). Los mecanismos MAR y MARX son menos restrictivos que el mecanismo MCAR (Enders, 2010; Asparouhov y Muthén, 2010).

Las técnicas tradicionales de tratamiento de datos faltantes tales como: eliminar todas las observaciones (fila) con datos faltantes (*listwise deletion*) o utilizar todas las observaciones (celdas) con información disponible (*pairwise deletion*) requieren que los datos faltantes cumplan con el

mecanismo -más restrictivo- MCAR (Enders, 2010). Si los datos faltantes no cumplen con el mecanismo MCAR y se aplica alguna de esas técnicas tradicionales de tratamientos de datos faltantes, las estimaciones de los parámetros resultantes serán sesgadas (Enders, 2010), incluidos los errores estándar y las pruebas estadísticas (ej. de bondad de ajuste) (Brown, 2015). Incluso, aunque los datos faltantes siguieran un mecanismo MCAR, Enders (2010) no recomienda el uso de esas técnicas tradicionales, sino que por el contrario, sugiere la aplicación de las técnicas modernas de tratamiento de datos faltantes, tales como: estimación por máxima verosimilitud o imputación múltiple. Estas técnicas modernas solo requieren del supuesto menos restrictivo de que los datos faltantes sigan un mecanismo MAR (Enders, 2010) o MARX (Asparouhov y Muthén, 2010).

El problema práctico que surge es que no existe ningún procedimiento estadístico o prueba que permita comprobar si el mecanismo de los datos faltantes es MAR (incluido MARX) (Enders, 2010). Adicionalmente, en el marco conceptual del M-PLUS, cuando se modela variables observadas dependientes categóricas, la probabilidad de datos faltantes sólo pueden ser función de las covariables  $X$ , y no de las variables observadas dependientes (Muthén y Muthén 1998-2015). Es decir, en M-PLUS cuando se modela variables observadas categóricas, los datos faltantes pueden ser predichos a través del mecanismo MARX (Asparouhov y Muthén, 2010). Similarmente, ante la imposibilidad de comprobar estadísticamente si los datos faltantes siguen un mecanismo MAR, Enders (2010) sugiere seguir una *estrategia inclusiva* para hacer más plausible que los datos faltantes sigan un mecanismo MAR. Dicha estrategia supone la inclusión de covariables  $X$  en el modelo, que puedan predecir la probabilidad de datos faltantes.

En el contexto de este trabajo, el tratamiento de los datos faltantes que se realiza sigue las recomendaciones de Enders (2010). En particular, se aplica una de las técnicas modernas de tratamiento de datos faltantes: la estimación por máxima verosimilitud con información completa (FILM por sus siglas en inglés) (Arbuckle, 1996; Asparouhov y Muthén, 2010). Es decir, se utiliza toda la información –la completa y la faltante- y los datos faltantes se estiman durante el proceso de estimación. Sin embargo, para implementar el enfoque FILM para el tratamiento de los datos faltantes y asegurar estimaciones de los parámetros eficientes e insesgadas, es preciso incorporar en el modelo un conjunto de covariables  $X$  que puedan predecir la probabilidad de datos faltantes. En otras palabras, hay que hacer plausible que los datos faltantes sigan un mecanismo MARX, para ello hay que seguir una estrategia inclusiva e identificar las covariables  $X$  que predicen los datos faltantes. Las covariables  $X$  identificadas que hacen plausible que los datos faltantes sigan un mecanismo MARX son: *edad, género, nivel SNI, área de conocimiento y tipo de institución*. En el Anexo 7 se presentan los resultados de los modelos logit estimados para explicar la probabilidad de datos faltantes en función de las covariables  $X$ . Finalmente, cabe mencionar que las covariables  $X$  identificadas se incorporan en la estimación del modelo estructural.

#### *VIII.3.3.4 Estimación: Estimador y paquete estadístico*

De acuerdo con Kline (2011) entre otros, cuando las variables observadas dependientes (endógenas) no son continuas y por lo tanto no se cumplen los supuestos de la teoría normal, es preciso aplicar otros métodos de estimación distintos del método de estimación de máxima verosimilitud.

En el contexto de este trabajo y debido a las propiedades métricas de las variables observadas dependientes, las estimaciones se realizan con el estimador mínimos cuadrados ponderados (WLS por sus siglas en inglés) robustos desarrollado por Muthén (1993). En particular, se utiliza el

estimador WLSMV que provee estimaciones WLS de los parámetros –utilizando solo la diagonal de la matriz de pesos  $W$ -, errores estándar robustos y media y varianza ajustada del estadístico de prueba Ji-Cuadrado (Brown, 2015; Muthén y Muthén, 1998-2015). Para una descripción detallada de los estimadores WLS robustos, sus respectivas funciones de ajuste y desempeño ver Muthén, du Toit y Spisic (1997).<sup>61</sup>

Las estimaciones se realizaron con el paquete estadístico M-PLUS versión 7.4 (Muthén y Muthén, 1998-2015). Este programa estadístico está diseñado específicamente para el modelado con variables latentes y, de acuerdo con Brown (2015), es el que ofrece las mejores opciones para la estimación de modelos AFC con variables categóricas.

#### VIII.3.4 Evaluación de la Validez del Modelo de Medida

Con el modelo de medida especificado y establecidos todos los aspectos del diseño, se puede proceder con su estimación y la evaluación sobre su validez. Los resultados obtenidos de la estimación se pueden observar en la sub-sección X.3.3 del capítulo X, aquí la atención se concentra en las pruebas que se realizan para evaluar la validez del modelo de medida propuesto.

De acuerdo con Hair et al. (2014), evaluar la validez del modelo de medida supone: i) analizar y comparar la teoría de medida propuesta en el modelo contra la realidad, representada por los datos; y, ii) estudiar la validez de constructo del modelo. Dicha evaluación provee evidencia para responder si el modelo es exacto y si efectivamente mide lo que busca medir.

En ese sentido, la validez del modelo de medida depende de dos evaluaciones, la que se realiza sobre la bondad de ajuste, y la que se hace sobre la validez de constructo. El AFC provee evidencia que permite confirmar si el modelo de medida teórico propuesto explica la realidad y si es válido.

##### VIII.3.4.1 Evaluación de la Bondad de Ajuste del Modelo de Medición

La bondad de ajuste indica en qué medida el modelo teórico especificado reproduce la matriz de correlación asintótica de las variables de respuesta latente. Se trata de observar la similitud –o diferencia- entre dicha matriz y la estimada. Cuanto menor sea la diferencia entre esas matrices, mayor será la bondad de ajuste del modelo de medida teórico propuesto con los datos empíricos.

Existen diversas medidas de bondad de ajuste que pueden clasificarse en medidas absolutas o incrementales, a continuación se describen las utilizadas en este trabajo.

La primera medida absoluta es el estadístico de prueba Ji-Cuadrado de bondad de ajuste ( $\chi^2$ ), sus grados de libertad ( $df$ ) y el tamaño de la muestra ( $N$ ). Con este estadístico, la hipótesis nula que se está poniendo a prueba es si las matrices de correlación observada – que es la matriz de correlación asintótica de las variables de respuesta latente- ( $S$ ) y estimada ( $\Sigma_k$ ) son iguales. Si el estadístico Ji-Cuadrado es significativo (cuando la probabilidad de que ambas matrices sean iguales en una población determinada el valor de  $p$  es bajo), entonces provee evidencia de que hay problemas con el ajuste. Sin embargo, es preciso notar la sensibilidad del estadístico de prueba Ji-Cuadrado a los grados de libertad del modelo y al tamaño de la muestra. (Hair et al., 2014; Klein 2011; Bollen, 1989).

---

<sup>61</sup> Para una discusión sobre las diferencias entre los estimadores WLS y WLS robustos ver Finney y Distefano (2006).

Sin embargo, hay un problema cuando el tamaño de muestra es grande. En ese contexto, se espera que el valor de probabilidad de la prueba Ji-cuadrado de bondad de ajuste detecte diferencias significativas entre dichas matrices cuando en realidad no existen. Este problema está ampliamente reconocido en la literatura, es por eso que se han desarrollado otras medidas alternativas y complementarias de bondad de ajuste, que –justamente- buscan corregir la sensibilidad al tamaño de la muestra y a los grados de libertad del modelo.

Así, la segunda medida absoluta es la *Raíz Cuadrada del Error Cuadrático Medio* (**RMSEA** por sus siglas en inglés). Esta es una medida adicional de ajuste absoluto y corrige tanto por la complejidad del modelo como por el tamaño de la muestra. De acuerdo con la literatura, valores bajos de RMSEA indican buen ajuste. Los valores de referencia de corte para el RMSEA son debatibles, mientras que algunos autores sostienen que valores de RMSEA menores a 0.05 son deseables; bajo ciertas condiciones, valores menores a 0.08 son indicadores de buen ajuste (Rigdon, 1996). Hair et al. (2014) y Bollen y Long (1993) señalan que cuando se analizan modelos con 30 variables observadas o más –como en el caso del modelo que se propone en este trabajo- y con un tamaño de muestra mayor a 250 observaciones, valores de RMSEA menores a 0.07 son esperables e indican buen ajuste.

La tercera medida de ajuste es el *Índice Tucker-Lewis* (**TLI** por sus siglas en inglés), que es una medida de ajuste relativo o incremental. A diferencia de las medidas de ajuste absolutas, las relativas o incrementales evalúan en qué medida el modelo estimado ajusta mejor en relación a algún modelo alternativo de base (o *modelo nulo*) (Hair et al., 2014). El *modelo nulo* asume que todas las variables observadas son independientes (no están correlacionadas). Así las medidas de ajuste relativo o incremental representan la mejora del ajuste debido a la especificación del modelo teórico realizada. En general, valores de TLI cercanos a 1 indican buen ajuste. De acuerdo con Hair et al. (2014) y Bollen y Long (1993), valores del TLI mayores a 0.90 indican buen ajuste.

La cuarta medida de ajuste es el *Índice de Ajuste Comparado* (**CFI** por sus siglas en inglés), que es –al igual que el TLI- una medida de ajuste relativo, y también valores del CFI mayores a 0.90 son asociados con modelos que ajustan satisfactoriamente (Hair et al., 2014; Bollen y Long, 1993).

Finalmente, el estadístico Ji-cuadrado también puede ser utilizado para comparar modelos. Es posible probar la significancia estadística de la diferencia en el valor de la Ji-cuadrado entre dos modelos. Se trata de la prueba de la diferencia de la Ji-Cuadrado para modelos anidados (Brown, 2015). El objetivo con esta prueba, en el contexto de este trabajo, es demostrar que el modelo de medida multifactorial de cinco factores propuesto ajusta significativamente mejor que un modelo de medida unidimensional (de un solo factor). Para ello, se realiza un procedimiento en dos etapas porque la diferencia de los valores de la Ji-cuadrado entre modelos anidados, que son estimados con WLSMV, no se distribuye como una Ji-cuadrado (Brown, 2015; Muthén y Muthén, 1998-2015; Muthén, 1998-2004). En la primera etapa se estima el modelo de medida multifactorial de cinco factores que se propone, es decir, el que cuenta con menor cantidad de restricciones ( $H_1$ ), y sus resultados se almacenan. En la segunda etapa, se estima el modelo más restringido ( $H_0$ ) -el que supone que la relación entre la evaluación de la investigación y la producción de conocimiento puede ser medida por un solo factor-, y se utilizan todos los resultados, incluidos los almacenados en la etapa 1, para calcular el valor de la prueba de la diferencia de la Ji-Cuadrado. Para una revisión formal, técnica y detallada de los cálculos y procedimientos para hallar el valor de la prueba de la diferencia de la Ji-Cuadrado ver Muthén (1998-2004)



#### VIII.3.4.2 Evaluación de la Validez de Constructo del Modelo de Medida

Tal como se ha señalado, uno de los principales objetivos del AFC -en el contexto de los modelos de ecuaciones estructurales-, es evaluar la validez de constructo de la teoría de medida que se está poniendo a prueba.

La *validez de constructo* se define como el grado en el que un conjunto de ítems de medida representan *verdaderamente* a los constructos latentes teóricos, que esos ítems buscan para medir (Hair, et al.2014: 618). Se trata de evaluar la precisión y en qué magnitud las medidas son exactas. En ese sentido, si existe evidencia que sugiere validez de constructo, se puede afirmar con confianza que los ítems de medida –observados en una muestra representativa- representan los valores verdaderos que existen en toda la población. A continuación se describen las cuatro pruebas de validación que se realizan en este trabajo.

De acuerdo con Hair et al. (2014) la validez de constructo del modelo de medida consiste en cuatro pruebas de validación.

La primera se denomina prueba de *validez aparente/contenido*: con esta prueba, de carácter cualitativo, se busca validar el contenido de los constructos. Para ello, un panel externo de expertos analizan el grado de coincidencia entre las definiciones conceptuales de cada una de los constructos (variables latentes) y las variables observadas indicadoras (ítems) que los operacionalizan. En el contexto de esta investigación, el panel de expertos externos está integrado por los miembros del tribunal de esta tesis, quienes en el marco del coloquio realizaron una primera evaluación de la validez aparente del modelo de medida propuesto. Dicha evaluación, dio como resultado que el contenido de los constructos es válido. Es decir, los evaluadores externos encontraron un alto grado de coincidencia entre las definiciones conceptuales de los constructos y los ítems que los componen.

La segunda se denomina prueba de *validez convergente*: con esta prueba, de carácter cuantitativo, se busca analizar la convergencia de los ítems, que son indicadores de constructos específicos. En otras palabras, los ítems de cada uno de los constructos deben compartir una proporción relativamente alta de varianza común. Cuanto mayor sea la varianza común compartida, mayor será la capacidad del constructo para explicar la variación individual y conjunta de esos ítems. De acuerdo con Hair et al. (2014), la prueba de validez convergente supone el análisis conjunto y complementario de:

- ✓ las *cargas factoriales estimadas no estandarizadas*: si son estadísticamente significativas, entonces hay evidencia que sugiere validez convergente;
- ✓ las *cargas factoriales estimadas estandarizadas*: si son estadísticamente significativas, lo suficientemente grandes y su signo es el esperado -de acuerdo a las definiciones conceptuales del constructo-, entonces hay evidencia que sugiere validez convergente<sup>62</sup>;

---

<sup>62</sup> La racionalidad detrás de esta guía general se asocia al concepto de *comunalidad*, varianza compartida entre los ítems o varianza extraída (Hair et al., 2014). El cuadrado de una carga factorial estandarizada representa en qué medida la variación de un ítem ( $y^*$ ) es explicada por el factor –varianza extraída- (Brown, 2015). A modo de ejemplo: una carga factorial estandarizada de 0.7 ( $\lambda = 0.7$ ) y significativa indica que ese factor está explicando aproximadamente el 50% de la varianza de ese ítem ( $y^*$ ), el restante 50% -que no es explicado por el factor- se contabiliza en la varianza residual.

- ✓ la *Varianza Extraída Promedio* (AVE por sus siglas en inglés): es un indicador resumen de convergencia. Se calcula como el promedio de la varianza extraída de los ítems por cada uno de los constructos latentes, y representa la comunalidad promedio. En el contexto del AFC, su cálculo está dado por:

$$AVE = \frac{\sum_{j=1}^p \lambda_j^2}{p}$$

Donde  $\lambda_j$  es la carga factorial estandarizada correspondiente al ítem  $j$  y  $p$  es el número total de ítems indicadores del constructo latente. Los valores de corte de referencia son discutibles. Algunos autores sostienen que un valor  $AVE \geq 0.5$  sugiere una convergencia adecuada, para otros alcanza con un valor  $AVE \geq 0.3$ . En el contexto de investigaciones exploratorias y en las ciencias sociales, valores mayores o iguales a 0.3 son considerados aceptables y sugieren buena convergencia; y,

- ✓ la *Confiabilidad Compuesta* (CR por sus siglas en inglés) (Dillon y Goldstein, 1984): también es una medida resumen de la validez de convergencia; y si bien hay debate en la literatura sobre cuál de todos los indicadores de confiabilidad es el más adecuado, y aunque el Coeficiente Alpha de Cronbach<sup>63</sup> es el más utilizado, en el contexto de este trabajo se optó por seleccionar el indicador CR porque es el recomendado en el contexto de los modelos de ecuaciones estructurales. Su cálculo está dado por:

$$CR = \frac{\left(\sum_{j=1}^p \lambda_j\right)^2}{\left(\sum_{j=1}^p \lambda_j\right)^2 + \sum_{j=1}^p \theta_j}$$

Donde  $\left(\sum_{j=1}^p \lambda_j\right)^2$  es el cuadrado de la sumatoria de las cargas factoriales estandarizadas de cada constructo y  $\sum_{j=1}^p \theta_j$  la sumatoria de la varianza de los términos de error o varianzas residuales de los ítems de cada uno de los constructos. Valores  $CR \geq 0.7$  sugieren una buena confiabilidad. Valores altos de CR sugieren consistencia interna, es decir, las variables observadas de medida consistentemente representan al mismo constructo.

La tercera se denomina prueba de validez discriminante: con esta prueba, de carácter cuantitativo, se busca analizar en qué medida un constructo es en realidad diferente de otro. La validez discriminante sugiere que el constructo es único y que captura algún fenómeno que otras medidas no, además también significa que los ítems de medida individuales representan a un único constructo. Si el AVE estimado es mayor que el cuadrado de la correlación estimada entre dos constructos o variables latentes, entonces hay evidencia que sugiere validez discriminante. La racionalidad es que si un constructo latente explica más de la variación de sus ítems (AVE) que lo que comparte con otro constructo (correlación al cuadrado), entonces existe evidencia que sugiere validez discriminante. Dicho de otro modo, los constructos discriminan entre sí.

---

<sup>63</sup> El Coeficiente Alpha de Cronbach provee resultados similares a la confiabilidad compuesta y también debe ser mayor que 0.7 (ver Nunnally y Bernstein, 1994).

Finalmente, la cuarta se denomina prueba de *validez nomológica*: con esta prueba, de carácter cuantitativo y conceptual, se estudian las correlaciones entre los constructos y se analiza su sentido conceptual. Si las correlaciones estimadas entre los constructos son consistentes con lo que se espera de acuerdo con la teoría, entonces existe evidencia que sugiere validez nomológica.

En la sub-sección X.3.3 del capítulo X se presentan todos los resultados derivados de la evaluación de la validez del modelo de medida.

#### VIII.4 Etapa 3: Explicación de la relación entre la evaluación de la investigación y la producción de conocimiento, el modelo de ecuaciones estructurales

A esta altura del proceso de investigación ya se tiene un modelo de medida válido y confiable, por lo que el objetivo de esta etapa es poner a prueba y validar el modelo teórico estructural que se propone para explicar la relación entre la evaluación de la investigación del desempeño individual de los investigadores y las formas diferenciales de producción de conocimiento.

En esta etapa, la atención se concentra en las relaciones teóricas/estructurales entre los constructos que se propone. En otras palabras, se contrastan el conjunto de hipótesis de trabajo orientadoras de esta investigación. Es decir, se pone a prueba y continúa con la validación de la teoría estructural que sustenta este trabajo. A continuación se describen y fundamentan todos los procedimientos que se realizan en esta investigación para alcanzar el objetivo de esta etapa.

##### VIII.4.1 Procedimientos y consideraciones para la especificación del modelo teórico estructural

La validez del modelo estructural que se propone depende del i) marco teórico propuesto, ii) de la validez y confiabilidad del modelo de medida; iii) de la bondad de ajuste del modelo teórico estructural; y iv) de la naturaleza -magnitud, significancia estadística, dirección y, por lo tanto, sentido conceptual- de los parámetros estructurales estimados.

Tal como se señaló, en esta etapa de la investigación ya se alcanzó un modelo de medida válido y confiable. Para avanzar en la validez de la teoría estructural que se propone -es decir, para poder realizar la evaluación de la bondad de ajuste del modelo teórico estructural y de la naturaleza de los parámetros estructurales estimados- es preciso en primer lugar especificar el modelo teórico estructural que se contrasta con la realidad y valida.

La especificación del modelo estructural teórico para evaluar la validez de las relaciones conceptuales propuestas para explicar teóricamente la relación entre la evaluación de la investigación individual y las formas diferenciales de producción de conocimiento en el contexto del SNI mexicano, supone un conjunto de transformaciones teóricas y de notación al modelo de medición desarrollado. A continuación se desarrollan estos aspectos.

El modelo estructural que se propone se basa en la teoría, el debate internacional y los hallazgos de investigaciones pasadas reportados en la literatura sobre las relaciones entre la evaluación del desempeño de la investigación y la dirección, contenido y características epistemológicas de la producción de conocimiento (ver capítulo V). Esas relaciones, en el contexto de esta investigación se exploran a través de los cinco constructos definidos en el marco teórico conceptual (capítulo V) y que se retoman en la tabla 28 de la sub-sección X.3.1 del capítulo X.

A diferencia del modelo AFC, donde la atención se enfocó en la relación entre las variables manifiestas ( $y^*$ ) y las variables latentes –constructos-, en el modelo teórico estructural el foco se pone sobre las relaciones entre las variables latentes. Ya no se trata de correlaciones entre las variables latentes –como en el AFC- sino relaciones de dependencia. Cada relación de dependencia entre las variables latentes da cuenta de las hipótesis de investigación de este trabajo y se corresponden con los parámetros libres a estimar en el modelo estructural; su representación gráfica se hace a través de senderos que muestran la relación entre variables latentes. Todas las relaciones son puestas a prueba simultáneamente en el mismo modelo estructural.

Los constructos que no están unidos por ningún sendero refiere a parámetros fijos (el parámetro se fija en 0) y significa que no hay una relación de dependencia directa entre esos conceptos en la teoría que se está poniendo a prueba. Adicionalmente, mientras que en el modelo de medida los cinco constructos son exógenos, en el modelo estructural algunos constructos son exógenos y otros son endógenos.

De hecho, la teoría que se pone a prueba estudia el efecto de los constructos exógenos (predictores) sobre los endógenos (resultado). Simultáneamente, un constructo puede ser endógeno (resultado) en una relación de dependencia y en la siguiente puede ser exógeno (predictor). Los constructos endógenos no son explicados completamente, por lo que se les asocia un término de error o varianza residual ( $\xi$ ). Así el modelo estructural considera los errores de medición (varianzas residuales de los ítems) y también el error de predicción teórico ( $\xi$ ). Los parámetros estimados se calculan de acuerdo a los efectos directos e indirectos en el modelo, y especificando relaciones de dependencia se establecen las restricciones sobre qué efectos se utilizan para la estimación.

Desde el punto de vista de la identificabilidad del modelo teórico estructural propuesto, cabe señalar que sus grados de libertad se calculan de la misma manera que en el modelo de medida. Sin embargo, en el modelo estructural el número de parámetros libres a estimar es menor que en el modelo de medida porque no se incluyen todas las relaciones estructurales entre los constructos. Se fijan en cero tres relaciones estructurales, a saber: i) la relación entre propensión modo 1 y dirección de la reforma del sistema de evaluación (prop\_m1 y dirref); ii) la relación entre propensión modo 2 y dirección de la reforma del sistema de evaluación (prop\_m2 y dirref); y, iii) la relación entre flexibilidad epistemológicas del sistema de evaluación y dirección de la reforma del sistema de evaluación (flxepsni y dirref). Es así que en lugar de estimar 106 parámetros -como se hizo en el modelo de medida-, ahora se estiman 103 parámetros.

El modelo teórico estructural que se propone está anidado dentro del modelo de medida (AFC) y es más parsimonioso porque contiene menor cantidad de parámetros libres a estimar con el mismo número de constructos (5) y de variables observadas de medida (33).

Además se trata de un modelo teórico estructural *recursivo*, porque todos los senderos entre constructos provienen solo del constructo predictor hacia otro constructo de resultado, no hay retroalimentación entre constructos (Hair et al., 2014). De acuerdo con este autor, los modelos estructurales recursivos nunca pueden tener menor cantidad de grados de libertad que el AFC cuando se utilizan la misma cantidad de constructos y variables observadas de medida.

Con datos de corte transversal, como los que se tienen disponibles para este estudio, no es plausible desarrollar un modelo estructural no recursivo porque es difícil sino imposible re-crear el conjunto

de condiciones necesarias para establecer relaciones recíprocas entre constructos, porque para ello es preciso considerar la secuencia temporal de los eventos (se requieren datos longitudinales) (Kline, 2011; Hair et al., 2014).

Dado que el modelo teórico estructural propuesto es un modelo recursivo, no contiene términos de interacción, el tamaño de la muestra es lo suficientemente grande (N=8053), cuenta con al menos tres variables observadas de medida por constructo, el modelo de medida está identificado y se estima un número menor de parámetros que en el AFC, es posible concluir que el modelo teórico estructural propuesto también está identificado. De hecho, y al igual que el modelo de medida, está sobre-identificado.

Si en el modelo teórico estructural se especificaran la misma cantidad de relaciones estructurales que relaciones de correlación entre constructos del modelo de medida, se obtendría el *modelo estructural saturado*. De acuerdo con Hair et al. (2014) los modelos estructurales saturados generalmente no representan demasiado interés porque usualmente no revelan más información que la obtenida con el análisis AFC. Las medidas de bondad de ajuste del modelo estructural saturado deberían ser las mismas que las obtenidas en el modelo de medida (AFC) (Hair et al., 2014). En el contexto de este trabajo se calculan y comparan dichas medidas de bondad de ajuste contra las obtenidas con el AFC para corroborar que no se cometió ningún error en la transformación desde el modelo de medida hacia el estructural.

Los restantes aspectos del diseño del estudio son consistentes con los planteados para el modelo de medida (AFC) (ver sub-sección VIII.3.3 de este capítulo).

En la sub-sección X.4.1 del capítulo X se especifica el modelo teórico estructural propuesto en este trabajo –se formaliza matemáticamente y se representa gráficamente a través de un diagrama de senderos- (figura 5).

Con el modelo teórico estructural especificado y fundamentado teóricamente, se puede proceder con su estimación, y a partir de esos resultados continuar con el proceso de validación del modelo teórico estructural propuesto.

#### VIII.4.2 Evaluación de la Validez del Modelo Teórico Estructural

Como se señaló, la validez del modelo teórico estructural depende de la validez y confiabilidad del modelo de medición, de la bondad de ajuste del modelo teórico estructural propuesto y de la naturaleza de los parámetros estructurales estimados.

Dado que a esta altura de la investigación ya se logró una teoría de medición confiable y válida, representada en el modelo de medida, a continuación y en primer lugar, se describen los procedimientos para analizar la bondad de ajuste del modelo estructural para evaluar el grado de aceptación del modelo teórico propuesto. En segundo lugar, se estudian los parámetros estructurales estimados (los senderos o relaciones hipotetizadas). Se trata del análisis de la magnitud, significancia estadística, dirección e interpretación conceptual de los parámetros estructurales estimados.

Si el modelo estructural presenta buen ajuste y los parámetros estructurales son significativos y en las direcciones hipotetizadas, entonces se puede concluir que el modelo teórico estructural es sostenido por la evidencia. Estos resultados junto con la validez del modelo de medida representan

evidencia fuerte para validar la teoría estructural que se propone para explicar la relación entre la evaluación del desempeño de investigación individual y las formas diferenciales que predominantemente orientan la producción de conocimiento de los investigadores en el contexto del SNI mexicano.

#### *VIII.4.2.1 Evaluación de la Bondad de Ajuste del Modelo Teórico Estructural*

Al igual que en el análisis de la bondad de ajuste del modelo de medida, la evaluación de la bondad de ajuste del modelo teórico estructural implica analizar en qué medida el modelo teórico estructural propuesto reproduce la matriz de correlación de las *variables de respuesta latente* –S–.

Sin embargo, y debido a que se trata de una teoría estructural recursiva, el modelo teórico estructural –recursivo– está anidado en el modelo de medida y por lo tanto no puede contener mayor cantidad de relaciones estructurales entre constructos que las correlaciones entre ellos posibles en el modelo de medida (Hair et al., 2014). Dicho de otro modo, el modelo estructural tiene mayor cantidad de grados de libertad que el modelo de medida –porque se estiman menor cantidad de relaciones entre los constructos–, y dado que todo lo demás se deja constante (el tamaño de la muestra, la cantidad de constructos y de ítems), el valor del estadístico Ji-Cuadrado nunca puede ser menor que el valor obtenido en el modelo de medida. Es decir, el modelo estructural no va a presentar un ajuste mejor que el modelo de medida.

En este sentido, la estrategia para evaluar la bondad de ajuste del modelo estructural implica comparar las medidas de bondad de ajuste del modelo estructural contra las obtenidas en el modelo de medida para analizar en qué medida el ajuste se degrada debido a la especificación de las relaciones estructurales.

Las medidas de bondad de ajuste que se analizan aquí son las mismas que se analizaron en el modelo de medida (Ji-Cuadrado, RMSEA, CFI, TLI), con los mismos valores de referencia de corte. Si el ajuste del modelo estructural es sustantivamente peor que el del modelo de medida, entonces se puede concluir que la teoría estructural carece de validez. (Hair et al., 2014)

Adicionalmente, otra medida para evaluar la bondad de ajuste del modelo estructural respecto al ajuste del modelo de medida es: la diferencia del estadístico Ji-Cuadrado entre esos dos modelos. El criterio indica que si esa diferencia es significativa, entonces el ajuste empeora significativamente (Hair et al., 2014). Sin embargo, Kline (2011:219) señala que con modelos no jerárquicos, de un solo nivel como el que se propone en este estudio, esa comparación se puede realizar pero la prueba estadística de la diferencia de la Ji-Cuadrado no puede interpretarse estadísticamente. Además de esto, y dado el efecto del tamaño de la muestra de este estudio sobre el valor de la Ji-Cuadrado, es probable que dicha diferencia sea significativa, cuando en realidad el ajuste es aceptable. En la subsección X.4.2.1 del capítulo X se presentan estos resultados.

#### *VIII.4.2.2 Evaluación sobre la naturaleza de los parámetros estructurales estimados*

La evaluación de la validez de la teoría estructural que se propone requiere, como se mencionó, una evaluación sobre la naturaleza de los parámetros a estimar en el modelo estructural.

De todas maneras y antes de estudiar la naturaleza de los parámetros del modelo estructural, es preciso analizar la estabilidad y confiabilidad de los parámetros. Para ello se comparan los parámetros estimados en el modelo teórico estructural propuesto y las cargas factoriales obtenidas

en la estimación del modelo de medida. El objetivo es analizar si existe evidencia que sugiera *confusión interpretativa* (“*interpretational confounding*”) (Kline, 2011; Hair et al., 2014).

Si hay evidencia que sugiera confusión interpretativa, significa que las estimaciones de las relaciones de medida en el modelo estructural para los distintos constructos son afectadas sustantivamente más por las relaciones estructurales especificadas en el modelo estructural que por las variables de medida específicas. Esto representa un problema y atenta contra la estabilidad y confiabilidad de la medición (genera confusión o dudas respecto al significado de los constructos), básicamente porque las cargas factoriales de los constructos –estimadas en el modelo de medida- y los parámetros de medición –estimados en el modelo estructural- deberían ser similares y no variar demasiado por los cambios que se realicen en la especificación del modelo estructural. De acuerdo con Hair et al. (2014) variaciones de 0.05 en valor absoluto o menores en la comparación de esos parámetros, son aceptables y no representan un problema de confusión interpretativa. En la sub-sección X.4.2.2 del capítulo X se presentan y discuten los resultados derivados de este análisis.

Luego de analizada la estabilidad de los parámetros, se procede –si los resultados obtenidos dan cuenta de la estabilidad de las estimaciones de medida- con el estudio de los parámetros estimados del modelo estructural.

Si dichos parámetros estructurales son estadísticamente significativos, sus signos son consistentes con las direcciones hipotetizadas (relaciones positivas o negativas) de acuerdo con la teoría y sus magnitudes son lo suficientemente grandes, entonces existe evidencia que apoya el modelo teórico propuesto (la teoría estructural) y además aumenta su validez.

La prueba estadística para la significancia de los parámetros asume una distribución normal en muestras grandes. El valor crítico para la prueba a dos colas –con un nivel alpha de 0.05- es en valor absoluto mayor que 1.96. (Muthén y Muthén, 1998-2015)

Analizar la magnitud de los parámetros estimados es importante porque si dicha magnitud es demasiado pequeña podría sugerir una alta probabilidad de estar cometiendo un “Error Tipo I” y con ello, cuestionar la validez de las relaciones teóricas estructurales de dependencia que se están estudiando.

Por esa razón, es preciso recordar que los resultados derivados de las pruebas de significación dependen de la magnitud de las relaciones (de los parámetros), del tamaño de la muestra y de la eficiencia del diseño estadístico de la investigación (Kish, 2004).

En ese sentido, la probabilidad de rechazar una hipótesis nula que en realidad es cierta –“Error Tipo I”- aumenta, teniendo en cuenta que: i) el tamaño muestral utilizado en esta investigación es muy grande (N=8053); ii) se trata de una investigación en el área de las ciencias sociales, lo que implica gran complejidad por la existencia de relaciones multivariantes, que simultáneamente dificultan el hallazgo de relaciones de independencia (parámetros no significativos) entre las variables de interés; y iii) existe una eficiencia del diseño estadístico relativamente menor debido, por un lado, a la especificación realizada (no se introducen variables de control para aislar potenciales efectos perturbadores de variables extrañas, que se pueden confundir con los efectos de las variables de interés), y por otro, a la aplicación del enfoque *pair wise deletion* para el tratamiento de los datos faltantes.

Es así que para minimizar la probabilidad de Error Tipo I, es decir, la probabilidad de que relaciones que en realidad son insignificantes aparezcan como estadísticamente significativas, en la siguiente sub-sección se describen los procedimientos para la especificación y validación del modelo teórico estructural condicional. Esto permitirá aumentar la eficiencia del diseño, reducir ciertos sesgos, y con ello, una evaluación más precisa y confiable de todas las relaciones estructurales (hipótesis) que se están poniendo a prueba (contrastando con la *realidad* representada por los datos).

#### VIII.4.3 Modelo teórico estructural condicional

El objetivo de especificar y estimar el modelo teórico estructural condicional es doble. Por un lado se trata de controlar los resultados que se derivan del modelo teórico estructural, aislando el efecto potencial que podrían tener -sobre las variables explicadas- ciertas variables perturbadoras. Por otro lado, se busca aplicar un tratamiento a los datos faltantes que contribuya a la obtención de estimaciones de los parámetros eficientes e insesgadas.

A diferencia del modelo teórico estructural sin condicionar, aquí se siguen las recomendaciones de Enders (2010) para asegurar estimaciones de los parámetros eficientes e insesgadas. En ese sentido, se aplica la estimación por máxima verosimilitud con información completa (FILM por sus siglas en inglés) (Asparouhov y Muthén, 2010) para el tratamiento de los datos faltantes.

Para ello, se incorporan en el modelo 19 covariables  $X$  de control -como predictoras de las variables latentes independientes propensión modo 1 y propensión modo 2-, útiles para: i) aumentar la eficiencia reduciendo los errores aleatorios y disminuyendo el sesgo de las variables perturbadoras, cuyos efectos pueden confundirse -si no se controlan- con los efectos de las variables explicativas (Kish, 2004); y, ii) predecir la probabilidad de datos faltantes, haciendo plausible que los datos faltantes sigan un mecanismo MARX (Asparouhov y Muthén, 2010).

Las covariables  $X$  de control, que condicionan el modelo, identificadas son: *edad*, *género*, *nivel SNI*, *área de conocimiento* y *tipo de institución* (ver tabla 41 de la sub-sección X.5.1 del capítulo X). Para identificar dichas variables de control, se estimaron un conjunto de modelos logit para explicar la probabilidad de datos faltantes de cada una de las variables observadas. En el Anexo 7 se presentan las estimaciones realizadas.

El objetivo en última instancia es aumentar la validez de la teoría que se contrasta y representa a través del modelo teórico estructural condicional.

Los procedimientos que se siguen en esta etapa de la investigación para validar el modelo teórico estructural condicional son esencialmente los mismos a los seguidos con el modelo teórico estructural, a saber: i) se especifica formalmente y representa gráficamente el modelo teórico estructural condicional y se analizan los aspectos vinculados con su identificabilidad; ii) se examina su bondad de ajuste; y iii) se analizan y comparan los parámetros estructurales del modelo teórico estructural condicional en relación al no condicional. Finalmente, se concluye sobre la validez de la teoría que se está poniendo a prueba. En las sub-secciones X.5.1, X.5.2 y X.5.3 del capítulo X se especifica formalmente el modelo teórico estructural condicional, se evalúa su bondad de ajuste y se estudia la naturaleza de los parámetros estructurales estimados del modelo estructural condicional.



## CAPÍTULO IX: Análisis Descriptivo

El objetivo de este capítulo es describir estadística y analíticamente las percepciones que tienen los investigadores categorizados en el SNI sobre: i) los ámbitos y motivaciones de la producción de conocimiento que realizan (sus propensiones modo 1 y modo 2 de producción de conocimiento); ii) la satisfacción que tienen con el proceso de evaluación del SNI; iii) la flexibilidad epistemológica que perciben del sistema de evaluación al que están sometidos; y iv) la dirección de la reforma del sistema de evaluación, de acuerdo con algunas características específicas de los investigadores categorizados en el Sistema.

Para ello, en la siguiente sección se realiza una caracterización de los investigadores del SNI de acuerdo a su: i) perfil demográfico (edad y género); ii) desempeño en el SNI (nivel de categorización, ascenso en el Sistema y si han participado como evaluadores en el SNI o en otras instancias de evaluación de la investigación); iii) pertenencia cognitiva y laboral (área de conocimiento y tipo de institución/organización en la que desempeñan sus actividades los investigadores); iv) percepción sobre los efectos generados por el SNI sobre su estrategia y desempeño como investigadores; y v) percepción sobre la necesidad de reformar el mecanismo de evaluación del Sistema.

Adicionalmente y en las secciones IX.2, IX.2.1 y IX.2.2 -sobre los ámbitos y motivaciones de la producción de conocimiento que realizan los investigadores del SNI-, se retoman las definiciones conceptuales de las propensiones modo 1 y modo 2 de producción de conocimiento, se describen las distribuciones de los ítems que operacionalizan estos conceptos, y se analiza cómo varían dichas distribuciones según los grupos de interés derivados de la caracterización de los investigadores SNI realizada. Además, y a partir de la validación derivada del AFC, se calcularon las puntuaciones factoriales de los individuos. Estas puntuaciones factoriales se utilizaron como índices, dando lugar a los Índices de Propensión Modo 1 y Propensión Modo 2. Estos índices permitieron ubicar a los investigadores del SNI en una escala que mide justamente sus propensiones Modo 1 y Modo 2 de producción de conocimiento e identificar en qué grupos -de acuerdo a la caracterización de los investigadores- hay mayores propensiones y de qué tipo (Modo 1 o Modo 2). En la sección IX.2.3 se presenta la síntesis de los resultados obtenidos de las propensiones Modo 1 y Modo 2 de producción de conocimiento en el SNI mexicano.

En la sección IX.3 se retoma la definición conceptual de satisfacción con el proceso de evaluación, se describen las distribuciones de los ítems que operacionalizan ese concepto, y se analizan las variaciones de las distribuciones según los grupos de interés. Aquí también, y a partir de la validación derivada del AFC, se calcularon las puntuaciones factoriales de los individuos, dando lugar al Índice de Satisfacción con el Proceso de Evaluación del SNI. Esto permitió ubicar a los investigadores del SNI en una escala según la satisfacción que tienen con el proceso de evaluación e identificar qué grupos están más o menos satisfechos. Al final de la sección se sintetizan los resultados obtenidos.

En la sección IX.4 se retoma la definición conceptual de flexibilidad epistemológica del sistema de evaluación, se describen las distribuciones de los ítems que operacionalizan este concepto, y se analizan las variaciones de las distribuciones según los grupos de interés. Aquí también, y a partir de la validación derivada del AFC, se calcularon las puntuaciones factoriales de los individuos, dando lugar al Índice de Flexibilidad Epistemológica del sistema de evaluación. Esto permitió ubicar a los investigadores del SNI en una escala según la flexibilidad epistemológica del sistema de evaluación

que perciben e identificar qué grupos perciben mayor y menor flexibilidad. Al final de la sección se sintetizan los resultados obtenidos.

En la sección IX.5 se retoma la definición conceptual de la dirección de la reforma del sistema de evaluación, se describen las distribuciones de los ítems que operacionalizan ese concepto, y se analizan las variaciones de las distribuciones según los grupos de interés. Aquí también, y a partir de la validación derivada del AFC, se calcularon las puntuaciones factoriales de los individuos, dando lugar al Índice de Dirección de la Reforma del sistema de evaluación. Esto permitió ubicar a los investigadores del SNI en una escala según la dirección definida de la reforma e identificar qué grupos presentan mayor o menor interés en la dirección de la reforma definida. Al final de la sección se sintetizan los resultados obtenidos.

Finalmente, la sección IX.6 concluye.

### IX.1 Caracterización de los investigadores categorizados en el SNI

El objetivo de esta sección es describir algunas características de los investigadores SNI. Esta caracterización de los investigadores se realiza, en primer lugar, desde el punto de vista demográfico atendiendo a la edad y el género de los investigadores. En segundo lugar, dicha caracterización se hace en base al desempeño de los investigadores en el SNI, focalizando en su nivel de avance académico en el sistema, si han ascendido de categoría y si han participado como evaluadores. En tercer lugar, se caracteriza a los investigadores del SNI de acuerdo a su pertenencia cognitiva y laboral, observando el área de conocimiento y el tipo de institución/organización en que adscriben sus actividades de investigación. Finalmente, la caracterización se realiza según la percepción que tienen los investigadores sobre: i) los efectos del SNI, y ii) la necesidad de reformar el mecanismo de evaluación del SNI. Las variables utilizadas en la descripción de las características de los investigadores son a la vez variables de control que determinan grupos de interés para el análisis.

Desde el punto de vista demográfico, en la Tabla 7 se puede observar la distribución de los investigadores categorizados en el SNI según su género, y en la Tabla 8 la distribución según los tramos de edad de los investigadores. Cerca de las dos terceras partes (64.2%) de los investigadores categorizados en el SNI son hombres y el 35.8% restante son mujeres (Ver Tabla 7).

*Tabla 7: Distribución de investigadores SNI por Género*

Género	Nro. de investigadores	% del total de investigadores
Mujeres	2881	35.77
Hombres	5174	64.23
<b>N total</b>	<b>8055</b>	<b>100.00</b>

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Tal como se muestra en la Tabla 8, más de las tres cuartas partes (75.9%) de los investigadores SNI tienen entre 35 y 59 años de edad. Los más jóvenes, menores de 35 años, representan al 9.7% , y los mayores, de 60 años y más, representan al 14.4% del total (ver Tabla 8).

Tabla 8: Distribución de investigadores SNI por Tramos de Edad

Tramos de Edad (años)	Nro. de investigadores	% del total de investigadores	% Acumulado
34 y menos	784	9.73	9.73
35 - 39	1357	16.85	26.58
40 - 44	1405	17.44	44.02
45 - 49	1151	14.29	58.31
50 - 54	1229	15.26	73.57
55 - 59	971	12.05	85.62
60 y más	1158	14.38	<b>100.00</b>
<b>N total</b>	<b>8055</b>	<b>100.00</b>	

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

En base al desempeño de los investigadores en el SNI, la tabla 9 muestra la distribución de los investigadores SNI según su nivel de categorización en el Sistema. Tal como se puede observar, la pirámide de categorización de los investigadores se caracteriza por una estructura más bien achatada, ya que el 75.9% de los investigadores del SNI están categorizados en los niveles más bajos del sistema. De hecho, en el 2013, la mayoría (55.9%) de los investigadores están categorizados en el Nivel I y el 20% en la categoría de Candidatos. Sólo el 24.1% de los investigadores están categorizados en los niveles más altos del Sistema, de los cuales el 16.5% corresponden a los investigadores Nivel II y apenas un 7.6% al Nivel III.

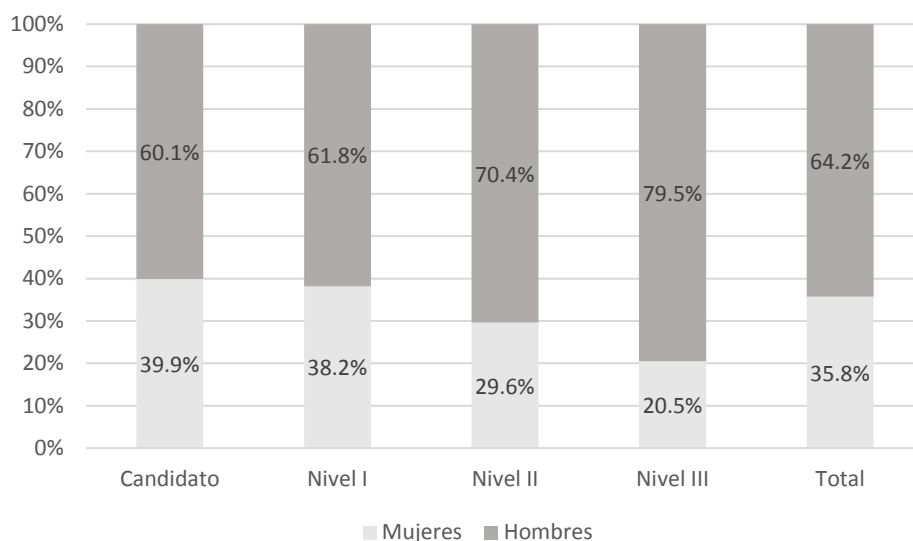
Tabla 9: Distribución de investigadores SNI por Nivel

Nivel de Categorización	Nro. de investigadores	% del total de investigadores	% Acumulado
Candidato	1611	20.00	20.00
Nivel I	4504	55.92	75.92
Nivel II	1328	16.49	92.41
Nivel III	611	7.59	<b>100.00</b>
<b>N total</b>	<b>8054 <sup>(1)</sup></b>	<b>100.00</b>	

Nota: (1) Hay un dato faltante.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Gráfico 3: Participación de Hombres y Mujeres en los niveles del SNI



Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Como se desprende del gráfico 3, existe cierto sesgo negativo en la participación de las mujeres conforme aumenta el nivel de categorización. Mientras que entre los Candidatos existen menos de 2 hombres por mujer, esa relación aumenta entre los Nivel III pasando a casi 4 hombres por cada mujer en el SNI.

En la Tabla 10 se puede observar la distribución de los investigadores SNI según si han ascendido de categoría cuando lo solicitaron o lo esperaban. La amplia mayoría (65.6%) de los investigadores categorizados en el SNI avanzaron en su nivel de categorización cuando lo solicitaron o lo esperaban, sólo poco más de la tercera parte (34.4%) de los investigadores SNI no han podido avanzar en su trayectoria académica en el Sistema. La participación de hombres y mujeres respecto al ascenso en el Sistema tiene una distribución similar al del total. En este sentido, no se encuentra evidencia que sugiera una inequidad potencial en el ascenso entre hombres y mujeres.

Tabla 10: Distribución de investigadores SNI por ascenso en el SNI

Ascendió en el Sistema cuando lo solicitó o lo esperaba	Nro. de investigadores	% del total de investigadores
No	2680	34.41
Sí	5108	65.59
<b>N total</b>	<b>7788<sup>(1)</sup></b>	<b>100.00</b>

Nota: (1) Hay 267 datos faltantes.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

En la Tabla 11 se muestra la distribución de investigadores SNI según si han participado como evaluadores, ya sea en el SNI o en otras instancias de evaluación de la investigación. El 39.3% de los investigadores categorizados en el Sistema ha desempeñado actividades de evaluación de la investigación en el SNI o en otras instancias de evaluación, mientras que el restante 60.7% de los

investigadores SNI no ha participado en ninguna instancia de evaluación de la investigación. La participación de hombres y mujeres como evaluadores tiene una distribución similar al del total. En este sentido, no se encuentra evidencia que sugiera un sesgo a favor –o en contra- de hombres o mujeres en el rol de evaluador.

*Tabla 11: Distribución de investigadores SNI según si ha participado como evaluador*

Ha participado como evaluador	Nro. de investigadores	% del total de investigadores
No	4593	60.69
Sí	2975	39.31
<b>N total</b>	<b>7568</b> <sup>(1)</sup>	<b>100.00</b>

Nota: (1) Hay 487 datos faltantes.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

De acuerdo a la pertenencia cognitiva y laboral de los investigadores SNI, la Tabla 12 muestra su distribución según el área de conocimiento a la que pertenecen y en la Tabla 13 se puede observar la distribución de los investigadores según el tipo de institución/organización en la que desempeñan sus actividades.

Los investigadores pertenecientes al área de Biología y Química son los que tienen mayor participación en el SNI, cerca de la quinta parte (19.8%) de los investigadores del Sistema trabaja en los campos científicos de Biología y Química. Le siguen los investigadores pertenecientes al área de Física, Matemáticas y Ciencias de la Tierra con el 15.2%. En una proporción apenas menor, se encuentran los investigadores pertenecientes a las áreas de Humanidades y Ciencias de la Conducta y de Ciencias Sociales con un 14.9% y 14.2% respectivamente. Finalmente, los investigadores pertenecientes a Ingenierías, Biotecnologías y Ciencias Agropecuarias y a Medicina y Ciencias de la Salud son los que menor participación tienen en el SNI con 12.9%, 12.6% y 10.4% respectivamente (ver Tabla 12).

*Tabla 12: Distribución de investigadores SNI según área de conocimiento*

Áreas	Nro. de investigadores	% del total de investigadores
Física, Matemáticas y Ciencias de la Tierra	1226	15.22
Biología y Química	1594	19.79
Medicina y Ciencias de la Salud	837	10.39
Humanidades y Ciencias de la Conducta	1203	14.93
Ciencias Sociales	1142	14.18
Biotecnología y Ciencias Agropecuarias	1015	12.60
Ingenierías	1038	12.89
<b>N total</b>	<b>8055</b>	<b>100.00</b>

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Desde el punto de vista laboral, es posible observar una gran concentración de los investigadores del SNI en Universidades, el 62.9% de los investigadores del Sistema desempeñan sus actividades en las Universidades (ver Tabla 13). En una proporción considerablemente menor, le siguen los investigadores que desempeñan sus actividades en los Centros Públicos de Investigación (CPI) del

CONACYT o sectorizados, el 12.4% de los investigadores del SNI trabajan en algún CPI<sup>64</sup>. El 11.7% de los investigadores del SNI desarrollan sus actividades en Institutos y Universidades Tecnológicas, y finalmente el 5.7% de los investigadores trabaja en el Gobierno, el 4.7% en los Institutos Nacionales de Salud y el restante 2.5% en otras instituciones (Ver Tabla 13).

*Tabla 13: Distribución de investigadores SNI según el tipo de institución en la que trabajan*

Tipo de Institución	Nro. de investigadores	% del total de investigadores
CPI de Ciencias Exactas y Naturales	371	4.62
CPI de Ciencias Sociales y Humanidades	209	2.60
CPI de Desarrollo Tecnológico y Servicios	77	0.96
CPI Sectorizado	341	4.25
Gobierno	456	5.68
Institutos Nacionales de Salud	380	4.73
Institutos y Universidades Tecnológicas	937	11.67
Universidades	5053	62.94
Otro	204	2.54
<b>N total</b>	<b>8028</b> <sup>(1)</sup>	<b>100.00</b>

Nota: (1) Hay 27 datos faltantes.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

La Tabla 14 muestra la distribución de los investigadores SNI según los efectos<sup>65</sup> que el sistema generó en la estrategia y desempeño de los investigadores.

*Tabla 14: Distribución de investigadores según Efectos del SNI*

Efectos que generó el SNI en la estrategia y desempeño de los investigadores <sup>(1)</sup>	Nro. de investigadores	% del total de investigadores
Negativos	431	5.82
Neutros (el SNI no tuvo efectos)	862	11.65
Positivos	6107	82.53
<b>N total</b>	<b>7400</b> <sup>(2)</sup>	<b>100.00</b>

Notas: (1) Las categorías de la variable son mutuamente excluyentes. (2) Hay 655 datos faltantes, de los cuales el 78% son datos faltantes propiamente dichos y el restante 22% corresponden a la categoría "No Sabe" que, dados los objetivos de este trabajo, se tratan como datos faltantes debido a que no aportan información relevante.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

La amplia mayoría (82.5%) de los investigadores indican que el SNI generó efectos positivos en su estrategia y desempeño como investigadores. El 11.6% indica que los efectos del SNI fueron neutros,

<sup>64</sup> Hay dos tipos de CPI: los agrupados en el Sistema-CONACYT y los que están asociados a Secretarías de Estado. Estos últimos son los sectorizados, y en el contexto de esta tesis se refieren como CPI Sectorizados.

<sup>65</sup> El diseño del cuestionario presenta una debilidad en esta pregunta porque plantea categorías mutuamente excluyentes cuando en realidad pueden no serlo. Así es posible que estén sub-representados los investigadores que perciben efectos positivos y negativos simultáneamente (Bensusán et al., 2014:65).

es decir no afectaron ni positiva ni negativamente su estrategia y desempeño, y apenas un 5.8% señala que el SNI generó efectos negativos (ver Tabla 14).

En la Tabla 15 se presenta la distribución de los investigadores SNI según su percepción sobre la necesidad de reformar el mecanismo de evaluación del SNI. A pesar de que para la amplia mayoría de los investigadores, el SNI ha generado efectos positivos sobre su estrategia y desempeño de investigación, las tres cuartas partes (75.4%) de los investigadores del Sistema considera que el mecanismo de evaluación del SNI debe reformarse (ver Tabla 15).

*Tabla 15: Distribución de investigadores SNI según su percepción sobre reformar el mecanismo de evaluación del SNI*

<b>El mecanismo de evaluación del SNI debe reformarse</b>	<b>Nro. de investigadores</b>	<b>% del total de investigadores</b>
No	1604	24.58
Sí	4922	75.42
<b>N total</b>	<b>6526<sup>(1)</sup></b>	<b>100.00</b>

Nota: (1) Hay 1529 datos faltantes, de los cuales el 19.3% son datos faltantes propiamente dichos y el restante 80.7% corresponden a la categoría "No Sabe" que, dados los objetivos de este trabajo, se tratan como datos faltantes debido a que no aportan información relevante.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

En síntesis, los investigadores del SNI en el 2013 se caracterizan por ser en su mayoría hombres y tener entre 35 y 59 años. Simultáneamente y desde el punto de vista de su desempeño, la amplia mayoría de los investigadores del SNI están categorizados en los niveles más bajos del Sistema, dos de cada tres investigadores han avanzado en su trayectoria académica del SNI y uno de cada tres ha participado como evaluador en el SNI o en otra instancia de evaluación de la investigación. Adicionalmente, los investigadores del SNI muestran una distribución por área cognitiva relativamente similar, aunque con cierto sesgo a favor de los investigadores de Biología y Química, y de Física, Matemáticas y Ciencias de la Tierra, y en detrimento de la participación de los investigadores de Medicina y Ciencias de la Salud, Biotecnologías y Ciencias Agropecuarias e Ingenierías. Desde el punto de vista laboral, los investigadores del SNI muestran una fuerte concentración de su trabajo en las Universidades. Finalmente, si bien la amplia mayoría de los investigadores SNI consideran que el Sistema generó efectos positivos sobre su estrategia y desempeño como investigadores, la amplia mayoría también considera que el mecanismo de evaluación del SNI debe reformarse.

## IX.2 Ámbitos y motivaciones de la producción de conocimiento: las propensiones Modo 1 y Modo 2 en el SNI

En este trabajo las conceptualizaciones Modo 1 y Modo 2 (Gibbons et al., 1994) sobre las formas de producción de conocimiento son retenidas como tipos ideales de producción de conocimiento y se aproximan u operacionalizan a partir de las percepciones de los investigadores SNI sobre los ámbitos y motivaciones que orientan su producción de conocimiento. La *propensión* refiere al grado de acuerdo que manifiestan los investigadores en cada uno de los ítems asociados a las dimensiones de dichos conceptos.

Tal como fueron definidas en el capítulo V:

- i) la *propensión modo 1* es el grado de acuerdo de los investigadores categorizados en el sistema de evaluación con aquella producción de conocimiento que se caracteriza por desarrollar procesos de investigación orientados y organizados disciplinariamente, involucrar a actores homogéneos (investigadores) y difundir el conocimiento producido por canales formales (publicación de artículos, libros y capítulos de libro), lo que representa su principal fuente de reconocimiento científico; y,
- ii) la *propensión modo 2* es el grado de acuerdo de los investigadores categorizados en el sistema de evaluación con aquella producción de conocimiento que se caracteriza por desarrollar procesos de investigación orientados por la búsqueda de soluciones a problemas dependientes del contexto y organizados inter/multi/transdisciplinariamente, involucrar a actores heterogéneos (investigadores y contrapartes interesadas en la solución de problemas) y difundir el conocimiento producido por canales informales (a través de las soluciones encontradas).

A continuación se presentan los resultados descriptivos de las propensiones modo 1 y modo 2.

#### IX.2.1 Propensión modo 1 de producción de conocimiento

En la Tabla 16 se presenta la distribución porcentual de investigadores SNI, según su grado de acuerdo con las afirmaciones expresadas en los ítems asociados a la propensión Modo 1 de producción de conocimiento.

Como puede observarse en la tabla 16, más de las tres cuartas partes de los investigadores SNI (77% aproximadamente) está “De acuerdo” o “Muy de acuerdo” con la afirmación de que *la investigación se orienta por su campo disciplinario*. Al analizar esta distribución según las diversas variables de control, es posible observar que las proporciones en el grado de acuerdo con esa afirmación se mantienen relativamente constantes en la mayoría de éstas. Sin embargo, entre los investigadores pertenecientes al área de Medicina y Ciencias de la Salud, el grado de acuerdo aumenta, ya que el 82.6% de ellos está “De acuerdo” o “Muy de acuerdo” con esa afirmación. Lo mismo sucede entre los investigadores que realizan sus actividades en los CPI Sectorizados, en el Gobierno y en los Institutos Nacionales de Salud, ya que el 81.3%, 81.7% y el 80.5% respectivamente indican estar “De acuerdo” o “Muy de acuerdo”. Finalmente, esa proporción asciende al 84.2% entre los investigadores que no quieren reformar el SNI.

El grado de acuerdo de los investigadores, si bien sigue siendo alto, disminuye ante la afirmación de que *el conocimiento se organiza disciplinariamente*, ya que el 64.5% de los investigadores manifestó estar “De acuerdo” o “Muy de acuerdo” con dicha afirmación (ver Tabla 16). Este grado de acuerdo disminuye levemente entre los investigadores Nivel II y Nivel III, el 60.4% de los primeros y el 62% de los segundos manifestaron estar “De acuerdo” o “Muy de acuerdo”. Los investigadores pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química y Biotecnología y Ciencias Agropecuarias son los que en mayor proporción están “De acuerdo” o “Muy de acuerdo” con la organización disciplinar del conocimiento (73.4%, 70% y 69.4% respectivamente). Esas proporciones alcanzan su mínimo entre los investigadores pertenecientes a las áreas de Ciencias Sociales y de Humanidades y Ciencias de la Conducta, solo el 55.5% de los primeros y el 57.5% de los segundos manifiesta estar “De acuerdo” o “Muy de acuerdo”. Mientras que el 73.6% de los investigadores que trabajan en los Institutos Nacionales de Salud y el 73% de los que desempeñan actividades en el Gobierno manifiestan estar “De acuerdo” o “Muy de acuerdo”, solo el 56.6% de



los que trabajan en los CPI –CONACYT- de Ciencias Sociales y Humanidades manifiesta ese grado de acuerdo con la idea de que el conocimiento se organiza disciplinarmente. El grado de acuerdo con esta afirmación también es menor entre los investigadores que no ascendieron en el sistema cuando lo solicitaron o lo esperaban (el 60.3% de ellos indica estar “De acuerdo” o “Muy de acuerdo”), entre los investigadores que perciben efectos negativos del SNI o neutros (52.6% y 53.5% indican esas categorías respectivamente). Finalmente, mientras que entre los investigadores que no quieren reformar el SNI, el 75.3% de ellos manifiesta estar “De acuerdo” o “Muy de acuerdo” con la afirmación de que el conocimiento se organiza disciplinarmente, entre los investigadores que sí quieren reformar el SNI, solo el 61.3% de ellos mantienen ese grado de acuerdo con dicha afirmación.

En una proporción relativamente similar, el 62.3% de los investigadores SNI está “De acuerdo” o “Muy de acuerdo” con que *la principal fuente de reconocimiento es la publicación de artículos científicos arbitrados*, mientras que solo el 45% aproximadamente está “De acuerdo” o “Muy de acuerdo” con que *la principal fuente de reconocimiento es el libro publicado* (ver Tabla 16). Los investigadores que en mayor proporción indican estar “De acuerdo” o “Muy de acuerdo” con la afirmación de que la principal fuente de reconocimiento es la publicación de artículos científicos en revistas arbitradas son: el 71.7% de los investigadores Nivel III; el 72.2% de los pertenecientes a Medicina y Ciencias de la Salud y el 68.9% de Biología y Química; el 69% de los que desempeñan actividades en el Gobierno y el 71.8% de los que trabajan en los Institutos Nacionales de Salud; y el 79% de los investigadores que no quieren reformar el SNI. El grado de acuerdo con esa afirmación disminuye entre los investigadores pertenecientes a las áreas de Humanidades y Ciencias de la Conducta, de Ciencias Sociales y de Ingenierías, quienes indican estar “Completamente en desacuerdo” o “Poco de acuerdo” en un 21.7%, 18.7%, 18.6% respectivamente. Lo mismo sucede entre los investigadores que desempeñan actividades en los CPI de Ciencias Sociales y Humanidades y de Desarrollo Tecnológico y Servicios, el 24.5% de los primeros y el 21% de los segundos manifiestan estar “Completamente en desacuerdo” o “Poco de acuerdo” con dicha afirmación. Finalmente, el grado de desacuerdo también es alto entre los investigadores que perciben que el SNI generó efectos negativos o neutros sobre su estrategia y desempeño de investigación; la tercera parte de los primeros y la cuarta parte de los segundos indica su completo desacuerdo o poco acuerdo. Adicionalmente, interesa señalar que si bien solo el 45% de los investigadores SNI está “De acuerdo” o “Muy de acuerdo” con que *la principal fuente de reconocimiento es el libro publicado*, esa proporción se incrementa sustantivamente entre los investigadores de Humanidades y Ciencias de la Conducta (63.7%) y de Ciencias Sociales (62.3%), supera el 50% entre los investigadores de 55 años y más, y alcanza a un 48.5% de los investigadores que desempeñan actividades en los CPI de Ciencias Sociales y Humanidades. Contrariamente, y al igual que en el caso de la afirmación sobre los artículos científicos como fuente principal de reconocimiento, los investigadores del área de Ingeniería (28.4%) y los que trabajan en los CPI de Desarrollo Tecnológico y Servicios (31.6%) también indican aquí estar “Completamente en desacuerdo” o “Poco de acuerdo” con que *la principal fuente de reconocimiento es el libro publicado*.

Estos resultados sugieren, en sintonía con la literatura, prácticas diferenciales de la difusión de los resultados de investigación relacionadas con las tradiciones de producción de conocimiento de los diversos campos científicos.

Tabla 16: Distribución porcentual de investigadores SNI, según su grado de acuerdo sobre los ámbitos y motivaciones de la producción de conocimiento (Propensión Modo 1)

Ítems	Completamente en desacuerdo	Poco de acuerdo	Algo de acuerdo	De acuerdo	Muy de acuerdo	Total
La investigación se orienta por su campo disciplinario	1.06	4.05	17.96	46.75	30.19	100.00
El conocimiento se organiza disciplinarmente	2.48	8.85	24.16	45.24	19.28	100.00
Los actores involucrados en la generación de conocimiento son únicamente académicos	13.37	26.47	32.06	22.58	5.52	100.00
La principal fuente de reconocimiento es la publicación de artículos científicos arbitrados	4.8	10.52	22.34	37.31	25.03	100.00
La principal fuente de reconocimiento es el libro publicado	6.04	17.48	31.53	32.96	11.99	100.00

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

El grado de acuerdo de los investigadores se revierte ante la afirmación de que *los actores involucrados en la generación de conocimiento son únicamente académicos*. De hecho, cerca del 40% de los investigadores SNI manifestó estar "Completamente en desacuerdo" o "Poco de acuerdo" con dicha afirmación, el 32.06% está "Algo de acuerdo" y solo poco más del 28% indicó estar "De acuerdo" o "Muy de acuerdo" (ver Tabla 16). Los que indican en mayor proporción estar "Completamente en desacuerdo" o "Poco de acuerdo" con esa afirmación son: el 43.6% de los Candidatos, poco más del 45% de los investigadores menores de 40 años, el 41.2% de los investigadores de Humanidades y Ciencias de la Conducta, el 43.4% de los de Ciencias Sociales, el 42% de los de Biotecnología y Ciencias Agropecuarias, el 44.5% de los de Ingenierías, el 54.5% de los que desempeñan actividades en los CPI de Ciencias Sociales y Humanidades, el 51.8% de los que lo hacen en los CPI Sectorizados y el 46% de los investigadores que trabajan en los CPI de Desarrollo Tecnológico y Servicios. La mitad de los investigadores que indicaron percibir efectos negativos o neutros del SNI y el 43% de los que indicaron que el SNI debe reformarse, también manifestaron estar "Completamente en desacuerdo" o "Poco de acuerdo" con la idea de que en la producción de conocimiento participan exclusivamente académicos.

Con los ítems (variables observadas) con los que se exploraron las percepciones en relación a las formas de producción de conocimiento (Modo 1 y Modo 2), se crearon los índices de *Propensión Modo 1* y *Propensión Modo 2* de los investigadores categorizados en el SNI. Estos índices refieren a las puntuaciones factoriales estandarizadas de los individuos. Las puntuaciones factoriales fueron estandarizadas para asegurar la comparabilidad de los índices y facilitar la interpretación de sus valores. Simultáneamente, y a partir de la subdivisión en cuatro tramos de la distribución de cada uno de los índices, se construyeron dos nuevas variables (una para cada uno) con cuatro categorías – o niveles del índice, a saber: "Muy bajo", "Bajo", "Alto" y "Muy alto". Los puntos de corte para establecer cada uno de esos cuatro niveles se determinaron de acuerdo a la consistencia entre el valor estandarizado de la puntuación factorial y las respuestas a los ítems de los investigadores.

La inmensa mayoría de los investigadores SNI exhiben una propensión Modo 1 de producción de conocimiento elevada. De hecho, el 56.1% de los investigadores del SNI cuenta con una propensión Modo 1 *Alta* y el 21% *Muy Alta*. Sólo el 1.6% de los investigadores tienen una propensión Modo 1 *Muy Baja* y el restante 21.3% *Baja*. Esas proporciones se mantienen relativamente constantes para la mayoría de las diversas variables de control.

Al analizar los intervalos de confianza para el valor promedio del índice según las diversas variables de control, es posible observar con un 95% de confianza, que los hombres, los Nivel III y los de 50 años o más tienen una propensión Modo 1 en promedio significativamente mayor, mientras que las mujeres, los Candidatos y los menores a 40 años son los que presentan una propensión Modo 1 en promedio significativamente menor. Complementariamente, vale aclarar que aunque las mujeres, los Candidatos y los menores a 40 años presenten propensiones Modo 1 en promedio menores, siguen manteniendo una gran proporción en los niveles de *Alta* y *Muy Alta* propensión Modo 1, a saber: el 75.7% de las mujeres –respecto al 77.81% de los hombres-, el 75% de los Candidatos –respecto al 79.5% de los Nivel III- y el 73.4% de los investigadores menores de 40 años –respecto al 80% de los de 50 años o más- tienen una propensión Modo 1 *Alta* o *Muy Alta*.

Adicionalmente, y al considerar el área de conocimiento de los investigadores, se puede constatar - como es esperable- y con un 95% de confianza que los investigadores pertenecientes a las áreas de Medicina y Ciencias de la Salud y de Biología y Química tienen una propensión Modo 1 en promedio significativamente mayor que los pertenecientes a los restantes campos científicos, siendo los de Ingenierías y –curiosamente- los que pertenecen al área de Física, Matemáticas y Ciencias de la Tierra los que exhiben una propensión Modo 1 significativamente menor. De todas maneras, el 72.8% de los investigadores pertenecientes al área de Ingenierías y el 73.9% de los de Física, Matemáticas y Ciencias de la Tierra –respecto al 85.8% de los pertenecientes a Medicina y Ciencias de la Salud y al 78.9% de los de Biología y Química- tienen una propensión Modo 1 *Alta* o *Muy Alta*.

Al observar los intervalos de confianza para el valor promedio del índice según el tipo de institución en la que desempeñan actividades los investigadores, curiosamente los que trabajan en el Gobierno y en los Institutos Nacionales de Salud tienen una propensión Modo 1 en promedio significativamente mayor respecto de los que trabajan en otro tipo de instituciones. De acuerdo con lo esperado, los investigadores que desempeñan actividades en los CPI de Desarrollo Tecnológico y Servicios, de Ciencias Sociales y Humanidades y de Ciencias Exactas y Naturales son los que presentan una propensión Modo 1 en promedio significativamente menor. De hecho, mientras que el 83.3% de los investigadores que trabajan en el Gobierno y el 83.2% de los que lo hacen en los Institutos Nacionales de Salud exhiben una propensión Modo 1 *Alta* o *Muy Alta*, los investigadores que desarrollan sus actividades en los CPI de Ciencias Sociales y Humanidades, de Desarrollo Tecnológico y Servicios y de Ciencias Exactas y Naturales y que presentan una propensión Modo 1 *Alta* o *Muy Alta* representan una proporción menor (61.7%, 70% y 71.7% respectivamente).

Finalmente, los investigadores que no perciben necesario reformar el mecanismo de evaluación del SNI, los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación, los que ascendieron en el sistema y los que sí han sido evaluadores presentan una propensión Modo 1 en promedio significativamente mayor. En estos casos, las variaciones más notorias de las distribuciones porcentuales de los investigadores según los niveles del índice de propensión Modo 1 se dan dependiendo de los efectos percibidos del SNI sobre la estrategia y

desempeño de investigación y de la percepción sobre la reforma del mecanismo de evaluación del sistema. De hecho, mientras que el 80% de los investigadores que perciben efectos positivos del SNI muestran una propensión Modo 1 *Alta o Muy Alta*, esa proporción baja a 59% entre los investigadores que perciben efectos negativos. A la vez, mientras que el 88% de los investigadores que no perciben necesario reformar el mecanismo de evaluación del SNI muestran una propensión Modo 1 *Alta o Muy Alta*, esa proporción baja a 72.1% entre los investigadores que sí perciben la necesidad de reformar el SNI.

#### IX.2.2 Propensión modo 2 de producción de conocimiento

En la Tabla 17 se presenta la distribución porcentual de investigadores SNI, según su grado de acuerdo con las afirmaciones expresadas en los ítems asociados a la propensión Modo 2 de producción de conocimiento.

La amplia mayoría de los investigadores SNI (84.3%) está “De acuerdo” o “Muy de acuerdo” con la afirmación de que *el conocimiento se organiza multi/inter/transdisciplinariamente* (ver Tabla 17). Si bien las proporciones en el grado de acuerdo con esa afirmación se mantienen relativamente similares para la mayoría de las diversas variables de control, es posible observar algunas variaciones respecto a la distribución global. De hecho, a medida que los investigadores ascienden en el SNI, el grado de acuerdo con la organización no disciplinar del conocimiento disminuye. Mientras que el 87.8% de los Candidatos manifiesta estar “De acuerdo” o “Muy de acuerdo” con esa afirmación, entre los Nivel III la proporción de investigadores que indican ese grado de acuerdo se reduce a 71%. Algo similar sucede al analizar la distribución por género. Entre las mujeres el grado de acuerdo con la organización no disciplinar del conocimiento es mayor que entre los hombres. Entre las mujeres el 87.5% está “De acuerdo” o “Muy de acuerdo”, mientras que entre los hombres esa proporción disminuye a 82.5%. Los investigadores pertenecientes al área de Medicina y Ciencias de la Salud son los que en mayor proporción (92%) están “De acuerdo” o “Muy de acuerdo” con la afirmación de que *el conocimiento se organiza multi/inter/transdisciplinariamente*. Lo mismo sucede entre los investigadores que trabajan en los Institutos Nacionales de Salud (92.7%), en los CPI de Desarrollo Tecnológico y Servicios (90.8%) y Sectorizados (87.3%) y en el Gobierno (90%).

Más de las tres cuartas partes de los investigadores categorizados en el SNI (77.4%) manifestó estar “De acuerdo” o “Muy de acuerdo” con que *la generación de conocimiento se orienta por la búsqueda de soluciones a problemas complejos (productivos, sociales, ambientales, de política pública)* (ver Tabla 17). Nuevamente aquí las proporciones en el grado de acuerdo con esa afirmación se mantienen para la mayoría de las variables de control, igualmente se pueden observar algunas variaciones respecto a la distribución global. En particular, conforme los investigadores avanzan en su trayectoria académica en el SNI, el grado de acuerdo con que la producción de conocimiento se orienta por la resolución de problemas de tipo productivos, sociales, ambientales o de política pública disminuye. Entre los Candidatos el 82.4% manifiesta estar “De acuerdo” o “Muy de acuerdo” con esa afirmación, mientras que entre los Nivel III esa proporción baja al 65.4%. Las variaciones en el grado de acuerdo también se pueden observar al analizar las distribuciones porcentuales según los campos científicos. Los investigadores pertenecientes a las áreas de Ciencias Sociales (86.9%), Biotecnologías y Ciencias Agropecuarias (84%), Ingenierías (83.6%) y Medicina y Ciencias de la Salud (82.6%) son los que en mayor proporción están “De acuerdo” o “Muy de acuerdo” con la afirmación de que la producción de conocimiento se orienta por la búsqueda de soluciones a problemas complejos. Ese grado de acuerdo disminuye entre los investigadores pertenecientes a Biología y

Química (67.7%) y entre los investigadores pertenecientes al área de Física, Matemáticas y Ciencias de la Tierra (68%). Al analizar las distribuciones porcentuales en el grado de acuerdo según el tipo de institución en la que trabajan los investigadores, se puede observar que los investigadores que en mayor proporción indican estar “De acuerdo” o “Muy de acuerdo” con dicha afirmación son: el 90.8% de los que desempeñan actividades en los CPI de Desarrollo Tecnológicos y Servicios, el 83.4% de los que trabajan en los CPI Sectorizados y el 85.6% de los que desarrollan sus actividades en el Gobierno.

En una proporción muy elevada, el 76.4% de los investigadores SNI está “De acuerdo” o “Muy de acuerdo” con la afirmación que sostiene que *la generación de conocimiento depende de la interacción entre académicos y las contrapartes interesadas en la solución de problemas* (ver Tabla 17). Las variaciones de las proporciones en el grado de acuerdo con esta afirmación, en relación a la distribución global, se observan cuando se controla por el nivel de categorización de los investigadores, su edad, el área de conocimiento a la que pertenecen y el tipo de institución en la que desempeñan sus actividades de investigación.

Tabla 17: Distribución porcentual de investigadores SNI, según su grado de acuerdo sobre los ámbitos y motivaciones de la producción de conocimiento (Propensión Modo 2)

Ítems	Completamente en desacuerdo	Poco de acuerdo	Algo de acuerdo	De acuerdo	Muy de acuerdo	Total
El objetivo de la generación de conocimiento es su aplicación	3.06	6.86	19.75	36.33	34.01	100.00
La generación de conocimiento se orienta por la búsqueda de soluciones a problemas complejos (productivos, sociales, ambientales, de política pública)	1.98	5.24	15.41	36.41	40.96	100.00
El conocimiento se organiza multi/inter/transdisciplinariamente	0.78	2.72	12.21	36.73	47.56	100.00
La generación de conocimiento depende de la interacción entre académicos y las contrapartes interesadas en la solución de problemas	1.60	5.46	16.53	38.00	38.41	100.00
La principal fuente de difusión del conocimiento es la solución encontrada (productos, procesos, marcas, patentes)	6.62	14.58	28.47	33.38	16.95	100.00

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Otra vez aquí, el grado de acuerdo con la idea de que la producción de conocimiento depende de la interacción entre actores heterogéneos (académicos y no académicos) disminuye a medida que los investigadores progresan en el nivel de categorización del SNI. Mientras que el 82.5% de los Candidatos manifiesta estar “De acuerdo” o “Muy de acuerdo” con esa afirmación, entre los Nivel III la proporción de investigadores que indica ese grado de acuerdo se reduce a 61.7%. Lo mismo sucede a medida que aumenta la edad de los investigadores, mientras que el 80% de los investigadores menores de 40 años está “De acuerdo” o “Muy de acuerdo” con la heterogeneidad de actores involucrados en la producción de conocimiento, entre los investigadores de 60 años y más dicha proporción con ese grado de acuerdo se reduce a 71.8%. El 84.6% de los investigadores

de Biotecnologías y Ciencias Agropecuarias, el 83.4% de los de Medicina y Ciencias de la Salud, el 79.5% de los de Ciencias Sociales y el 78.5% de los de Ingenierías están “De acuerdo” o “Muy de acuerdo” con que la producción de conocimiento depende de la interacción entre académicos y las contrapartes interesadas en la solución de problemas. Lo mismo sucede según el tipo de institución en las que desarrollan actividades los investigadores. Es posible apreciar que los investigadores que en mayor proporción indican estar “De acuerdo” o “Muy de acuerdo” con dicha afirmación son: el 85.5% de los que trabajan en los CPI de Desarrollo Tecnológico y Servicios, el 85% de los que lo hacen en los Institutos Nacionales de Salud y el 81.7% de los que se desempeñan en el Gobierno.

La gran mayoría (70.3%) de los investigadores SNI está “De acuerdo” o “Muy de acuerdo” con la afirmación que señala que *el objetivo de la generación de conocimiento es su aplicación* (ver Tabla 17). Las proporciones en el grado de acuerdo con esa afirmación son similares para la mayoría de las variables de control, de todas maneras es posible observar ciertas variaciones respecto a la distribución global. En particular, a medida que los investigadores ascienden en el SNI, el grado de acuerdo con que el objetivo de la producción de conocimiento es la aplicación de los resultados de investigación disminuye. Entre los Candidatos, el 77.2% de ellos está “De acuerdo” o “Muy de acuerdo” con dicha afirmación, mientras que entre los Nivel III, que mantienen ese grado de acuerdo, dicha proporción baja a 55.3%. Otra vez aquí, el 82.3% de los investigadores de Ingenierías, el 80% de los de Medicina y Ciencias de la Salud y el 79.8% de los de Biotecnologías y Ciencias Agropecuarias están “De acuerdo” o “Muy de acuerdo” con la aplicación como objetivo de la producción de conocimiento. Lo mismo sucede según el tipo de institución en las que desarrollan actividades los investigadores. Los investigadores que en mayor proporción indican estar “De acuerdo” o “Muy de acuerdo” con dicha afirmación son: el 90.8% de los que trabajan en los CPI de Desarrollo Tecnológico y Servicios, el 81.6% de los que lo hacen en los CPI Sectorizados, el 81% que desarrollan sus actividades en el Gobierno y el 77.3% de los que se desempeñan en los Institutos Nacionales de Salud.

Finalmente, el grado de acuerdo entre los investigadores disminuye apreciablemente ante la afirmación que dice que *la principal fuente de difusión del conocimiento es la solución encontrada (productos, procesos, marcas, patentes)*. De hecho, poco más de la mitad (50.3%) de los investigadores SNI está “De acuerdo” o “Muy de acuerdo”, el 28.5% está “Algo de acuerdo” y el restante 21.2% está “Completamente en desacuerdo” o “Poco de acuerdo” con esa afirmación (ver tabla 17). Estas proporciones varían, en relación a la distribución global, si se observa la distribución según el nivel de categorización de los investigadores, el área de conocimiento a la que pertenecen y el tipo de institución en la que desempeñan sus actividades de investigación. Conforme los investigadores avanzan en los niveles del SNI, el grado de acuerdo con que la principal fuente de difusión del conocimiento es la solución encontrada disminuye. Mientras que más de la mitad de los Candidatos (54.8%) indican estar “De acuerdo” o “Muy de acuerdo” con esa afirmación, entre los Nivel III ese grado de acuerdo se mantiene sólo para el 42.3% de los investigadores en ese nivel. Al analizar la distribución según el campo científico al que pertenecen los investigadores, es posible observar aquí también que los investigadores pertenecientes a Ingenierías, Biotecnologías y Ciencias Agropecuarias y Medicina y Ciencias de la Salud son los que en mayor proporción (59.2%, 59.1% y 54.6% respectivamente) indican estar “De acuerdo” o “Muy de acuerdo” con la idea de que la principal fuente de difusión del conocimiento es la solución encontrada. Similarmente y según el tipo de institución en las que desarrollan actividades los investigadores, se puede observar que los

investigadores que en mayor proporción indican estar “De acuerdo” o “Muy de acuerdo” con dicha afirmación son: el 64.5% de los que trabajan en los CPI de Desarrollo Tecnológico y Servicios, el 63.3% de los que se desempeñan en el Gobierno y el 62.4% de los que desarrollan sus actividades en los CPI Sectorizados.

Aquí también, al igual que con la propensión Modo 1, la inmensa mayoría de los investigadores SNI exhiben una propensión Modo 2 de producción de conocimiento elevada. El 49.6% de ellos presenta una propensión *Alta* y el 27.5% *Muy Alta*. Únicamente el 2.1% y el 20.8% de los investigadores SNI muestra una propensión *Muy Baja* y *Baja* respectivamente. Si bien esas proporciones se mantienen relativamente similares, es posible observar algunas variaciones de esta distribución global según el nivel de categorización, el campo científico en el que inscriben sus actividades los investigadores, el tipo de institución en la que trabajan, su edad, género, el tipo de efectos que perciben sobre sus estrategias y desempeño como investigadores y según la percepción que tienen respecto a la necesidad de reformar o no reformar el SNI. A continuación se describen estos resultados y se analizan los intervalos de confianza para el valor promedio del índice propensión Modo 2.

A diferencia de lo que se observó con los intervalos de confianza para el valor promedio del índice propensión Modo 1, aquí se puede constatar con un 95% de confianza, que las mujeres, los Candidatos y Nivel I y los jóvenes (menores a 35 años) tienen una propensión Modo 2 en promedio significativamente mayor, mientras que los Nivel II y III y los investigadores de 60 años y más presentan una propensión Modo 2 en promedio significativamente menor. Aunque los investigadores Nivel II y III y los de 60 años y más presenten propensiones Modo 2 en promedio menores, el 68.7% de los Nivel II y el 57.8% de los Nivel III y el 73.3% de los investigadores de 60 años y más tienen una propensión Modo 2 *Alta* o *Muy Alta*.

De acuerdo con lo esperado, se puede afirmar con un 95% de confianza que los investigadores pertenecientes a las áreas de Ingenierías, Biotecnologías y Ciencias Agropecuarias, Medicina y Ciencias de la Salud y Ciencias Sociales exhiben una propensión Modo 2 en promedio significativamente mayor, mientras que los de Física, Matemáticas y Ciencias de la Tierra, Biología y Química y Humanidades y Ciencias de la Conducta una significativamente menor. De hecho, mientras que aproximadamente el 85% de los investigadores pertenecientes a Ingenierías, Biotecnologías y Ciencias Agropecuarias, Medicina y Ciencias de la Salud y Ciencias Sociales tienen una propensión Modo 2 *Alta* o *Muy Alta*, esa proporción se reduce a 68% aproximadamente entre los investigadores pertenecientes a las restantes áreas.

Al controlar los resultados según el tipo de institución en la que desempeñan actividades los investigadores, es posible observar, con un 95% de confianza, que los investigadores que trabajan en los CPI de Desarrollo Tecnológico y Servicios, y en los Sectorizados, en el Gobierno y en los Institutos Nacionales de Salud tienen una propensión Modo 2 en promedio significativamente mayor, mientras que los que trabajan en los CPI de Ciencias Sociales y Humanidades y de Ciencias Exactas y Naturales y en las Universidades presentan una propensión Modo 2 en promedio significativamente menor. De hecho, el 94.8% de los investigadores que desarrollan sus actividades en los CPI de Desarrollo Tecnológico y Servicios, y el 86.5% de los que trabajan en los Sectorizados, el 87.5% de los que lo hacen en el Gobierno y el 83.4% de los Institutos Nacionales de Salud presentan una propensión Modo 2 *Alta* o *Muy Alta*. Esas proporciones descienden entre los investigadores que trabajan en los CPI de Ciencias Sociales y Humanidades y de Ciencias Exactas y

Naturales, y entre los que se desempeñan en las Universidades (67%, 72.2% y 75.4% respectivamente).

Finalmente, mientras que el 80.6% de los investigadores que perciben efectos positivos del SNI muestran una propensión Modo 2 *Alta o Muy Alta*, esa proporción baja a 63.1% entre los investigadores que perciben efectos negativos. Simultáneamente, el 81.5% de los investigadores que no perciben necesario reformar el mecanismo de evaluación del SNI muestran una propensión Modo 2 *Alta o Muy Alta*, esa proporción baja a 75.2% entre los investigadores que sí quieren reformar el SNI.

### IX.2.3 Una síntesis de los resultados

La amplia mayoría, más de las tres cuartas partes (77%), de los investigadores pertenecientes al SNI mexicano exhiben propensiones Modo 1 y Modo 2 de producción de conocimiento *Alta o Muy Alta*. Estos resultados sugieren que una parte importante de la comunidad académica y científica de México –la contenida en el SNI- mayoritariamente coincide o legitima ambas formas de producción de conocimiento. Sin embargo, la importancia relativa de las propensiones Modo 1 y Modo 2 varían de acuerdo a ciertas características diferenciales de los investigadores.

Los investigadores que exhiben una propensión Modo 1 en promedio mayor son los hombres, los más grandes (50 años y más) y los categorizados en el Nivel III. Por su parte, las mujeres, los más jóvenes (menores de 35 años), los Candidatos y los Nivel I muestran una propensión Modo 2 en promedio mayor. Estos resultados sugieren que conforme avanza la edad de los investigadores hay una cierta adaptación y/o aprendizaje respecto de lo que el sistema de evaluación del SNI premia y valora. En los gráficos 4, 5 y 6 se pueden observar las distribuciones de los índices de Propensión Modo 1 y Propensión Modo 2 para el total, por género, por tramos de edad y por nivel de categorización respectivamente.

Gráfico 4: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Género)

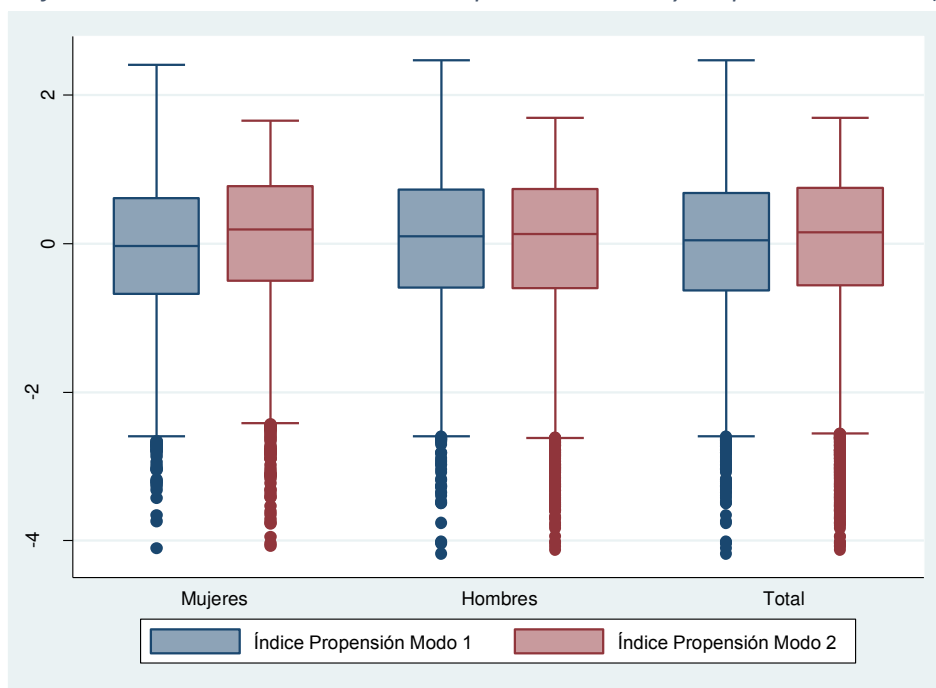




Gráfico 5: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Tramos de Edad)

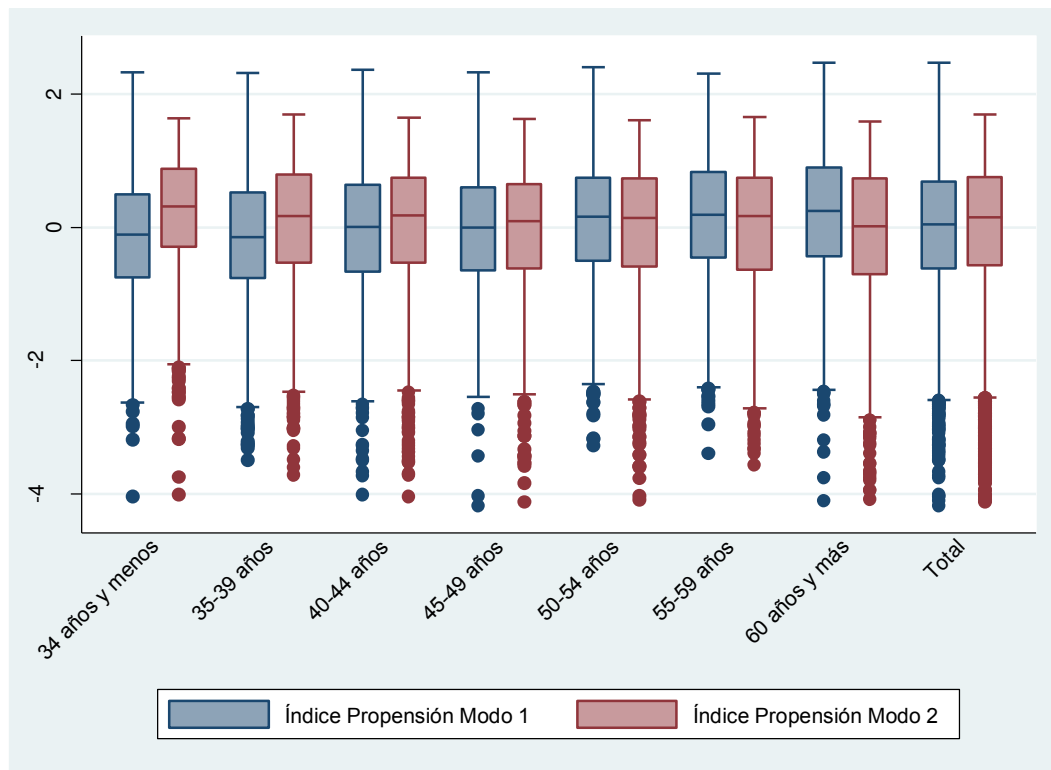
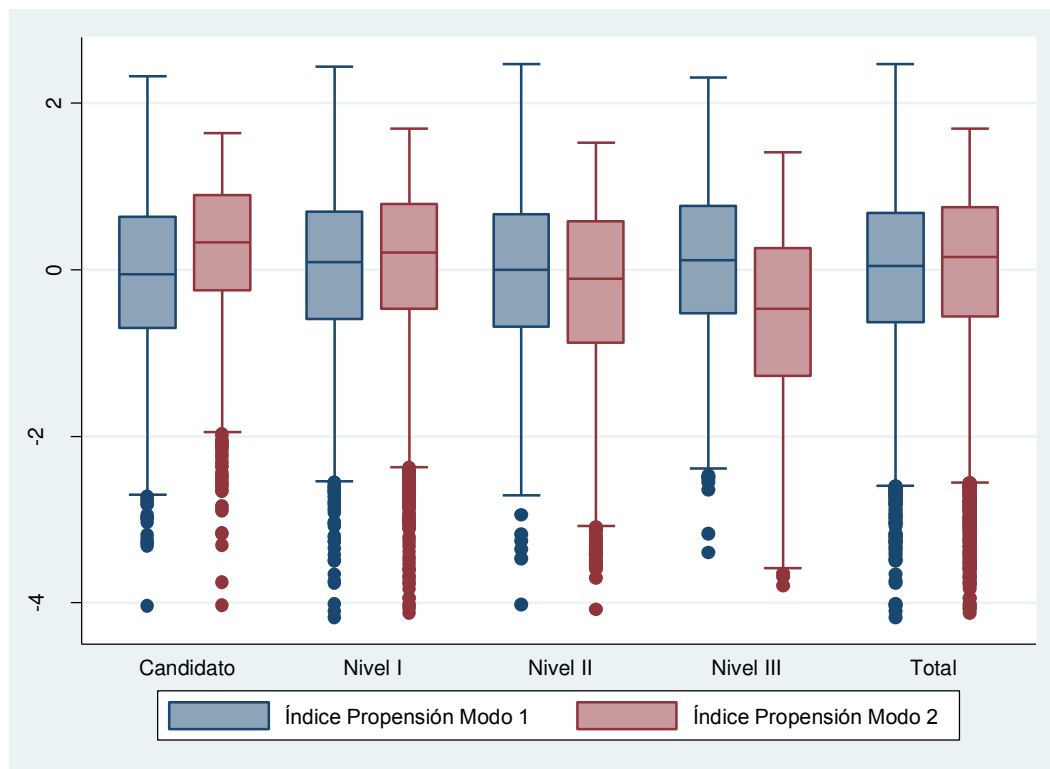


Gráfico 6: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Nivel)



Además, los investigadores que han avanzado en las categorías o trayectoria académica del Sistema y los que han sido evaluadores en el SNI o en otras instancias de evaluación, mientras que presentan en promedio una propensión Modo 1 de producción de conocimiento significativamente mayor, no muestran diferencias significativas respecto al promedio cuando se trata de la propensión Modo 2. Nuevamente aquí, estos resultados sugieren por una parte que el avance en la trayectoria académica del SNI se vincula más con una propensión Modo 1 de producción de conocimiento que con una propensión Modo 2. Adicionalmente, y por otra parte, estos resultados también sugieren que la experiencia o práctica de evaluación, que desempeñan los investigadores, los induce a reproducir cierta forma de producción de conocimiento, en este caso la asociada con el Modo 1.

La importancia relativa de las propensiones Modo 1 y Modo 2 también varía dependiendo del área de conocimiento y el tipo de institución en la que trabajan los investigadores. Las excepciones a este resultado son para los casos de los investigadores de Medicina y Ciencias de la Salud y para los que trabajan en el Gobierno o en los Institutos Nacionales de Salud que presentan en promedio propensiones significativamente mayores tanto en Modo 1 como en Modo 2. Esto puede explicarse por el hecho de que en dichas categorías se engloba a grupos muy heterogéneos de investigadores en relación a sus intereses, motivaciones, objetivos y orientación de la investigación que realizan.

De todas maneras, mientras que los investigadores de Biología y Química presentan una propensión Modo 1 en promedio significativamente mayor, los investigadores de Ingenierías, Biotecnologías y Ciencias Agropecuarias y Ciencias Sociales exhiben una propensión Modo 2 en promedio significativamente mayor. Este resultado sugiere que no hay una única manera de producir conocimiento y que las características específicas que de dicha producción asuma en términos de

dónde se genera, con qué objetivos u orientación, cómo se organiza, quiénes participan, cómo se difunde y legitima, está asociado al campo científico en el que los investigadores realizan sus investigaciones. El gráfico 7 muestra las distribuciones de los índices de Propensión Modo 1 y Propensión Modo 2 para el total y por área de conocimiento.

Finalmente, los investigadores que trabajan en los CPI de Desarrollo Tecnológico y Servicios y en los CPI Sectorizados presentan una propensión Modo 2 en promedio significativamente mayor. Este resultado sugiere que la producción de conocimiento y las características específicas que asuma se asocia también con el contexto institucional/organizacional en el que se desarrolla, es decir de sus objetivos y misiones institucionales/organizacionales. El gráfico 8 presenta las distribuciones de los índices de Propensión Modo 1 y Propensión Modo 2 para el total y por el tipo de institución/organización en la que trabajan los investigadores.

Gráfico 7: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Área de conocimiento)

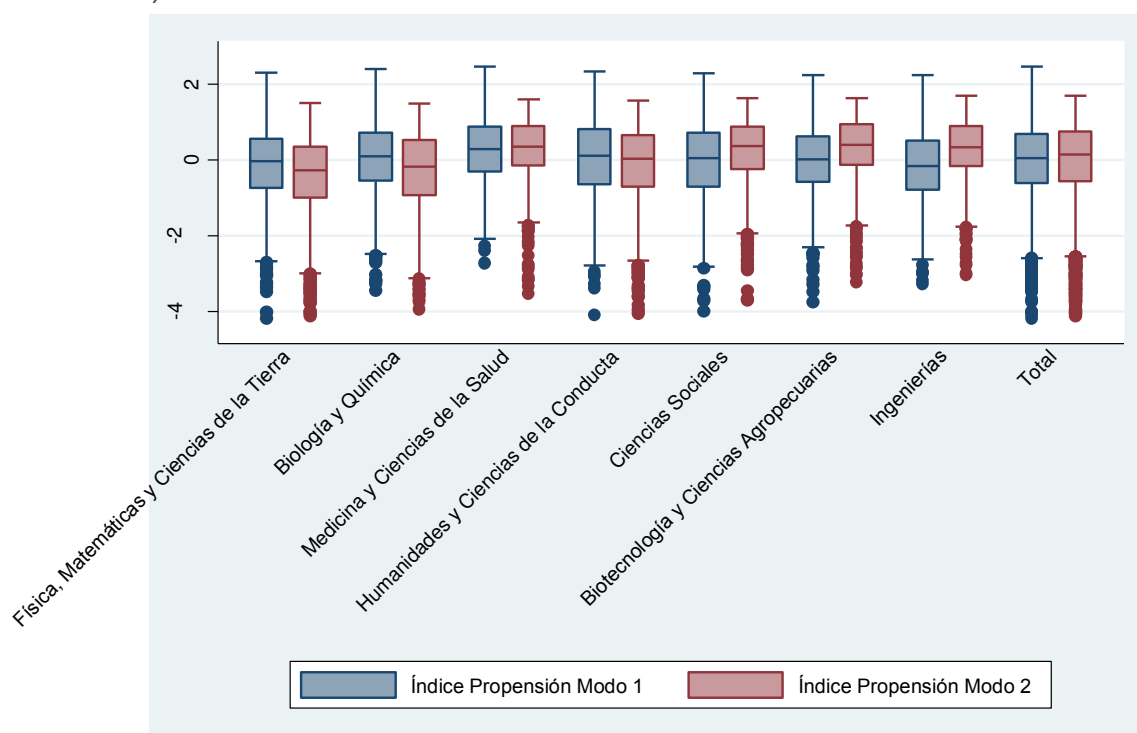
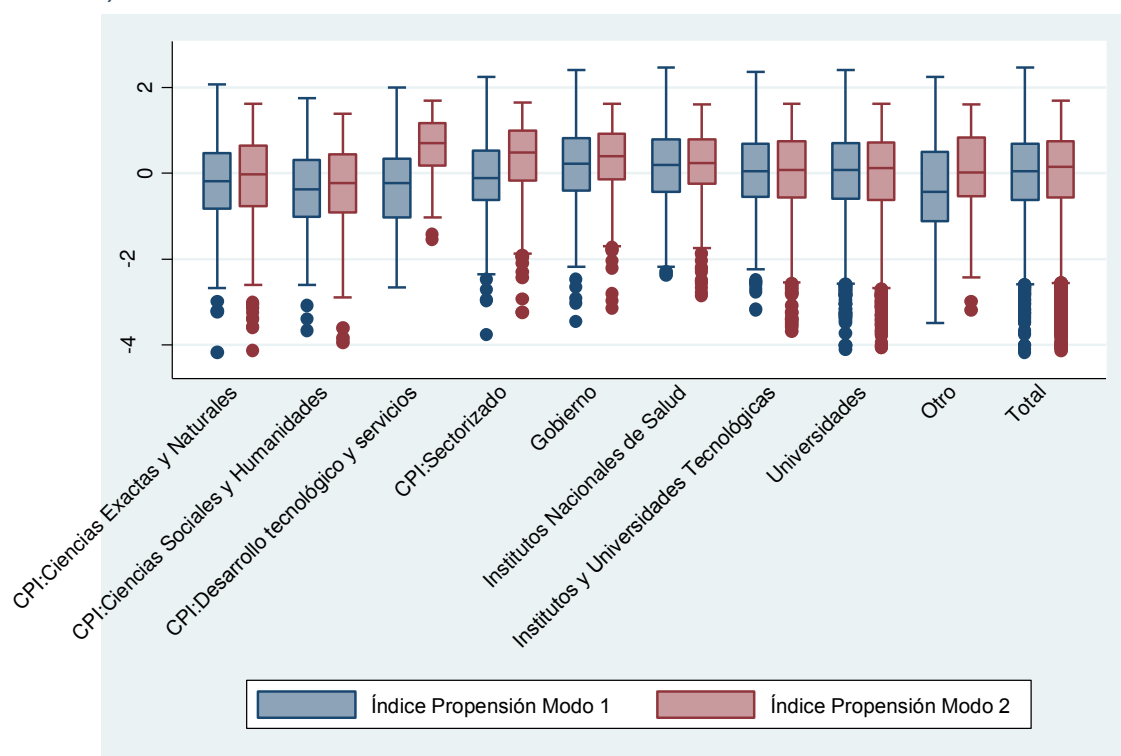


Gráfico 8: Distribución de los Índices Propensión Modo 1 y Propensión Modo 2 (Total y por Tipo de Institución)



### IX.3 Satisfacción con el proceso de evaluación del SNI

Tal como fue definida en el capítulo V, la *satisfacción con el proceso de evaluación es el nivel de satisfacción que tienen los investigadores categorizados en el sistema de evaluación respecto a si el proceso de evaluación es capaz de reconocer el desempeño, es decir, si cuenta con la información necesaria para elaborar juicios académicos apropiados sobre el desempeño de los investigadores e incluye todas sus contribuciones importantes, si los criterios y mecanismos de selección de los evaluadores garantizan evaluaciones académicas apropiadas en las diversas áreas de conocimiento y si la composición de las comisiones dictaminadoras respetan los equilibrios regionales e institucionales y dan garantías respecto a la objetividad de la evaluación.*

La satisfacción con el proceso de evaluación se aproxima u operacionaliza a partir de la percepción que tienen los investigadores sobre: i) si la información solicitada por el SNI sobre el desempeño de los investigadores es la necesaria para lograr juicios académicos apropiados; ii) si la información solicitada no es excesiva; iii) si la información solicitada no deja afuera contribuciones importantes del desempeño de los investigadores; iv) si los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garantizan evaluaciones académicas apropiadas en el área de conocimiento de los investigadores; v) si los requisitos para ser evaluador son pertinentes; vi) si la manera en que se integran las comisiones dictaminadoras es adecuada; vii) si la composición de las comisiones dictaminadoras considera suficientemente el equilibrio regional; viii) si la composición de las comisiones dictaminadoras considera suficientemente el equilibrio institucional; y ix) si la composición de las comisiones dictaminadoras considera otros requisitos (además del Nivel III) para

dar más garantías de una evaluación objetiva. En la Tabla 18 se presenta la distribución porcentual de investigadores SNI, según su percepción sobre estos ítems.

Desde el punto de vista de la relevancia de la información solicitada por el SNI para la evaluación del desempeño de los investigadores, la amplia mayoría de los investigadores (82.6% para el ingreso o reingreso no vigente y el 76.3% para reingreso vigente) considera que dicha información es la necesaria para lograr juicios académicos apropiados; el 97.3% considera que no es excesiva y el 80.4% indica que no deja afuera aportaciones importantes (ver tabla 18). Estas distribuciones si bien se mantienen relativamente similares, es posible observar algunas variaciones respecto al promedio según el nivel de categorización, el área de conocimiento y el tipo de institución en la que trabajan los investigadores. También se observan algunas otras variaciones cuando se controla por el éxito que han tenido en materia de ascenso en el sistema, la percepción que tienen sobre los efectos que el sistema de evaluación ha generado sobre sus estrategias y desempeño como investigadores y respecto a la necesidad de reformar o no el sistema.

Mientras que el 86% aproximadamente de los investigadores Nivel III y de los pertenecientes al área de Medicina y Ciencias de la Salud considera que la información solicitada por el SNI es la necesaria para lograr juicios académicos apropiados -tanto para el ingreso o reingreso no vigente y como para el reingreso vigente-, esa proporción se reduce a 76.9% aproximadamente entre los Candidatos, a 75.4% entre los investigadores pertenecientes al área de Ciencias Sociales y a 73% entre los investigadores de Humanidades y Ciencias de la Conducta. Adicionalmente, y también en proporciones elevadas los que consideran que la información solicitada por el sistema de evaluación es la necesaria para lograr juicios académicos apropiados -tanto para el ingreso o reingreso no vigente y como para el reingreso vigente- son: el 96% de los investigadores que no quieren reformar el SNI, el 85% de los que ascendieron de nivel cuando lo solicitaron o lo esperaban y el 83.4% de los que perciben efectos positivos del Sistema sobre sus estrategias y desempeño como investigadores. Esas proporciones disminuyen a 70% entre los que sí quieren reformar el SNI, a 68.8% entre los que no ascendieron en el sistema y a 51.7% entre los que perciben efectos negativos del SNI. Finalmente, el 21.8% de los Candidatos, el 22.8% de los investigadores de Ciencias Sociales y el 26% de los de Humanidades y Ciencias de la Conducta, y el 34.3% de los que desempeñan actividades en los CPI de Ciencias Sociales y Humanidades, el 42.9% de los que perciben efectos negativos del SNI, el 29.7% de los que no ascendieron y el 28.5% de los que perciben necesario reformar el mecanismo de evaluación son los que en mayor proporción indican que la información solicitada por el SNI deja afuera aportaciones importantes de su desempeño.

*Tabla 18: Distribución porcentual de investigadores SNI, según su percepción sobre los diversos ítems asociados a la Satisfacción con el Proceso de Evaluación*

Ítems	No	Sí	Total
La información que solicita el SNI es la necesaria para lograr un juicio académico apropiado (ingreso o reingreso no vigente)	17.44	82.56	100
La información que solicita el SNI es la necesaria para lograr un juicio académico apropiado (reingreso vigente)	23.74	76.26	100
La información que solicita el SNI es excesiva	97.29	2.71	100
La información que solicita el SNI deja afuera aportaciones importantes	80.37	19.63	100

Los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garanticen una evaluación académica apropiada en su área de conocimiento	35.64	64.36	100
Los requisitos para ser evaluador son pertinentes	21.90	78.10	100
La manera en que se integran las comisiones dictaminadoras es adecuada	38.09	61.91	100
La manera en que se integran las comisiones dictaminadoras no considera suficientemente el equilibrio regional	95.57	4.43	100
La manera en que se integran las comisiones dictaminadoras no considera suficientemente el equilibrio institucional	95.55	4.45	100
La manera en que se integran las comisiones dictaminadoras no considera otros requisitos (además de estar en el Nivel III) para dar más garantías de una evaluación objetiva	86.29	13.71	100

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Desde el punto de vista de los criterios y mecanismos para designar a los evaluadores que participarán del proceso de evaluación, el 78.1% de los investigadores del SNI considera que los requisitos para ser evaluador son pertinentes y el 64.4% considera que los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garantizan evaluaciones académicas apropiadas en su área de conocimiento (ver Tabla 18).

Los investigadores que en mayor proporción indican que los requisitos para ser evaluador son pertinentes son: el 84.9% de los Nivel III, el 83.3% de los de Medicina y Ciencias de la Salud, el 82.2% de los de Física, Matemáticas y Ciencias de la Tierra y el 88.9% de los que desempeñan sus actividades en los CPI de Desarrollo Tecnológico y Servicios. Además, entre los investigadores que progresaron en las categorías del SNI, entre los que perciben efectos positivos del Sistema y entre los que no quieren reformar el mecanismo de evaluación, la proporción de ellos que indica que los requisitos para ser evaluador son pertinentes alcanza 85.5%, 82.9% y 95.7% respectivamente. Por el contrario, los investigadores que en mayor proporción indican que los requisitos para ser evaluador no son pertinentes son: el 29% de los Nivel II, el 27.7% de los de Humanidades y Ciencias de la Conducta, el 25.4% de los de Ciencias Sociales, el 26.9% de los de Ingenierías, el 26.7% de los que trabajan en los CPI de Ciencias Exactas y Naturales y el 40.7% de los que lo hacen en los de Ciencias Sociales y Humanidades, y el 23.7% de los que trabajan en los Institutos y Universidades Tecnológicas. Adicionalmente, entre los investigadores que perciben que el SNI generó efectos negativos en su estrategia y desempeño de investigación, entre los que consideran que los mecanismos de evaluación del SNI deben reformarse y entre los que no ascendieron en el SNI, el 58.6%, el 31.6% y 36.8% de ellos respectivamente considera que los requisitos para ser evaluador no son pertinentes.

Mientras que el 73.4% de los investigadores más jóvenes (menores de 40 años) y el 73.9% de los Candidatos indican que los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garantizan una evaluación académica apropiada en su área de conocimiento, el 40.6% de los investigadores más grandes (mayores de 54 años), el 43.5% de los Nivel II y el 40% de los Nivel III indica lo contrario. El 72.1% de los investigadores pertenecientes al área de Medicina y Ciencias de la Salud, el 70.8% de los que desempeñan actividades en los CPI de Desarrollo Tecnológico y Servicios, el 72.5% de los que trabajan en el Gobierno y el 70.9% de los que lo hacen en los Institutos

Nacionales de Salud son lo que indican en mayor proporción que los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garantizan una evaluación académica apropiada en su área de conocimiento. En contraposición, el 44.3% de los investigadores pertenecientes al área de Humanidades y Ciencias de la Conducta y el 41.2% de los de Ciencias Sociales, el 40.7% de los investigadores que desempeñan sus actividades en los CPI de Ciencias Exactas y Naturales y el 51.9% de los que lo hacen en los CPI de Ciencias Sociales y Humanidades indican lo contrario. Finalmente, mientras que el 73.8% de los investigadores que progresaron en las categorías del SNI, el 70.3% de los que perciben efectos positivos del sistema y el 90.3% de los que no quieren reformar el sistema de evaluación indican que los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garantizan una evaluación académica apropiada en su área de conocimiento, el 54.1% de los que no subieron de categoría, el 75.6% de los que perciben efectos negativos del SNI y el 48.2% de los que sí consideran necesario reformar el sistema indican lo contrario.

La composición de las comisiones dictaminadoras es el aspecto que genera mayores diferencias entre los investigadores. De hecho, mientras que el 61.9% de los investigadores categorizados en el SNI considera que la manera en que se integran las comisiones dictaminadoras es adecuada, el restante 38.1% indica que no es adecuada. Adicionalmente, el 4.4% de los investigadores SNI indica que la manera en que se integran las comisiones dictaminadoras no es adecuada porque no considera suficientemente el equilibrio regional y otro 4.4% porque no considera suficientemente el equilibrio institucional. Finalmente, el 13.7% de los investigadores SNI señala que la manera en que se integran las comisiones dictaminadoras no es adecuada porque no considera otros requisitos (además de estar en el Nivel III) para dar más garantías de una evaluación objetiva. Importa señalar que la amplia mayoría de los investigadores no señalaron afirmativamente ninguna de estos tres últimos ítems (ver Tabla 18).

Los investigadores que en mayor proporción indican que la manera en que se integran las comisiones dictaminadoras es adecuada son los más jóvenes (menores de 40 años), los Candidatos, los investigadores pertenecientes a las áreas de Biología y Química (69%), Medicina y Ciencias de la Salud (70.5%), y los pertenecientes a Biotecnología y Ciencias Agropecuarias (69.2%), también los que trabajan en los Institutos Nacionales de Salud (71%), en el Gobierno (72.6%) y en los CPI de Desarrollo Tecnológico y Servicios (75.5%) y los Sectorizados (66.2%). De hecho, mientras que el 70.4% de los Candidatos y el 72.8% de los investigadores menores de 40 años indican que la manera en que se integran las comisiones dictaminadoras es adecuada, el 46.8% de los Nivel II, el 43% de los Nivel III y el 43.5% de los investigadores mayores de 54 años indican lo contrario. Los investigadores pertenecientes a las áreas de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías son los que en mayor proporción señalan que la manera en que se integran dichas comisiones no es adecuada (49.6%, 45% y 44% respectivamente). Lo mismo sucede entre los investigadores que se desempeñan en los CPI de Ciencias Exactas y Naturales y en los de Ciencias Sociales y Humanidades, el 49.3% de los primeros y el 57.1% de los segundos indican que la manera en que se integran las comisiones dictaminadoras no es adecuada. Por último, mientras que el 72.2% de los investigadores que ascendieron en las categorías del SNI, el 68.8% de los que perciben efectos positivos del sistema y el 91.5% de los que no quieren reformar el sistema de evaluación indican que la manera en que se integran las comisiones dictaminadoras es adecuada, el 57.6% de los que

no subieron de categoría, el 83.4% de los que perciben efectos negativos del SNI y el 53% de los que sí consideran necesario reformar el sistema señalan lo contrario.

Los investigadores que en mayor proporción indican que la manera en que se integran las comisiones dictaminadoras no es adecuada porque no considera suficientemente el equilibrio regional son el 6.2% de los investigadores pertenecientes al área de Ciencias Sociales y el 6.8% de los que trabajan en los CPI de Ciencias Sociales y Humanidades. Entre los que indican en mayor proporción que la manera en que se integran las comisiones dictaminadoras no es adecuada porque no considera suficientemente el equilibrio institucional se encuentran: el 5.9% de los investigadores Nivel III, el 6.2% de los investigadores pertenecientes al área de Ingenierías, el 6.8% de los que trabajan en los CPI de Ciencias Sociales y Humanidades y el 6.1% de los que lo hacen en los de Desarrollo Tecnológico y Servicios, y el 8.9% de los que trabajan en los Institutos Nacionales de Salud. Finalmente, los investigadores que señalan en mayor proporción que la manera en que se integran las comisiones dictaminadoras no es adecuada porque no considera otros requisitos además de estar en el Nivel III son: el 19.4% de los investigadores Nivel II, el 18.3% de los de Ingenierías, el 17.8% de los investigadores que se desempeñan en los CPI de Ciencias Exactas y Naturales, el 18.8% de los que lo hacen en los CPI de Ciencias Sociales y Humanidades y el 17.9% de los que trabajan en los Institutos y Universidades Tecnológicas.

Al igual que en la construcción de los índices de propensión Modo 1 y Modo 2, aquí también, con los ítems que se exploraron las opiniones de los investigadores en relación a la satisfacción con el proceso de evaluación del SNI, se creó el *Índice de Satisfacción con el Proceso de Evaluación* de los investigadores categorizados en el SNI. Este índice refiere a las puntuaciones factoriales estandarizadas de los individuos. Las puntuaciones factoriales fueron estandarizadas para asegurar la comparabilidad con los otros índices y facilitar la interpretación de sus valores. Simultáneamente, y a partir de la subdivisión en cuatro tramos de la distribución del índice, se construyó una nueva variable con cuatro categorías o niveles del índice de *Satisfacción*, a saber: “*Muy baja*”, “*Baja*”, “*Alta*” y “*Muy alta*”. Los puntos de corte para establecer cada uno de esos cuatro niveles se determinaron de acuerdo a la consistencia entre el valor estandarizado de la puntuación factorial y las respuestas a los ítems de los investigadores.

Cerca de las dos terceras partes de los investigadores SNI presentan gran Satisfacción con el Proceso de Evaluación. En particular, el 46.3% de los investigadores SNI tienen una satisfacción *Muy alta*, el 19% *Alta*, el 21.7% *Baja* y el restante 13% *Muy baja*. Si bien estas proporciones se mantienen relativamente similares, es posible observar algunas variaciones cuando se controla por ciertas variables. A continuación se analizan esos resultados junto con los intervalos de confianza para el valor promedio que asume el índice.

De acuerdo con los intervalos de confianza para el valor promedio del índice de *Satisfacción con el Proceso de Evaluación* según las diversas variables de control, es posible constatar con un 95% de confianza, que los investigadores Nivel III, los pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra y Biotecnologías y Ciencias Agropecuarias y los que trabajan en el Gobierno y en los Institutos Nacionales de Salud son los que tienen una Satisfacción con el proceso de evaluación en promedio significativamente mayor. Por su parte, los investigadores Nivel II, los de Humanidades y Ciencias de la Conducta, los de Ciencias Sociales y los de Ingenierías, y los que desempeñan sus actividades en los CPI de Ciencias Sociales y



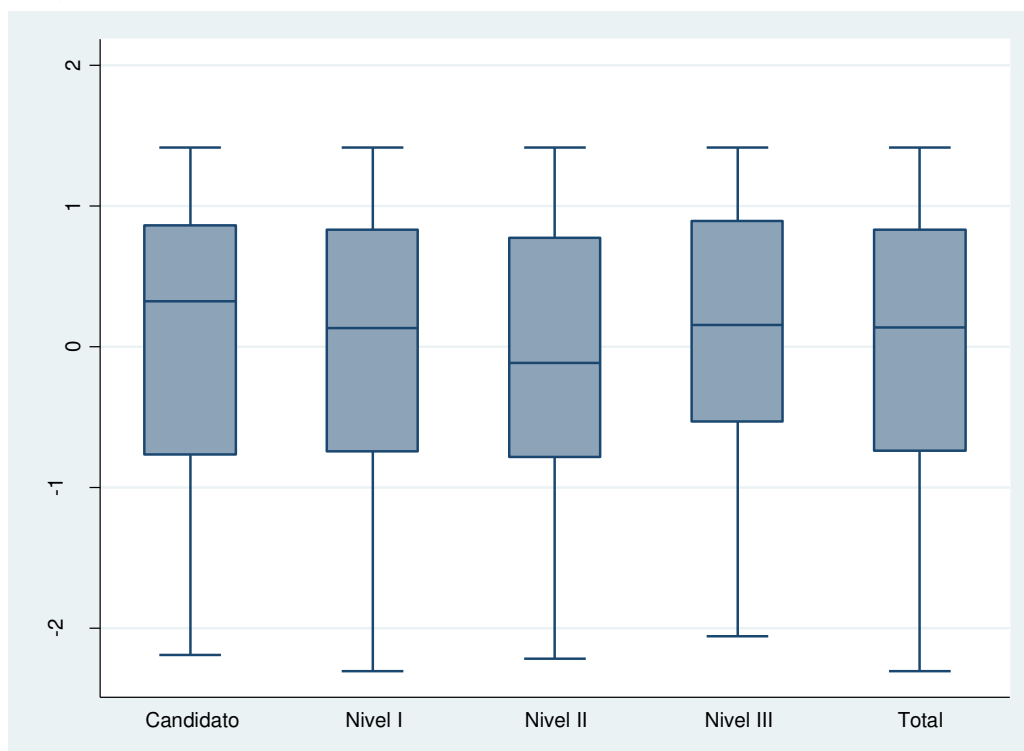
Humanidades y en los de Ciencias Exactas y Naturales son los que exhiben una Satisfacción con el proceso de evaluación en promedio significativamente menor. Mientras que el 71.2% de los investigadores Nivel III tienen una *Satisfacción con el Proceso de Evaluación Alta o Muy alta*, el 39% de los Nivel II exhiben una *Satisfacción Baja o Muy Baja*. Algo similar sucede según el campo científico y el tipo de instituciones en la que desarrollan sus investigaciones los investigadores. Mientras que entre los investigadores pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra y Biotecnologías y Ciencias Agropecuarias y los que trabajan en el Gobierno y en los Institutos Nacionales de Salud el 70% de ellos aproximadamente tiene una *Satisfacción con el Proceso de Evaluación Alta o Muy alta*, el 40% aproximadamente de los investigadores de Humanidades y Ciencias de la Conducta, los de Ciencias Sociales y los de Ingenierías, y los que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades y en los de Ciencias Exactas y Naturales muestran un nivel de Satisfacción *Baja o Muy Bajo*.

Finalmente, los investigadores que ascendieron en las categorías del sistema, los que no perciben necesario reformar el mecanismo de evaluación del SNI y los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación presentan una *Satisfacción con el Proceso de Evaluación* en promedio significativamente mayor que los investigadores que no han ascendido en el sistema, los que sí quieren reformar y los que perciben efectos negativos o neutros del SNI. De hecho, mientras que el 71.7% de los investigadores que perciben efectos positivos del SNI muestra una *Satisfacción Alta o Muy Alta*, el 81.9% de los que perciben efectos negativos y el 53.1% de los que no perciben ningún tipo de efecto muestran una *Satisfacción con el proceso de evaluación Baja o Muy Baja*. A la vez, el 74.7% de los investigadores que han avanzado en las categorías del SNI muestra una *Satisfacción Alta o Muy alta*, sin embargo entre los investigadores que no han ascendido, más de la mitad de ellos (53.2%) exhibe una *Baja o Muy baja Satisfacción con el Proceso de Evaluación*. Por último, mientras que casi la totalidad (94.6%) de los investigadores que no quieren reformar el SNI presentan una *Satisfacción Alta o Muy alta*, casi la mitad (49.9%) de los investigadores que sí quieren reformar muestran una *Baja o Muy baja Satisfacción con el Proceso de Evaluación*.

En síntesis, la mayoría, cerca de las dos terceras partes (65.3%) de los investigadores pertenecientes al SNI mexicano exhiben una *Satisfacción con el Proceso de Evaluación del SNI Alta o Muy Alta*. Sin embargo, el nivel de Satisfacción con el Proceso de Evaluación del SNI varía dependiendo de ciertas características diferenciales de los investigadores, tales como: las de desempeño (nivel de categorización y ascenso en el sistema), las asociadas a la pertenencia cognitiva y laboral (área de conocimiento y tipo de institución), las vinculadas al tipo de efectos generados por el SNI en la estrategia y desempeño de los investigadores y las relacionadas a la necesidad de reformar o no el mecanismo de evaluación del SNI.

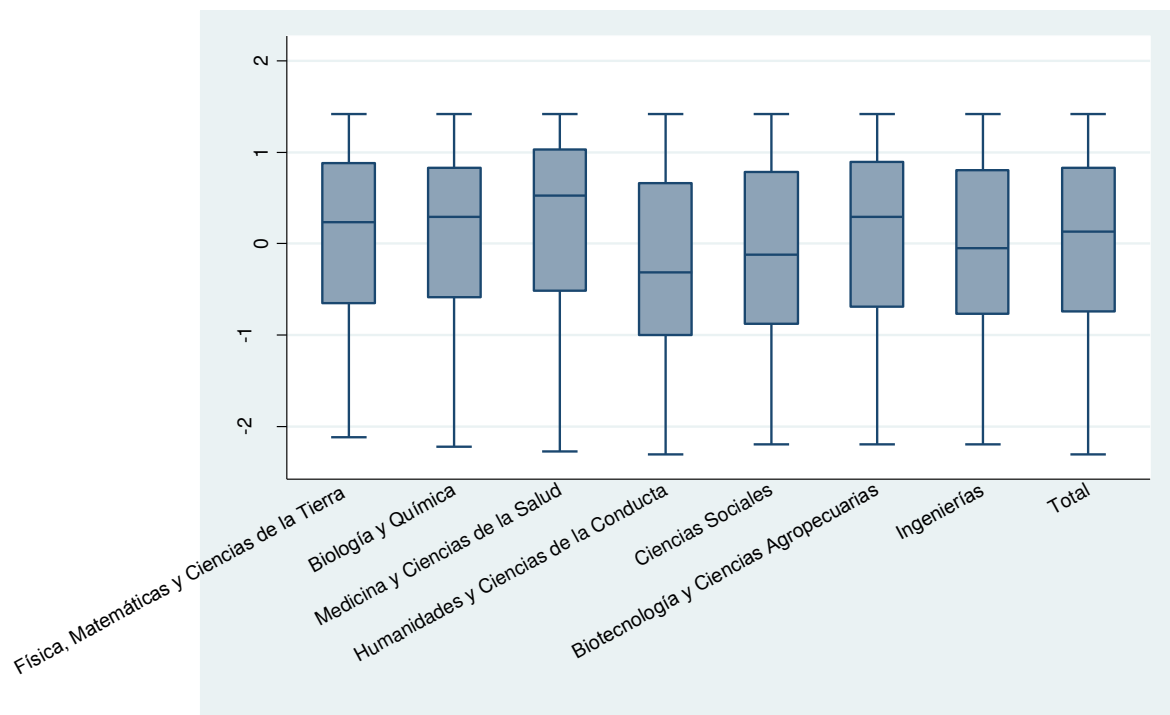
Desde el punto de vista del desempeño de los investigadores se observa que mientras que los Nivel III y los investigadores que ascendieron de categoría en el Sistema son los que muestran una *Satisfacción con el Proceso de Evaluación del SNI* en promedio significativamente mayor; los investigadores Nivel II y los que no han avanzado en las categorías o trayectoria académica del SNI son los que exhiben una *Satisfacción con el Proceso de Evaluación* en promedio significativamente menor. En el gráfico 9 se pueden observar la distribución del índice de Satisfacción con el Proceso de Evaluación del SNI para el total y por nivel de categorización.

Gráfico 9: Distribución del índice de Satisfacción con el Proceso de Evaluación del SNI (Total y por Nivel)



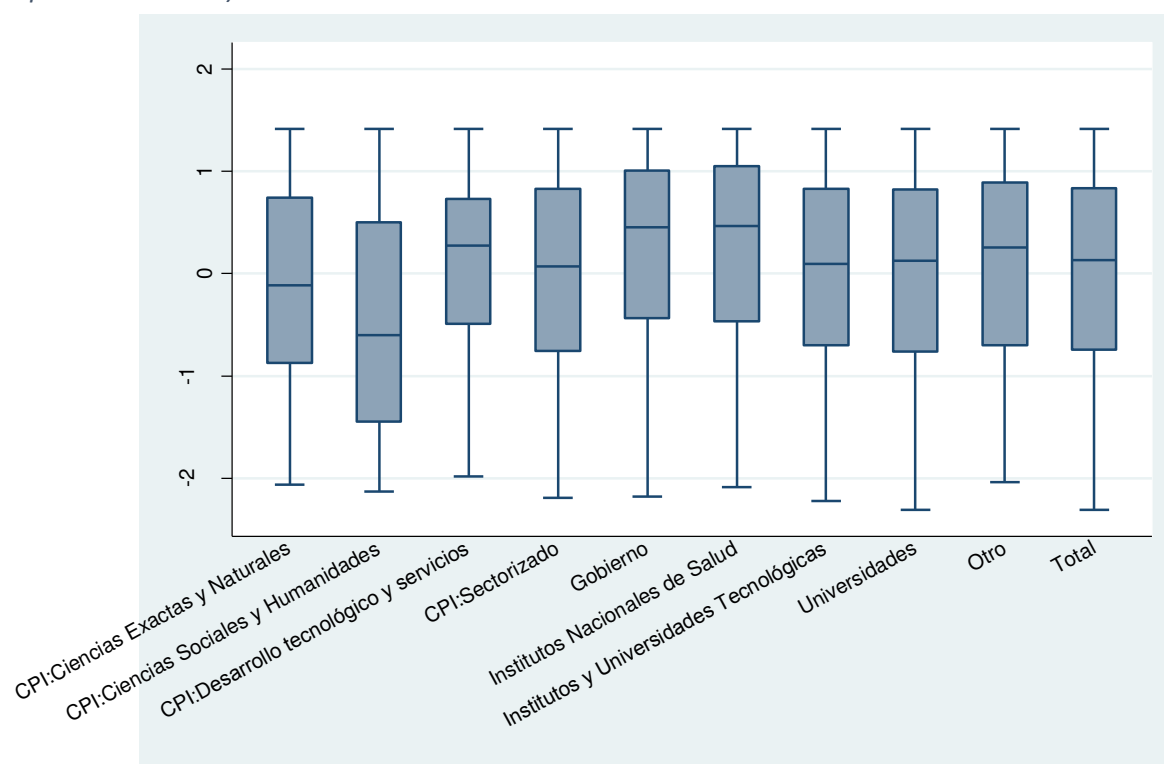
La Satisfacción con el Proceso de Evaluación también varía dependiendo del área de conocimiento y el tipo de institución en la que trabajan los investigadores. Los investigadores que en promedio están significativamente más satisfechos con el proceso de evaluación son los pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra y Biotecnología y Ciencias Agropecuarias. Por su parte, los investigadores de Humanidades y Ciencias de la Conducta, de Ciencias Sociales y de Ingenierías son los que exhiben una satisfacción con el proceso de evaluación en promedio significativamente menor. El gráfico 10 muestra la distribución del índice de Satisfacción con el Proceso de Evaluación del SNI para el total y por área de conocimiento.

Gráfico 10: Distribución del índice de Satisfacción con el Proceso de Evaluación del SNI (Total y por Área de conocimiento)



Desde el punto de vista del ámbito laboral, los investigadores que trabajan en el Gobierno y en los Institutos Nacionales de Salud son los que muestran una Satisfacción con el Proceso de Evaluación del SNI en promedio significativamente mayor. Mientras que los investigadores que en promedio exhiben una satisfacción significativamente menor son los que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades y de Ciencias Exactas y Naturales. El gráfico 11 muestra la distribución del índice de Satisfacción con el Proceso de Evaluación del SNI para el total y por el tipo de institución en la que trabajan los investigadores.

Gráfico 11: Distribución del índice de Satisfacción con el Proceso de Evaluación del SNI (Total y por Tipo de Institución)



Finalmente, mientras que los investigadores que indican que no es necesario reformar el mecanismo de evaluación, o que el SNI generó efectos positivos sobre sus estrategias y desempeño de investigación, exhiben una Satisfacción con el Proceso de Evaluación en promedio significativamente mayor, los investigadores que señalan que sí debe reformarse el SNI o que el Sistema generó efectos negativos o neutros, muestran una Satisfacción con el Proceso de Evaluación significativamente menor en promedio.

#### IX.4 Flexibilidad epistemológica del sistema de evaluación del SNI

En este trabajo la *flexibilidad epistemológica del sistema de evaluación* del SNI refiere a la percepción que tienen los investigadores en relación a si el sistema de evaluación facilita ciertas características epistemológicas de la producción de conocimiento.

Tal como fue definida en el capítulo XX, la *flexibilidad epistemológica del sistema de evaluación* es en qué medida los investigadores categorizados en el sistema de evaluación perciben que dicho sistema de evaluación permite diversidad de perfiles y de trayectorias de investigación, procesos de investigación de largo plazo y colectivos, el desarrollo de estrategias de investigación originales y novedosas, y la obtención de resultados de investigación suficientes –confiables– y científicamente relevantes.

La *flexibilidad epistemológica del sistema de evaluación* se aproxima u operacionaliza a partir de las percepciones de los investigadores categorizados en el SNI sobre: i) si los períodos de evaluación del SNI permiten lograr resultados suficientes y relevantes en la investigación que realizan los

investigadores; ii) si los criterios de evaluación para cada categoría reconocen los distintos perfiles de desempeño de los investigadores; iii) si la integración de las comisiones dictaminadoras considera suficientemente la diversidad disciplinaria; iv) si el SNI no genera estandarización de trayectorias; v) si los criterios de evaluación permiten asumir estrategias de investigación innovadoras; vi) si el SNI no genera dificultades para involucrarse en investigaciones de larga duración; y vii) si el SNI no genera dificultades para trabajar en equipo por la presión para publicar individualmente. En la Tabla 19 se presenta la distribución porcentual de investigadores SNI, según su percepción sobre estos ítems.

*Tabla 19: Distribución porcentual de investigadores SNI, según su percepción sobre si el SNI facilita ciertas características epistemológicas de la producción de conocimiento (Flexibilidad Epistemológica del sistema de evaluación del SNI)*

Ítems	No	Sí	Total
Los períodos de evaluación del SNI para el reingreso vigente permiten lograr resultados suficientes y relevantes en la investigación que Ud. realiza	50.55	49.45	100
Los criterios de evaluación para cada categoría reconocen los distintos perfiles de desempeño de los investigadores	45.13	54.87	100
Los criterios de evaluación le permiten asumir estrategias de investigación innovadoras	53.02	46.98	100
La integración de las comisiones dictaminadoras no considera suficientemente la diversidad disciplinaria	85.77	14.23	100
El SNI le generó dificultades para involucrarse en investigaciones de larga duración	96.73	3.27 <sup>1</sup>	100
El SNI le generó dificultades para trabajar en equipo por la presión para publicar individualmente	97.45	2.55 <sup>1</sup>	100
El SNI generó estandarización de trayectorias	98.85	1.15 <sup>1</sup>	100

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014). Notas: 1/Tal como fue indicado en la nota a pie de página 65, es importante recordar aquí que dada la debilidad del diseño del formulario, las respuestas sobre los efectos negativos del SNI, en este caso asociadas a estos ítems pueden estar sub-representadas.

Las percepciones de los investigadores del SNI se encuentran divididas en relación a si los períodos entre evaluaciones son adecuados para lograr resultados de investigación suficientes –confiables- y científicamente relevantes. De hecho, mientras que el 49.5% de los investigadores SNI considera que los períodos de evaluación permiten lograr resultados suficientes y relevantes en la investigación que realizan, el 50.5% considera lo contrario (ver Tabla 19).

Al controlar este resultado según la edad y trayectoria de los investigadores, es posible observar que para los más jóvenes (menores de 45 años) y aquellos que están iniciando su carrera académica (Candidatos) los períodos entre evaluaciones resultan escasos para obtener resultados suficientes y científicamente relevantes (58.9% y 67.4% respectivamente), mientras que para los más grandes (50 años y más) y los avanzados o consolidados (Niveles II y III) la percepción sobre los períodos de evaluación es la contraria, el 56.5% de los investigadores más grandes, el 61.2% de los Nivel II y el 72.6% de los Nivel III consideran que dichos períodos son los requeridos para obtener resultados de investigación suficientes y científicamente relevantes. Vale la pena señalar que conforme se avanza en la trayectoria académica en el SNI, también aumentan los períodos entre las evaluaciones de desempeño de los investigadores. Mientras que los Candidatos y los Nivel I deben someterse a un

nuevo proceso de evaluación cada tres años, los investigadores Niveles II y III deben hacerlo cada cinco años o más (ver capítulo VI).

Los investigadores pertenecientes a las áreas de Humanidades y Ciencia de la Conducta y los de Ciencias Sociales son los que perciben en mayor proporción los escasos períodos de tiempo entre evaluaciones para alcanzar resultados de investigación suficientes y científicamente relevantes (56.6% y 53.8% respectivamente). Siendo el 53.4% de los investigadores pertenecientes a Física, Matemáticas y Ciencias de la Tierra y el 52.9% de los de Medicina y Ciencias de la Salud los que consideran en mayor proporción que los períodos entre evaluaciones permiten alcanzar resultados de investigación suficientes y relevantes.

Si se considera el tipo de institución en la cual desarrollan sus actividades los investigadores, también es posible encontrar algunas diferencias respecto a la distribución global. En particular, los investigadores que trabajan los CPI de Desarrollo Tecnológico y Servicios y los que lo hacen en los de Ciencias Sociales y Humanidades son los que consideran en mayor proporción que los tiempos entre evaluaciones no son los necesarios para obtener resultados de investigación suficientes y científicamente relevantes (62.3% y 54.8% respectivamente). Por el contrario, el 58.8% de los investigadores que desarrollan sus actividades en los CPI de Ciencias Exactas y Naturales, el 55.2% de los que lo hacen en los CPI Sectorizados, y el 58.1% de los que trabajan en los Institutos Nacionales de Salud son los que en mayor proporción indican que los períodos entre evaluaciones permiten alcanzar resultados de investigación suficientes y científicamente relevantes.

La mayoría de las mujeres (53.3%) percibe que los períodos de evaluación son escasos para alcanzar resultados de investigación suficientes e intelectualmente relevantes, esa proporción baja a 49% entre los hombres. La percepción de la escasez de tiempo disponible entre evaluaciones para obtener resultados de investigación suficientes y científicamente relevantes es mayor cuando se controla según los investigadores que no ascendieron en las categorías del SNI (58.6%), los que perciben efectos negativos o neutros del SNI (77.6% y 58.5% respectivamente) y los que consideran que es necesario reformar el sistema de evaluación (60.2%).

Las opiniones de los investigadores SNI sobre la diversidad de trayectorias y perfiles de investigación que el sistema de evaluación reconoce y valora también están divididas. En particular, mientras que el 54.9% de los investigadores categorizados en el Sistema manifiesta que los criterios de evaluación reconocen los distintos perfiles de desempeño de los investigadores, el 45.1% indica que dichos criterios no reconocen la diversidad de perfiles de investigación. Adicionalmente, y en una proporción sustantivamente menor, el 14.2% de los investigadores SNI indica que la integración de las comisiones dictaminadoras no considera suficientemente la diversidad disciplinaria, y apenas un 1.1% señala que el SNI generó estandarización de trayectorias de investigación<sup>66</sup> (ver Tabla 19).

Al controlar estos resultados según el nivel de categorización de los investigadores, es posible observar que mientras que para cerca de las dos terceras partes (64%) de los investigadores Nivel III los criterios de evaluación sí reconocen los distintos perfiles de investigación, más de la mitad (52%) de los Nivel II señalan lo contrario.

---

<sup>66</sup> La proporción tan baja en este ítem tienen que ver con el diseño del formulario, esta pregunta fue respondida únicamente por los que percibieron efectos negativos del SNI.

Los investigadores pertenecientes a las áreas de Medicina y Ciencias de la Salud, Física, Matemáticas y Ciencias de la Tierra, Biología y Química y Biotecnología y Ciencias Agropecuarias indican en mayor proporción (64.6%, 59.7%, 58.2% y 58.2% respectivamente) que dichos criterios sí reconocen la diversidad de perfiles, mientras que el 58.1% de los de Humanidades y Ciencias de la Conducta y el 49.1% de los de Ingeniería consideran todo lo contrario. Al controlar por el tipo de institución en la que desempeñan actividades los investigadores, es posible observar que el 64% de los que trabajan en los Institutos Nacionales de Salud y el 63.8% de los que lo hacen en el Gobierno son los que indican en mayor proporción que los criterios de evaluación sí reconocen los distintos perfiles de investigación. Los que en mayor proporción indican lo contrario, son el 63.3% de los investigadores que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades. La percepción de que los criterios de evaluación no reconocen los distintos perfiles de investigación se agudiza entre los investigadores que no ascendieron en el SNI (63.8%), los que perciben efectos negativos o neutros del SNI (84.8% y 66.6% respectivamente) y los que consideran que es necesario reformar el sistema de evaluación (59.9%).

Las opiniones de los investigadores SNI también se dividen cuando se les consulta sobre el grado de novedad y originalidad que pueden enfrentar en sus investigaciones dados los criterios de evaluación vigentes. Para la mayoría de los investigadores SNI, el riesgo intelectual que pueden asumir en sus investigaciones es relativamente bajo. De hecho, el 53% de los investigadores categorizados en el sistema señala que los criterios de evaluación del SNI no les permiten asumir estrategias de investigación innovadoras, mientras que el 47% restante indica lo contrario (ver Tabla 19).

Al controlar los resultados según la edad y el nivel de categorización de los investigadores en el SNI, se puede observar que los que indican en mayor proporción que los criterios de evaluación no les permiten asumir estrategias de investigación innovadoras son el 57% de los investigadores más grandes (55 años y más), el 60.6% de los Nivel II y el 54.7% de los Nivel III.

Al observar cómo varían las proporciones globales de acuerdo con campo científico al que pertenecen los investigadores, el 60.6% de los investigadores de Humanidades y Ciencias de la Conducta y el 54.6% de los de Física, Matemáticas y Ciencias de la Tierra son los que señalan en mayor proporción que los criterios de evaluación no les permiten asumir investigaciones más novedosas, originales y por lo tanto, más riesgosas e innovadoras. En contraposición, el 55.8% de los investigadores de Medicina y Ciencias de la Salud y el 53% de los de Biotecnologías y Ciencias Agropecuarias señalan que los criterios de evaluación sí les permiten asumir estrategias de investigación innovadoras.

Según el tipo de institución en la que desempeñen actividades los investigadores, los que en mayor proporción indican que los criterios de evaluación no les permiten asumir estrategias de investigación innovadoras son: el 56.9% de los que trabajan en los CPI de Ciencias Exactas y Naturales, el 75.3% de los que lo hacen en los CPI de Ciencias Sociales y Humanidades, el 60% de los que desempeñan sus actividades en los CPI de Desarrollo Tecnológico y Servicios y el 56.8% de los que lo hacen en los Institutos y Universidades Tecnológicas. En contraste, el 51.5% de los que trabajan en los CPI Sectorizados, el 57.4% de los que lo hacen en el Gobierno y el 51% de los que se desempeñan en los Institutos Nacionales de Salud son los que en mayor proporción señalan que los criterios de evaluación sí les permiten asumir estrategias de investigación innovadoras.

La percepción de que los criterios de evaluación no permiten asumir estrategias de investigación innovadoras se hace mayor entre los investigadores que no ascendieron en el SNI (67.4%), los que perciben efectos negativos o neutros del SNI (85.9%) y los que consideran que es necesario reformar el sistema de evaluación (65.5%). En sentido contrario, el 54.3% de los investigadores sí ascendieron en el SNI, el 54% de los que perciben efectos positivos del SNI y el 78.2% de los que consideran que no es necesario reformar el sistema de evaluación perciben que los criterios de evaluación sí les permiten asumir estrategias de investigación innovadoras.

Desde el punto de vista de las consecuencias del SNI sobre las características temporales de la investigación, casi la totalidad (96.7%) de los investigadores SNI considera que el Sistema no les generó dificultades para involucrarse en investigaciones de larga duración<sup>67</sup> (ver Tabla 19). Los que en mayor proporción indican que el SNI les generó dificultades para involucrarse en investigaciones de largo plazo son: el 4.8% de los Nivel II, el 4.6% de los de Humanidades y Ciencias de la Conducta, el 10.7% de los que trabajan en los CPI de Ciencias Sociales y Humanidades, el 5.3% de los que se desempeñan en los CPI de Ciencias Exactas y Naturales, el 5.5% de los que no ascendieron en el SNI, el 4.8% de los que consideran necesario reformar el Sistema y el 56.1% de los que perciben efectos negativos del SNI sobre su estrategia y desempeño de investigación.

Desde el punto de vista de las consecuencias del SNI sobre la investigación colectiva o individual, casi la totalidad (97.4%) de los investigadores SNI considera que el Sistema no les generó dificultades para trabajar en equipo<sup>68</sup> (ver Tabla 19). Los que en mayor proporción indican que el SNI les generó dificultades para trabajar en equipo son: el 3.5% de los Nivel II, el 3.4% de los de Humanidades y Ciencias de la Conducta, el 7% de los que trabajan en los CPI de Ciencias Sociales y Humanidades, el 4.7% de los que se desempeñan en los CPI de Ciencias Exactas y Naturales, el 8.7% de los que lo hacen en los CPI de Desarrollo Tecnológico y Servicios, el 4.8% de los que no ascendieron en el SNI, el 3.9% de los que consideran necesario reformar el Sistema y el 43.8% de los que perciben efectos negativos del SNI sobre su estrategia y desempeño de investigación.

Al igual que en la construcción de los índices de propensión Modo 1 y Modo 2, y el índice que mide la Satisfacción con el Proceso de Evaluación aquí también, con los ítems que se exploraron las opiniones de los investigadores en relación a si el SNI facilita ciertas características epistemológicas de la producción de conocimiento, se creó el *índice de Flexibilidad Epistemológica del sistema de evaluación* del SNI. De la misma manera que en los índices anteriores, éste refiere a las puntuaciones factoriales estandarizadas de los individuos. Las puntuaciones factoriales fueron estandarizadas para asegurar la comparabilidad con los otros índices y facilitar la interpretación de sus valores. Simultáneamente, y a partir de la subdivisión en cuatro tramos de la distribución del índice, se construyó una nueva variable con cuatro categorías o niveles del índice de *Flexibilidad epistemológica del sistema de evaluación*, a saber: “Muy baja”, “Baja”, “Alta” y “Muy alta”. Los puntos de corte para establecer cada uno de esos cuatro niveles se determinaron de acuerdo a la consistencia entre el valor estandarizado de la puntuación factorial y las respuestas a los ítems de los investigadores.

---

<sup>67</sup> La proporción tan baja en este ítem tienen que ver con el diseño del formulario, esta pregunta fue respondida únicamente por los que percibieron efectos negativos del SNI.

<sup>68</sup> La proporción tan baja en este ítem tienen que ver con el diseño del formulario, esta pregunta fue respondida únicamente por los que percibieron efectos negativos del SNI.



Nuevamente aquí, cerca de las dos terceras partes de los investigadores perciben una gran *Flexibilidad epistemológica del sistema de evaluación*. De hecho, el 36.2% de los investigadores SNI perciben una Flexibilidad Epistemológica *Muy alta* y el 26.8% *Alta*. La tercera parte de los investigadores (33.6%) percibe una Flexibilidad epistemológica del sistema de evaluación *Baja* y el restante 3.4% una *Muy Baja*. Si bien estas proporciones se mantienen relativamente similares, es posible observar algunas variaciones cuando se controla por ciertas variables. A continuación se analizan esos resultados junto con los intervalos de confianza para el valor promedio que asume el índice.

De acuerdo con los intervalos de confianza para el valor promedio del índice de *Flexibilidad epistemológica del sistema de evaluación* según las diversas variables de control, es posible constatar con un 95% de confianza, que los investigadores Nivel III, los más jóvenes (menores de 35 años), los pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra y Biotecnologías y Ciencias Agropecuarias y los que trabajan en el Gobierno y en los Institutos Nacionales de Salud perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente mayor. Por su parte, los investigadores Nivel II, los más grandes (60 años y más), las mujeres, los de Humanidades y Ciencias de la Conducta, los de Ciencias Sociales y los de Ingenierías, y los que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades los que perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente menor.

Mientras que el 71.7% de los investigadores Nivel III y el 71.9% de los más jóvenes (menores de 35 años) perciben una *Flexibilidad epistemológica del sistema de evaluación Alta o Muy alta*, el 42.4% de los Nivel II, 41.3% de los más grandes (60 años y más) y el 38.7% de las mujeres perciben una *Flexibilidad Epistemológica Baja o Muy Baja*. Algo similar sucede según el campo científico y el tipo de instituciones en la que desarrollan sus investigaciones los investigadores. Los investigadores pertenecientes a las áreas de Medicina y Ciencias de la Salud (71.6%), Biología y Química (67.4%), Física, Matemáticas y Ciencias de la Tierra (67.7%) y Biotecnologías y Ciencias Agropecuarias (67.4%) y los investigadores que trabajan en el Gobierno (72.6%) y en los Institutos Nacionales de Salud (70%) son los que perciben en mayor proporción una *Flexibilidad epistemológica del sistema de evaluación Alta o Muy alta*. El 49.6% de los investigadores de Humanidades y Ciencias de la Conducta, el 41.6% de los de Ciencias Sociales, el 40.3% de Ingenierías y el 56.5% de los que trabajan en los CPI de Ciencias Sociales y Humanidades son los que perciben en mayor proporción una *Flexibilidad epistemológica del sistema de evaluación Baja o Muy baja*.

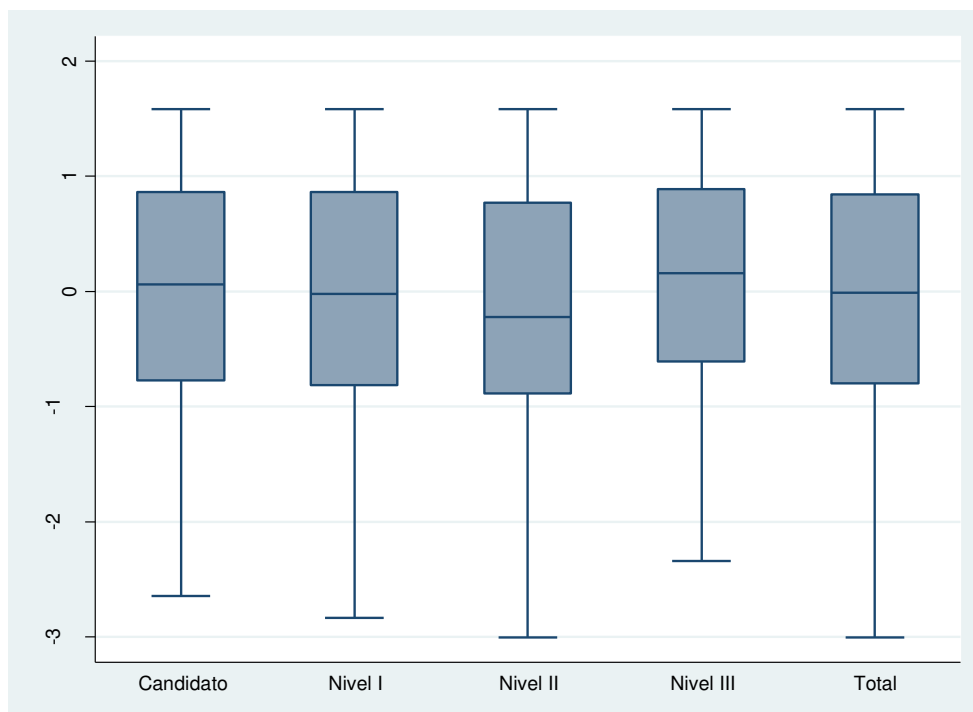
Finalmente, los investigadores que ascendieron en las categorías del SNI, los que no perciben necesario reformar el mecanismo de evaluación del SNI y los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación también perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente mayor que los investigadores que no han ascendido en el sistema, que los que sí quieren reformar y los que perciben efectos negativos o neutros del SNI. De hecho, mientras que el 70.3% de los investigadores que perciben efectos positivos del SNI se ubican en las categorías de *Alta o Muy Alta Flexibilidad epistemológica del sistema de evaluación*, el 88.2% de los que perciben efectos negativos y el 58.9% de los que no perciben ningún tipo de efecto también perciben una *Flexibilidad epistemológica del sistema de evaluación Baja o Muy Baja*. Simultáneamente, mientras que el 72% de los investigadores que han avanzado en las categorías del SNI perciben una *Alta o Muy Alta Flexibilidad*

*epistemológica del sistema de evaluación*, entre los que no ascendieron el 54.8% perciben una *Baja o Muy baja Flexibilidad Epistemológica*. Por último, mientras que casi la totalidad (95.6%) de los investigadores que no quieren reformar el SNI perciben una *Flexibilidad epistemológica del sistema de evaluación Alta o Muy alta*, más de la mitad (52.6%) de los investigadores que sí quieren reformar perciben una *Baja o Muy baja Flexibilidad epistemológica del sistema de evaluación*.

En síntesis, la mayoría, cerca de las dos terceras partes (63.7%) de los investigadores pertenecientes al SNI mexicano perciben una *Flexibilidad epistemológica del sistema de evaluación Alta o Muy Alta*. Sin embargo, el grado de Flexibilidad epistemológica del sistema de evaluación varía dependiendo de ciertas características diferenciales de los investigadores, tales como: las de desempeño (nivel de categorización y ascenso en el sistema), las asociadas a la pertenencia cognitiva y laboral (área de conocimiento y tipo de institución), las vinculadas al tipo de efectos generados por el SNI en la estrategia y desempeño de los investigadores y las relacionadas a la necesidad de reformar o no el mecanismo de evaluación del SNI.

Desde el punto de vista del desempeño de los investigadores se observa -al igual que lo que sucedía con el índice de Satisfacción con el Proceso de Evaluación- que mientras que los Nivel III y los investigadores que ascendieron de categoría en el Sistema son los que perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente mayor; los investigadores Nivel II y los que no han avanzado en las categorías o trayectoria académica del SNI son los que perciben una Flexibilidad Epistemológica en promedio significativamente menor. En el gráfico 12 se puede observar la distribución del índice de Flexibilidad epistemológica del sistema de evaluación para el total y por nivel de categorización.

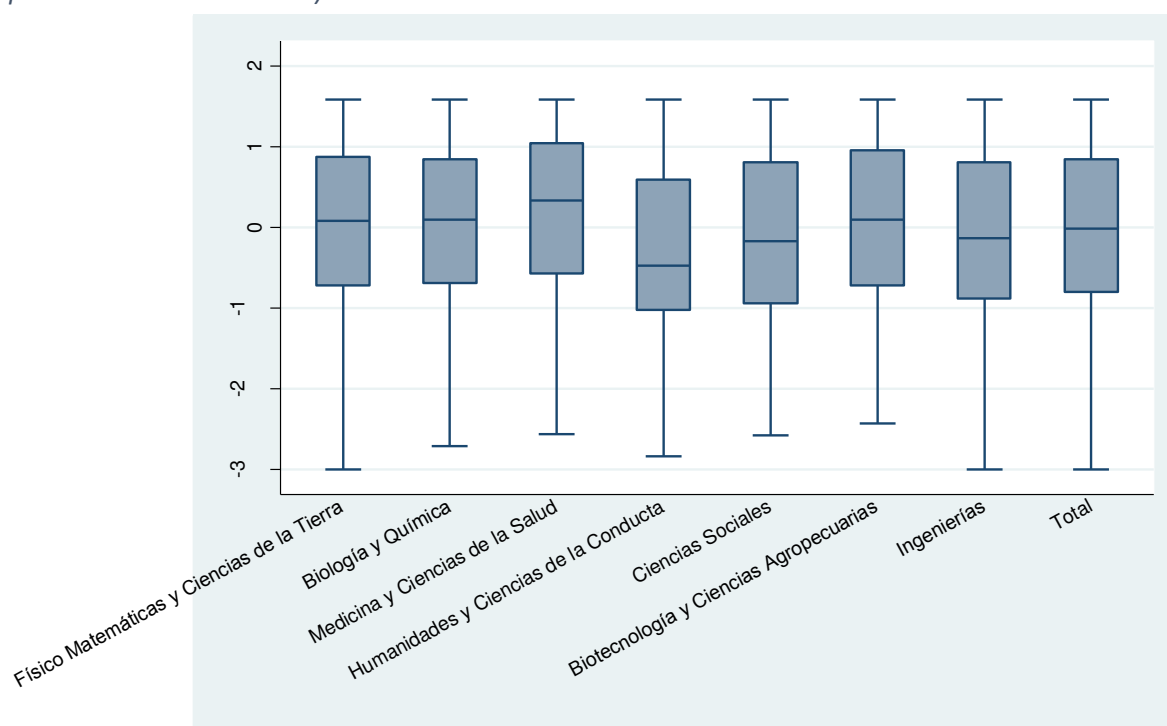
*Gráfico 12: Distribución del índice de Flexibilidad epistemológica del sistema de evaluación (Total y por Nivel)*



Al igual de lo que se observó con el índice de Satisfacción con el Proceso de Evaluación del SNI, el grado de Flexibilidad epistemológica del sistema de evaluación percibido por los investigadores varía dependiendo del área de conocimiento y el tipo de institución en la que trabajan los investigadores.

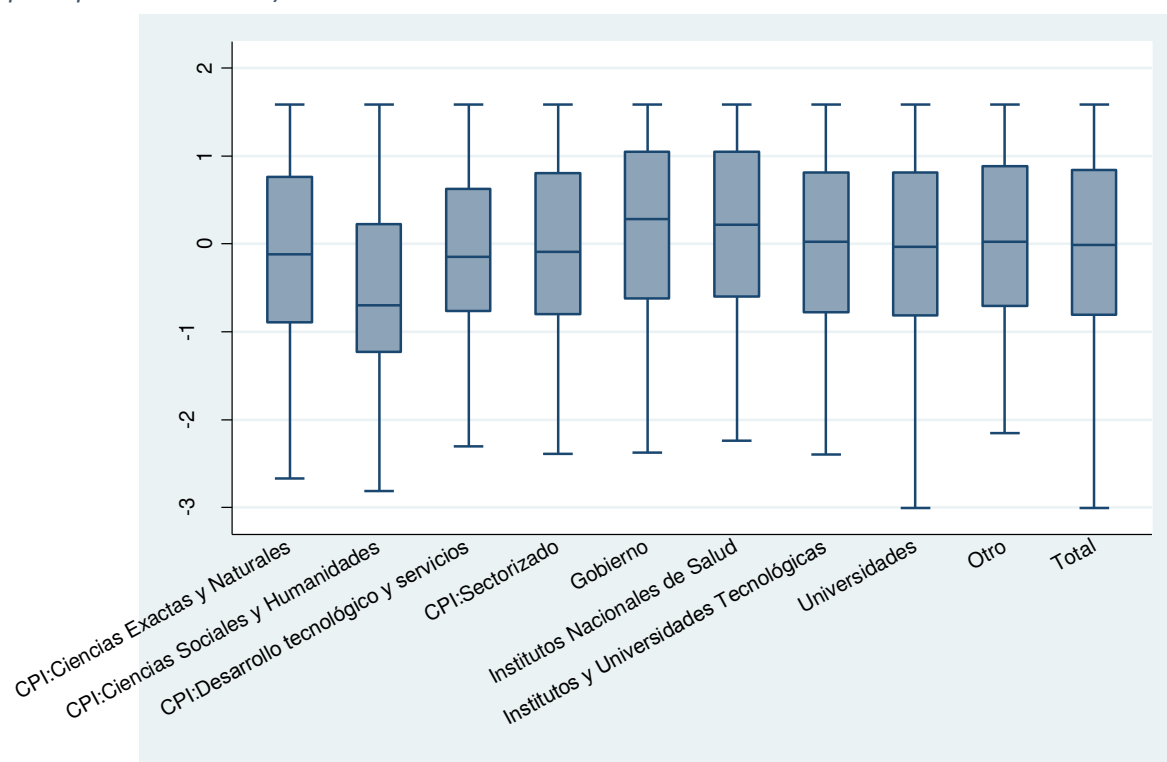
De igual manera que lo que sucedía con el índice de Satisfacción con el Proceso de Evaluación, los investigadores que perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente mayor son los pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra y Biotecnología y Ciencias Agropecuarias. Por su parte, los investigadores de Humanidades y Ciencias de la Conducta, de Ciencias Sociales y de Ingenierías además de ser los que exhiben una menor satisfacción con el proceso de evaluación, también son los que perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente menor. El gráfico 13 muestra la distribución del índice de Flexibilidad epistemológica del sistema de evaluación para el total y por área de conocimiento.

*Gráfico 13: Distribución del índice de Flexibilidad epistemológica del sistema de evaluación (Total y por Área de conocimiento)*



Desde el punto de vista del ámbito laboral, el resultado es similar al obtenido con el índice de Satisfacción con el Proceso de Evaluación. Nuevamente aquí, los investigadores que trabajan en el Gobierno y en los Institutos Nacionales de Salud son los que perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente mayor. Por su parte, los investigadores que perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente menor son los que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades. El gráfico 14 muestra la distribución del índice de Flexibilidad epistemológica del sistema de evaluación para el total y por el tipo de institución en la que trabajan los investigadores.

Gráfico 14: Distribución del índice de Flexibilidad Epistemológica del sistema de evaluación (Total y por Tipo de Institución)



Finalmente y coincidente con los resultados obtenidos con el índice de Satisfacción con el Proceso de Evaluación del SNI, aquí también sucede que mientras que los investigadores que indican que no es necesario reformar el mecanismo de evaluación, o que el SNI generó efectos positivos sobre sus estrategias y desempeño de investigación, perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente mayor, los investigadores que señalan que sí debe reformarse el SNI o que el Sistema generó efectos negativos o neutros, perciben una Flexibilidad epistemológica del sistema de evaluación en promedio significativamente menor.

### IX.5 Dirección de la reforma del sistema de evaluación

En este trabajo la *Dirección de la reforma del sistema de evaluación del SNI* refiere al interés que tienen los investigadores en reformar el sistema de evaluación hacia una evaluación más cualitativa, que cambia la composición disciplinaria de las comisiones dictaminadoras y que valore mejor la investigación colectiva, inter/multi/transdisciplinaria y aplicada u orientada a la resolución de problemas dependientes del contexto.

Tal como fue definida en el capítulo V, la *Dirección de la reforma del sistema de evaluación* es el interés que tienen los investigadores categorizados en el sistema de evaluación de reformarlo en una dirección que cambia la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, mejor consideración de la investigación multi/inter/transdisciplinaria y que valore más los resultados de la investigación aplicada u orientada a la resolución de problemas.

En la Tabla 20 se presenta la distribución porcentual de investigadores SNI, según su interés de reformar el sistema de evaluación según los ítems asociados a la dirección de la reforma, a saber: i) reformular los criterios de evaluación; ii) modificar la composición disciplinaria de las comisiones dictaminadoras por área; iii) dar más peso a la evaluación cualitativa; iv) dar mayor reconocimiento a la investigación colectiva; v) considerar adecuadamente la investigación multi/inter/transdisciplinaria; y vi) dar más peso a los resultados de la investigación aplicada.

El 69% de los investigadores SNI considera que hay que reformular los criterios de evaluación, el 56.1% indica que hay que modificar la composición disciplinaria de las comisiones dictaminadoras por área y el 57.2% sostiene que es necesario dar más peso a la evaluación cualitativa del desempeño de los investigadores. (ver Tabla 20).

*Tabla 20: Distribución porcentual de investigadores SNI, según su interés de reformar el sistema de evaluación según los ítems asociados a la dirección de la reforma (Dirección de la Reforma del sistema de evaluación del SNI)*

Ítems	No	Sí	Total
Dar mayor reconocimiento a la investigación colectiva	48.46	51.54	100.00
Considerar adecuadamente la investigación multi/inter/transdisciplinaria	32.97	67.03	100.00
Dar más peso a los resultados de la investigación aplicada	70.06	29.94	100.00
Reformular los criterios de evaluación	31.01	68.99	100.00
Modificar la composición disciplinaria de las comisiones dictaminadoras por área	43.91	56.09	100.00
Dar más peso a la evaluación cualitativa	42.76	57.24	100.00

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Ese patrón de respuesta es muy similar cuando se observan los resultados según las diversas variables de control, tanto por la afirmativa como por la negativa. De hecho, los grupos de investigadores que en mayor proporción indican que es preciso reformular los criterios de evaluación, también son los que en mayor proporción indican que es necesario modificar la composición disciplinaria de las comisiones dictaminadoras. Similarmente, los grupos de investigadores que en mayor proporción indican que no hay que reformular los criterios de evaluación, también son los que en mayor proporción indican que no es necesario modificar la composición disciplinaria de las comisiones dictaminadoras.

Mientras que el 78.9% de los investigadores de Humanidades y Ciencias de la Conducta, el 72.6% de los de Ciencias Sociales, el 72.4% de los de Ingeniería, el 73.7% de los investigadores que trabajan en los CPI de Ciencias Exactas y Naturales, el 77% de los que lo hacen en los CPI de Ciencias Sociales y Humanidades y el 84.4% de los que desempeñan actividades en los CPI de Desarrollo Tecnológico y Servicios indican en mayor proporción que hay que reformular los criterios de evaluación; el 42.4% de los Nivel III, el 37.1% de los investigadores de Física, Matemáticas y Ciencias de la Tierra, el 40.1% de los de Medicina y Ciencias de la Salud, el 40.6% de los que trabajan en el Gobierno y el 37.3% de los que lo hacen en los Institutos Nacionales de Salud son los que en mayor proporción indican lo contrario.

Similarmente, los que en mayor proporción señalan que es preciso modificar la composición disciplinaria de las comisiones dictaminadoras por área son: el 66.8% de los investigadores de Humanidades y Ciencias de la Conducta, el 62.4% de los de Ciencias Sociales, el 59.3% de los de Ingeniería, el 68.9% de los que trabajan en los CPI de Ciencias Sociales y Humanidades y el 83.6% de los que se desempeñan en los CPI de Desarrollo Tecnológico y Servicios. Nuevamente aquí, el 55.8% de los Nivel III, el 50.5% de los investigadores de Física, Matemáticas y Ciencias de la Tierra, el 53.4% de los de Medicina y Ciencias de la Salud, el 54.9% de los que trabajan en el Gobierno y el 49.4% de los que lo hacen en los Institutos Nacionales de Salud son los que señalan en mayor proporción que no hay que modificar la composición disciplinaria de las comisiones dictaminadoras por área.

El interés de reformular los criterios de evaluación es mayor cuando se controla según los investigadores que no ascendieron en las categorías del SNI (82.6%), los que perciben efectos negativos o neutros del SNI (95.2% y 85.7% respectivamente) y los que consideran que es necesario reformar el sistema de evaluación (92.6%). En contraposición, el 38.8% de los investigadores que avanzaron en las categorías del SNI y el 36.1% de los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación consideran que no hay que reformular los criterios de evaluación.

Mientras que el 73.4% de los investigadores que no ascendieron en el SNI, el 83.6% de los que perciben efectos negativos sistema, el 72.5% de los que no perciben ningún efecto del SNI y el 77% de los que consideran necesario reformar el sistema, sostienen que hay que modificar la composición disciplinaria de las comisiones dictaminadoras por área, el 53.7% de los que subieron de categoría en el sistema y el 36.1% de los que perciben efectos positivos del SNI consideran lo contrario.

En relación al cambio en el tipo de evaluación, los investigadores que en mayor proporción indican que hay que dar más peso a la evaluación cualitativa son el 61.9% de los Candidatos, el 76.7% de los investigadores pertenecientes a Humanidades y Ciencias de la Conducta, el 67.5% de los de Ciencias Sociales, el 70.7% de los investigadores que trabajan en los CPI de Ciencias Sociales y Humanidades, el 59.9% de los que trabajan en Universidades, el 60.8% de la mujeres, el 62% de los investigadores más grandes (55 años y más) y el 60% de los que han sido evaluadores. La percepción sobre la necesidad de dar más peso a la evaluación cualitativa es mayor entre los investigadores que no ascendieron en el SNI (69.7%), los que perciben efectos negativos o neutros del SNI (82.6% y 71.3% respectivamente) y los que consideran que es necesario reformar el sistema de evaluación (77.1%).

Las opiniones de los investigadores se dividen aún más cuando son consultados por el cambio en el reconocimiento y valoración que el sistema de evaluación hace de la producción de conocimiento colectiva. Poco más de la mitad (51.5%) de los investigadores categorizados en el SNI considera que hay que dar mayor reconocimiento a la investigación colectiva, mientras que el 48.5% restante indica exactamente lo contrario (ver Tabla 20).

La percepción de reformar el SNI en una dirección que valore más la investigación colectiva se reduce conforme aumenta la edad y el nivel de categorización de los investigadores. Mientras que el 54.6% de los investigadores más jóvenes (menores de 45 años) y el 60.5% de los Candidatos señala que hay que dar mayor reconocimiento a la investigación colectiva, el 51.7% de los investigadores más grandes (50 años o más), el 54% de los Nivel II y el 64.7% de los Nivel III consideran lo contrario.

Adicionalmente, mientras que el 56.2% de las mujeres indica que hay que dar mayor reconocimiento a la investigación colectiva, el 51% de los hombres señala lo contrario.

Los investigadores pertenecientes a las áreas de Humanidades y Ciencias de la Conducta (56.7%), Ciencias Sociales (55.4%), Biotecnologías y Ciencias Agropecuarias (52.6%) e Ingenierías (55.9%), y los que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades (61.4) y los que lo hacen en los CPI de Desarrollo Tecnológico y Servicios (74.2%) son los que en mayor proporción consideran que el SNI debe reformarse en la dirección de dar mayor reconocimiento a la investigación colectiva. Por su parte, los que en mayor proporción indican que no hay que dar mayor reconocimiento a la investigación colectiva son: el 55.5% de los investigadores pertenecientes a Física, Matemáticas y Ciencias de la Tierra, el 51.5% de los de Biología y Química, el 52.6% de los de Medicina y Ciencias de la Salud, el 54.5% de los que se desempeñan en el Gobierno y el 53.2% de los que lo hacen en los Institutos y Universidades Tecnológicas.

Mientras que el 53% de los investigadores que avanzaron en las categorías del SNI y el 51.2% de los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación consideran que no hay que dar mayor reconocimiento de la investigación colectiva, el 59.8% de los que no ascendieron en el sistema, el 71.3% de los que perciben efectos negativos, el 58.2 de los que perciben efectos neutros y el 69.7% de los que quieren reformar el SNI, indican que el cambio debe ir en la dirección de dar mayor reconocimiento a la investigación colectiva.

La amplia mayoría de los investigadores coincide cuando son consultados por el cambio en el reconocimiento y valoración que el sistema de evaluación hace de la producción de conocimiento inter/multi/transdisciplinar. De hecho, mientras que el 67% de los investigadores SNI señalan que el SNI debe reformarse en la dirección de considerar adecuadamente la investigación multi/inter/transdisciplinaria, solo el 33% restante considera lo contrario (ver Tabla 20).

La percepción de reformar el SNI en una dirección que considere adecuadamente la investigación inter/multi/transdisciplinar se reduce a medida que aumenta el nivel de categorización de los investigadores. Mientras que el 70.4% de los Candidatos señala que el SNI debe reformarse en la dirección de considerar adecuadamente la investigación inter/multi/transdisciplinar, el 45.5% de los investigadores Nivel III indica exactamente lo contrario. Entre las mujeres la percepción de reformar el SNI en una dirección que considere adecuadamente la investigación inter/multi/transdisciplinar es mayor que entre los hombres. Mientras que el 70.9% de las mujeres indican que el SNI debe reformarse en una dirección que valore adecuadamente la investigación inter/multi/transdisciplinar, esa proporción se reduce a 64.9% entre los hombres.

Los investigadores pertenecientes a las áreas de Humanidades y Ciencias de la Conducta (76.3%), Ciencias Sociales (72.3%) e Ingenierías (70.3%), y los que desempeñan sus actividades en los CPI de Ciencias Sociales y Humanidades (76.3%) y los que lo hacen en los CPI de Desarrollo Tecnológico y Servicios (80%) son los que en mayor proporción consideran que el SNI debe reformarse en la dirección de considerar adecuadamente la investigación inter/multi/transdisciplinar. Por su parte, los que en mayor proporción indican que el SNI debe reformarse pero no en esa dirección son: el 38.6% de los investigadores pertenecientes a Física, Matemáticas y Ciencias de la Tierra, el 36.7% de los de Biología y Química, el 41% de los de Medicina y Ciencias de la Salud, el 38.7% de los que se desempeñan en el Gobierno, el 37.8% de los que lo hacen en los Institutos y Universidades Tecnológicas y el 38.7% de los que trabajan en los Institutos Nacionales de Salud.

El 39.9% de los investigadores que avanzaron en las categorías del SNI y el 36.9% de los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación consideran que el SNI debe reformarse pero no en la dirección de considerar adecuadamente la investigación inter/multi/transdisciplinar. Por el contrario, el 79.5% de los que no ascendieron en el sistema, el 88.7% de los que perciben efectos negativos, el 78.7% de los que perciben efectos neutros y el 90.4% de los que quieren reformar el SNI, indican que el cambio del SNI debe ir en la dirección de considerar adecuadamente -en la evaluación del desempeño de los investigadores- la investigación inter/multi/transdisciplinar.

En relación con el cambio en el reconocimiento y valoración que el sistema de evaluación hace de la producción de conocimiento cuando su orientación predominante es la aplicación, menos de la tercera parte (29.9%) de los investigadores coincide en que el SNI debe reformarse en la dirección de dar más peso a los resultados de la investigación aplicada en la evaluación del desempeño de los individuos (ver Tabla 20). Por su parte, la amplia mayoría (70.1%) de los investigadores SNI consideran lo contrario.

Si bien la propuesta de dar mayor peso a los resultados de la investigación aplicada en la evaluación es la que recibe menor aceptación entre los investigadores, vale destacar que este resultado varía si se considera a los investigadores pertenecientes a los campos científicos de Ingeniería y Biotecnología y Ciencias Agropecuarias, y a los investigadores que desempeñan sus actividades en los CPI de Ciencias Exactas y Naturales, de Desarrollo Tecnológico y Servicios, en los CPI Sectorizados y en los Institutos y Universidades Tecnológicas.

De hecho, casi la mitad (49.3%) de los investigadores de Ingeniería y algo más de la tercera parte (35.3%) de los investigadores de Biotecnología y Ciencias Agropecuarias consideran que en la evaluación de su desempeño sí se le debe dar mayor peso a los resultados de la investigación aplicada. Asimismo, más de las dos terceras partes (67.7%) de los investigadores que se desempeñan en los CPI de Desarrollo Tecnológico y Servicios, casi la mitad (49.6%) de los investigadores que lo hacen en los CPI Sectorizados, y algo más del 35% de los que trabajan en los CPI de Ciencias Exactas y Naturales o en los Institutos y Universidades Tecnológicas también consideran que en la evaluación del desempeño de los investigadores se le debe dar mayor peso a los resultados de la investigación aplicada. Esto sugiere la necesidad de ponderar la producción de conocimiento de acuerdo a sus objetivos y resultados de investigación, al mismo tiempo que implica entenderla de acuerdo al contexto en el que se produce el conocimiento. Finalmente, el 41.1% de los investigadores que quieren reformar el SNI, indican que el cambio debe ir en la dirección de dar mayor peso a los resultados de la investigación aplicada.

Al igual que en la construcción de los índices de propensión Modo 1 y Modo 2, Satisfacción con el Proceso de Evaluación y el que mide la Flexibilidad Epistemológica del sistema de evaluación del SNI, aquí también, con los ítems que se exploraron las opiniones de los investigadores en relación al cambio del SNI en cuanto a los criterios y tipo de evaluación, y la mejor valoración de la investigación no disciplinar, colaborativa y orientada por la búsqueda de soluciones a problemas específicos (investigación aplicada), se creó el *índice de Dirección de la Reforma del sistema de evaluación SNI*. De la misma manera que en los índices anteriores, este índice refiere a las puntuaciones factoriales estandarizadas de los individuos. Las puntuaciones factoriales fueron estandarizadas para asegurar la comparabilidad con los otros índices y facilitar la interpretación de



sus valores. Simultáneamente, y a partir de la subdivisión en cuatro tramos de la distribución del índice, se construyó una nueva variable con cuatro categorías o niveles del índice de *Dirección de la reforma del sistema de evaluación*, a saber: “Muy bajo interés”, “Bajo interés”, “Alto interés” y “Muy alto interés”. Los puntos de corte para establecer cada uno de esos cuatro niveles se determinaron de acuerdo a la consistencia entre el valor estandarizado de la puntuación factorial y las respuestas a los ítems de los investigadores.

Las opiniones de los investigadores categorizados en el SNI en relación a la *Dirección de la reforma del sistema de evaluación* están divididas. Aunque la mayoría (58.6%) de los investigadores tienen un *Muy bajo o Bajo interés* de reformar el SNI en la dirección de reformular los criterios de evaluación, modificar la composición disciplinaria de las comisiones dictaminadoras por área, dar más peso a la evaluación cualitativa, dar mayor reconocimiento a la investigación colectiva, considerar adecuadamente la investigación multi/inter/transdisciplinaria, y dar más peso a los resultados de la investigación aplicada; una proporción importante de investigadores está interesada en esa *Dirección de la reforma del sistema de evaluación* (41.4%). En particular, para el 33.5% de los investigadores categorizados, esa *Dirección de la reforma del sistema de evaluación* es de *Alto interés* y para el 7.9% de *Muy Alto interés*.

Si bien estas proporciones se mantienen relativamente similares, es posible observar algunas variaciones cuando se controla por ciertas variables. A continuación se analizan esos resultados junto con los intervalos de confianza para el valor promedio que asume el índice.

De acuerdo con los intervalos de confianza para el valor promedio del índice de *Dirección de la reforma del sistema de evaluación* según las diversas variables de control, es posible constatar con un 95% de confianza, que los investigadores pertenecientes a los campos científicos de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías y los que trabajan en los CPI de Ciencias Sociales y Humanidades y en los CPI de Desarrollo Tecnológico y Servicios son los que con un promedio significativamente mayor están interesados en reformar el SNI en la dirección de reformular los criterios de evaluación, modificar la composición disciplinaria de las comisiones dictaminadoras por área, dar más peso a la evaluación cualitativa, dar mayor reconocimiento a la investigación colectiva, considerar adecuadamente la investigación multi/inter/transdisciplinaria, y dar más peso a los resultados de la investigación aplicada. Por su parte, los investigadores Nivel III, los hombres, los de Física, Matemáticas y Ciencias de la Tierra, los de Biología y Química, los de Medicina y Ciencias de la Salud, y los que desempeñan sus actividades en el Gobierno y en los Institutos Nacionales de Salud son los que perciben esa *Dirección de la reforma del sistema de evaluación* con un interés en promedio significativamente menor.

Mientras que el 53.4% de los investigadores de Humanidades y Ciencias de la Conducta, el 46.7% de los de Ciencias Sociales, el 45.8% de los de Ingenierías, el 54.6% de los que trabajan en los CPI de Ciencias Sociales y Humanidades y el 55.9% de los que lo hacen en los CPI de Desarrollo Tecnológico y Servicios muestran en mayor proporción un *Alto o Muy Alto interés* de esa *Dirección de la reforma del sistema de evaluación*; el 66.8% de los investigadores Nivel III, el 59.7% de los hombres, el 65.6% de los de Física, Matemáticas y Ciencias de la Tierra, el 63.3% de los de Biología y Química, el 66.3% de los de Medicina y Ciencias de la Salud, el 67.5% de los que trabajan en el Gobierno y el 63.9% de los que lo hacen en los Institutos Nacionales de Salud son los que muestran en mayor proporción un *Bajo o Muy Bajo interés* de la reforma del SNI en la dirección señalada.

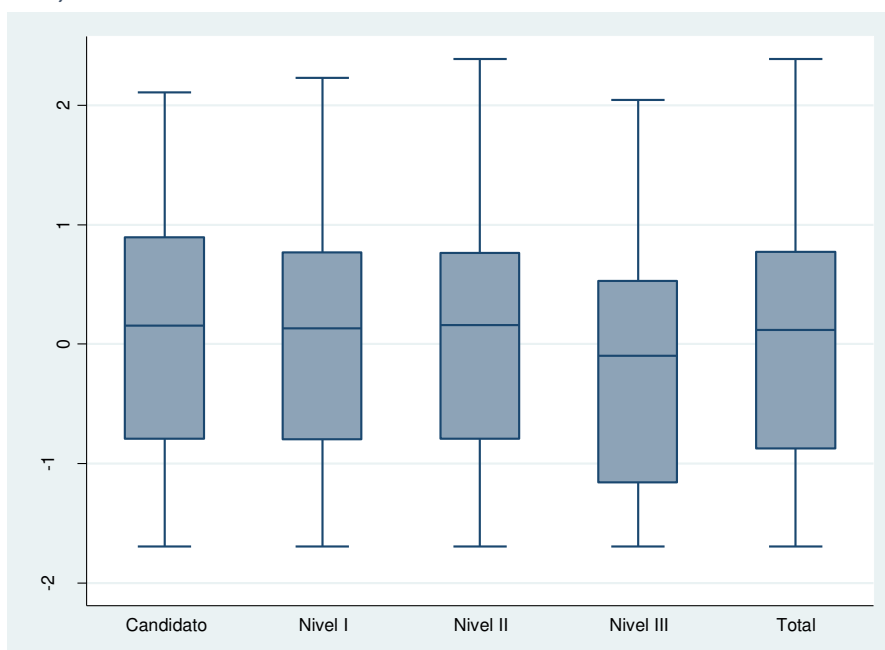
Finalmente, los investigadores que ascendieron en las categorías del SNI y los que perciben efectos positivos del sistema sobre su estrategia y desempeño de investigación exhiben en promedio un *índice de Dirección de la reforma del sistema de evaluación* significativamente menor que los investigadores que no han ascendido en el sistema, que los que sí quieren reformar y los que perciben efectos negativos o neutros del SNI. De hecho, mientras que el 63.2% de los investigadores que perciben efectos positivos del SNI y el 66.2% de los que ascendieron en los niveles del SNI se ubican en las categorías de *Bajo o Muy Bajo interés* en la *Dirección de la Reforma del sistema de evaluación del SNI* señalada, el 66.5% de los investigadores que perciben efectos negativos o neutros y el 56.7% de los que no ascendieron se ubican en el lado opuesto del índice de *Dirección de la reforma del sistema de evaluación*, mostrando un *Muy alto o Alto interés*. Por último, el 44.4% de los investigadores que han sido evaluadores y el 63% de los que quieren reformar el SNI, coinciden en un Alto o Muy Alto interés en la *Dirección de la reforma del sistema de evaluación* señalada.

En síntesis, el interés de los investigadores categorizados en el SNI en la Dirección de la Reforma del sistema de evaluación del SNI está dividido. Mientras que el 41.4% de los investigadores tiene un Muy alto o Alto interés de reformar el SNI en la dirección de reformular los criterios de evaluación, modificar la composición disciplinaria de las comisiones dictaminadoras por área, dar más peso a la evaluación cualitativa, dar mayor reconocimiento a la investigación colectiva, considerar adecuadamente la investigación multi/inter/transdisciplinaria, y dar más peso a los resultados de la investigación aplicada; el 58.6% restante tiene un Bajo o Muy bajo interés en esa Dirección de la reforma del sistema de evaluación.

El interés en esa Dirección de la Reforma del sistema de evaluación del SNI varía dependiendo de ciertas características diferenciales de los investigadores, tales como: las de desempeño (nivel de categorización, ascenso en el sistema y si han sido evaluadores), las asociadas a la pertenencia cognitiva y laboral (área de conocimiento y tipo de institución) y las vinculadas al tipo de efectos generados por el SNI en la estrategia y desempeño de los investigadores.

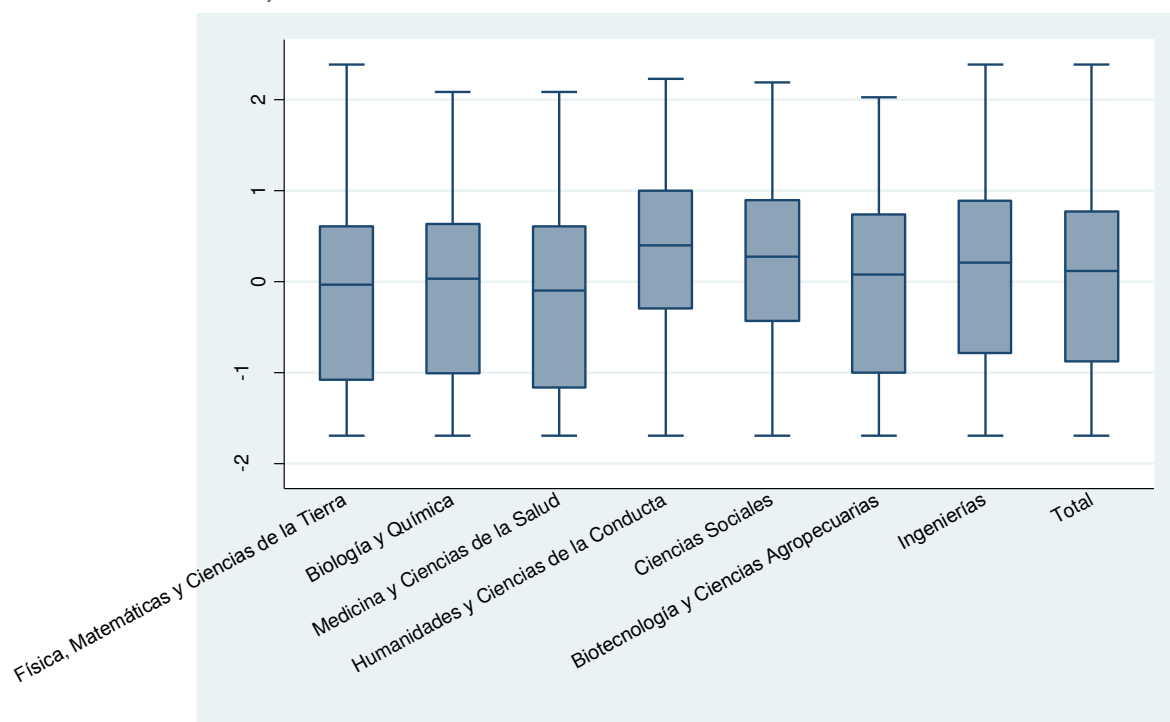
Desde el punto de vista del desempeño y el género de los investigadores se observa –de acuerdo con lo esperado y consistentemente con los resultados de los índices de Satisfacción con el Proceso de Evaluación y de Flexibilidad epistemológica del sistema de evaluación- que los Nivel III, los hombres y los investigadores que ascendieron de categoría en el Sistema son los que presentan un interés en promedio significativamente menor en la Dirección de la Reforma del sistema de evaluación del SNI. Por su parte, las mujeres, los investigadores que no ascendieron en el SNI y los que han participado como evaluadores, ya sea en el SNI o en otras instancias de evaluación de la investigación, son los que exhiben un interés en promedio significativamente mayor en la Dirección de la Reforma del sistema de evaluación del SNI. Estos resultados muestran que los nivel III -la élite de investigadores- son los que en mayor medida consideran que no se debe reformar ni el proceso y tipo de evaluación, ni cambiar el reconocimiento y valoración que el sistema de evaluación hace de la producción de conocimiento colectiva, aplicada e inter/multi/transdisciplinaria, probablemente porque así como están les permitió no solo avanzar sino que alcanzar la categoría más alta del sistema. A la vez, estos resultados sugieren que los investigadores con experiencia de evaluación perciben esos cambios muy positivamente. En el gráfico 15 se puede observar la distribución del índice de Dirección de la reforma del sistema de evaluación para el total y por nivel de categorización.

Gráfico 15: Distribución del índice de Dirección de la reforma del sistema de evaluación (Total y por Nivel)



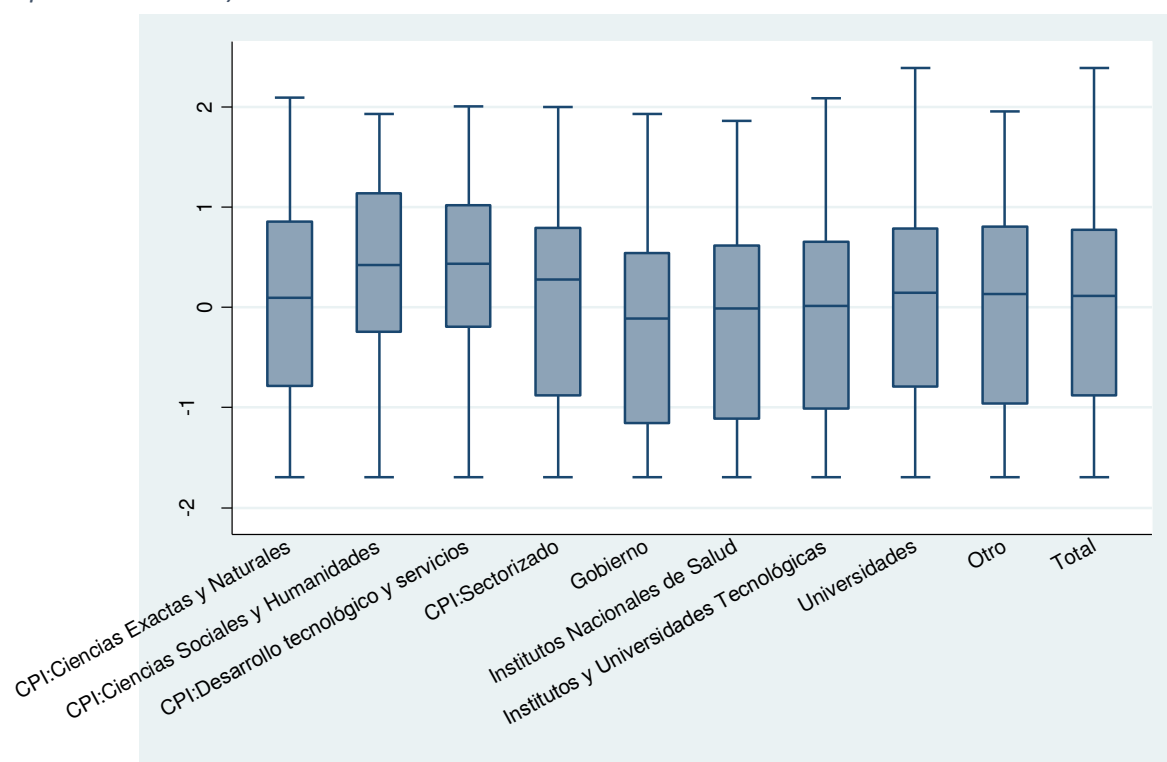
De acuerdo con lo esperado y consistentemente con los resultados de los índices de Satisfacción con el Proceso de Evaluación y de Flexibilidad epistemológica del sistema de evaluación, los investigadores pertenecientes a las áreas de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías son los que exhiben un interés en promedio significativamente mayor en la Dirección de la reforma del sistema de evaluación propuesta. Por su parte, los investigadores que muestran un interés en promedio significativamente menor en dicha Dirección de la reforma del sistema de evaluación son los pertenecientes a las áreas de Física, Matemáticas y Ciencias de la Tierra, Biología y Química y Medicina y Ciencias de la Salud. Estos resultados sugieren que el interés en la dirección de la reforma del sistema de evaluación depende de las características específicas de las áreas de conocimiento. De acuerdo con lo esperado y la literatura, los investigadores de las áreas que exhiben mayor interés en la dirección de la reforma del sistema de evaluación son justamente aquellas cuyas prácticas de producción y difusión del conocimiento difieren de las ciencias exactas y naturales, y cuyos productos son relativamente menos valorados en el sistema de evaluación del SNI. El gráfico 16 muestra la distribución del índice de Dirección de la reforma del sistema de evaluación para el total y por área de conocimiento.

Gráfico 16: Distribución del índice de Dirección de la reforma del sistema de evaluación (Total y por Área de conocimiento)



Desde el punto de vista del ámbito laboral, de acuerdo con lo esperado y consistentemente con los resultados de los índices de Satisfacción con el Proceso de Evaluación y Flexibilidad epistemológica del sistema de evaluación; los investigadores que trabajan en los CPI de Ciencias Sociales y Humanidades y en los CPI de Desarrollo Tecnológico y Servicios son los que exhiben un interés en promedio significativamente mayor en la Dirección de la reforma del sistema de evaluación propuesta. Por su parte, los investigadores que muestran un interés en promedio significativamente menor en dicha Dirección de la reforma del sistema de evaluación son los investigadores que trabajan en el Gobierno y en los Institutos Nacionales de Salud. El gráfico 17 muestra la distribución del índice de Dirección de Reforma del SNI para el total y por el tipo de institución en la que trabajan los investigadores.

Gráfico 17: Distribución del índice de Dirección de la reforma del sistema de evaluación (Total y por Tipo de Institución)



Finalmente, los investigadores que perciben efectos negativos o neutros del SNI sobre su estrategia y desempeño de investigación y los que consideran que el mecanismo de evaluación del SNI debe reformarse, son los que exhiben un interés en promedio significativamente mayor en la Dirección de la reforma del sistema de evaluación propuesta. Por su parte, los investigadores que perciben efectos positivos del SNI y, como es esperable, los que consideran que el mecanismo de evaluación del SNI no debe reformarse, muestran un interés en promedio significativamente menor en la Dirección de la reforma del sistema de evaluación propuesta.

## IX.6 Conclusión

La amplia mayoría (77%) de los investigadores pertenecientes al SNI mexicano exhiben propensiones Modo 1 y Modo 2 de producción de conocimiento *Alta o Muy Alta*. Estos resultados sugieren que una parte importante de la comunidad académica y científica de México mayoritariamente coincide o legitima ambas formas de producción de conocimiento. Sin embargo, las propensiones Modo 1 y Modo 2 promedio varían de acuerdo a ciertas características diferenciales de los investigadores.

Los investigadores hombres, los de 50 años y más y los categorizados en el Nivel III son los que exhiben una mayor propensión Modo 1 en promedio. Por el contrario, las investigadoras mujeres, los más jóvenes (menores de 35 años), los Candidatos y los Nivel I muestran una mayor propensión Modo 2 en promedio.

Estos resultados sugieren que, conforme avanza la edad de los investigadores y sus niveles de categorización, hay una cierta adaptación de los investigadores respecto a los modos que

predominantemente orientan la producción de conocimiento que realizan. En particular, hacia una mayor propensión modo 1. Esto podría sugerir que el sistema de evaluación del SNI reconoce y valora mejor el Modo 1 de producción de conocimiento que el Modo 2, ya que el nivel de categorización más alto del sistema se asocia con una mayor propensión modo 1. Los investigadores aprenden de las sucesivas evaluaciones del SNI a las que se someten, respecto de lo que el sistema reconoce y valora, y adaptan su comportamiento en materia de producción de conocimiento para avanzar en las categorías del SNI, ya que de ello depende su prestigio académico, buena parte de su ingreso mensual y el mejor acceso a recursos externos para financiar la investigación que realizan. Esta explicación es consistente, ya que los investigadores que han ascendido en el nivel de categorización del sistema y más aún, los que se han desempeñado como evaluadores, también exhiben una mayor propensión modo 1 en promedio.

La importancia relativa de las propensiones Modo 1 y Modo 2 también varía dependiendo del área de conocimiento y el tipo de institución/organización en la que trabajan los investigadores. Mientras que los investigadores del área de Biología y Química presentan una mayor propensión modo 1 en promedio, los investigadores de Ingenierías, Biotecnologías y Ciencias Agropecuarias y Ciencias Sociales, y los que trabajan en los CPI Sectorizados y en los CPI de Desarrollo Tecnológico y Servicios exhiben una mayor propensión modo 2 en promedio. Este resultado sugiere que las características específicas que asuma la producción de conocimiento están asociadas a los campos científicos en el que los investigadores realizan sus investigaciones y al contexto institucional/organizacional al que pertenecen -objetivos y misiones institucionales/organizacionales-.

La excepción a esa regla son los investigadores de Medicina y Ciencias de la Salud y los que trabajan en el Gobierno o en los Institutos Nacionales de Salud ya que presentan propensiones significativamente mayores en promedio tanto en Modo 1 como en Modo 2. Esto puede explicarse por el hecho de que en dichas categorías se engloba a grupos muy heterogéneos de investigadores en relación a sus intereses, motivaciones, objetivos y orientación de la investigación que realizan.

Aunque la mayoría (65.3%) de los investigadores pertenecientes al SNI mexicano exhiben una Satisfacción con el Proceso de Evaluación del SNI *Alta* o *Muy Alta*, una proporción importante –el restante 34.7%- tiene una satisfacción *Baja* o *Muy Baja*. El nivel de Satisfacción con el Proceso de Evaluación del SNI también varía dependiendo de ciertas características diferenciales de los investigadores.

Los investigadores que tienen una mayor satisfacción con el proceso de evaluación del SNI en promedio son los categorizados en el Nivel III, los que ascendieron de categoría en el SNI, los que pertenecen a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física-Matemáticas y Ciencias de la Tierra, y Biotecnología y Ciencias Agropecuarias, los que trabajan en el Gobierno y en los Institutos Nacionales de Salud, los que consideran que no hay que reformar el SNI y los que perciben efectos positivos sobre su estrategia y desempeño de investigación.

Por el contrario, los investigadores que tienen una menor satisfacción con el proceso de evaluación del SNI en promedio son los categorizados en el Nivel II, los que no ascendieron de categoría en el SNI, los que pertenecen a las áreas de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías, los que trabajan en los CPI de Ciencias Sociales y Humanidades y en los CPI de Ciencias Exactas y Naturales, los que consideran que sí hay que reformar el SNI y los que perciben efectos negativos o neutros sobre su estrategia y desempeño de investigación. Estos resultados sugieren

que el proceso de evaluación del SNI no es capaz de reconocer de igual manera el desempeño de investigación de los investigadores de las diversas áreas de conocimiento y según el contexto institucional/organizacional al que pertenecen. En particular, estos resultados sugieren que el proceso de evaluación del SNI tiene dificultades para reconocer el desempeño de los investigadores de las áreas de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías, y de los que trabajan en los CPI de Ciencias Sociales y Humanidades y en los CPI de Ciencias Exactas y Naturales.

Aunque la mayoría (63%) de los investigadores pertenecientes al SNI mexicano perciben una *Flexibilidad Epistemológica del sistema de evaluación del SNI Alta o Muy Alta*, una proporción importante –el restante 37%– percibe una flexibilidad epistemológica *Baja o Muy Baja*. La flexibilidad epistemológica del sistema de evaluación, que perciben los investigadores, también varía según algunas características diferenciales de ellos.

Los investigadores que perciben una mayor flexibilidad epistemológica del sistema de evaluación del SNI en promedio son los categorizados en el Nivel III, los más jóvenes (menores de 35 años), los pertenecientes a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra y Biotecnologías y Ciencias Agropecuarias y los que trabajan en el Gobierno y en los Institutos Nacionales de Salud, los que ascendieron de categoría en el SNI, los que consideran que no hay que reformar el SNI y los que perciben efectos positivos sobre su estrategia y desempeño de investigación.

Por el contrario, los investigadores que perciben una menor flexibilidad epistemológica del sistema de evaluación del SNI en promedio son los categorizados en el Nivel II, los de 60 años y más, las mujeres, los que pertenecen a las áreas de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías, los que trabajan en los CPI de Ciencias Sociales y Humanidades, los que no ascendieron de categoría en el SNI, los que consideran que sí hay que reformar el SNI y los que perciben efectos negativos o neutros sobre sus trayectorias de investigación. Estos resultados sugieren que el sistema de evaluación del SNI no es igualmente flexible desde el punto de vista de las características epistemológicas que permite de los procesos de producción de conocimiento realizados por los investigadores de las diversas áreas de conocimiento y según el contexto institucional/organizacional al que pertenecen.

Las características de los investigadores que tienen mayor satisfacción con el proceso de evaluación del SNI son las mismas que las que caracterizan a los investigadores que perciben mayor flexibilidad epistemológica del sistema de evaluación del SNI, y que además, los investigadores que tienen menor satisfacción comparten las mismas características con aquellos que perciben menor flexibilidad.

Simultáneamente, y en línea con lo argumentado respecto a que la menor satisfacción de los investigadores con el proceso de evaluación se asocia con las dificultades que éste tiene para reconocer el desempeño de los investigadores de las áreas de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías, y de los que trabajan en los CPI de Ciencias Sociales y Humanidades, es probable que la menor flexibilidad epistemológica del sistema de evaluación que perciben los investigadores de esas áreas y que trabajan en esos CPI esté asociada a que el sistema de evaluación no permite o tiene dificultades para permitir la diversidad de perfiles de investigación. Este argumento resulta convincente o se refuerza al observar que los investigadores que no ascendieron en las categorías del SNI son también los que en promedio tienen menor satisfacción

con el proceso de evaluación y perciben una menor flexibilidad epistemológica del sistema de evaluación.

Finalmente, el interés de los investigadores categorizados en el SNI en la dirección de la reforma del sistema de evaluación del SNI está dividido. Mientras que el 41.4% de los investigadores tiene un *Muy alto* o *Alto interés* de reformar el SNI en la dirección de cambiar la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, mejor consideración de la investigación multi/inter/transdisciplinaria y que valore más los resultados de la investigación aplicada u orientada a la resolución de problemas; el 58.6% restante tiene un *Bajo* o *Muy bajo interés* en esa dirección de la reforma del sistema de evaluación del SNI. Nuevamente aquí, y como es razonable esperar, el interés en esa dirección de la reforma del sistema de evaluación del SNI varía dependiendo de ciertas características diferenciales de los investigadores.

Los investigadores que exhiben mayor interés en esa dirección de la reforma del sistema de evaluación del SNI en promedio son las investigadoras mujeres, los investigadores pertenecientes a los campos científicos de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías, los que trabajan en los CPI de Ciencias Sociales y Humanidades y en los CPI de Desarrollo Tecnológico y Servicios, los que no han ascendido en las categorías del SNI, los que han participado como evaluadores y los que perciben efectos negativos en su estrategia y desempeño de investigación.

Contrariamente, los investigadores que exhiben menor interés en esa dirección de la reforma del sistema de evaluación del SNI en promedio son los investigadores categorizados en el Nivel III, los hombres, los que pertenecen a las áreas de Medicina y Ciencias de la Salud, Biología y Química, Física, Matemáticas y Ciencias de la Tierra, los que trabajan en el Gobierno y en los Institutos Nacionales de Salud, los que ascendieron de categoría en el SNI, y los que perciben efectos positivos sobre su estrategia y desempeño de investigación.

Estos resultados muestran que el interés de los investigadores de reformar el sistema de evaluación del SNI en la dirección planteada se asocia con las características específicas de las diversas áreas de conocimiento. En sintonía con los hallazgos de investigaciones pasadas reportados en la literatura (ver capítulo IV), los investigadores de las áreas que exhiben mayor interés en esta dirección de la reforma del sistema de evaluación son justamente aquellas cuyas prácticas de producción y difusión del conocimiento difieren de las ciencias exactas y naturales, y cuyos productos son relativamente menos valorados en el sistema de evaluación del SNI. A la vez, son los nivel III, que representan a la élite científica, los que exhiben en promedio un menor interés en cambiar el tipo de evaluación hacia uno más cualitativo y que reconozca y valore mejor la producción de conocimiento colectiva, aplicada e inter/multi/transdisciplinaria, probablemente porque así como está les permitió no solo avanzar en su trayectoria académica dentro del sistema, sino que además les permitió alcanzar la categoría más alta dentro del sistema.

Estos resultados son consistentes con los obtenidos en relación a la flexibilidad epistemológica del sistema de evaluación percibida y la satisfacción con el proceso de evaluación que tienen los investigadores. Los investigadores que exhiben menor interés en esa dirección de la reforma del sistema de evaluación del SNI comparten características con los que perciben mayor flexibilidad epistemológica y con los que tienen mayor satisfacción con el proceso de evaluación del SNI, a saber: los investigadores Nivel III, los que pertenecen a las áreas de Medicina y Ciencias de la Salud, Biología



y Química, Física, Matemáticas y Ciencias de la Tierra, los que trabajan en el Gobierno y en los Institutos Nacionales de Salud, los que ascendieron de categoría en el SNI, y los que perciben efectos positivos sobre su estrategia y desempeño de investigación. A la vez, los investigadores que exhiben mayor interés en esa dirección de la reforma del sistema de evaluación del SNI comparten características con los que perciben menor flexibilidad epistemológica y con los que también tienen menor satisfacción con el proceso de evaluación del SNI, a saber: los investigadores pertenecientes a los campos científicos de Humanidades y Ciencias de la Conducta, Ciencias Sociales e Ingenierías, los que trabajan en los CPI de Ciencias Sociales y Humanidades, los que no han ascendido en las categorías del SNI y los que perciben efectos negativos en su estrategia y desempeño de investigación.

Atendiendo a todos estos resultados, se puede argumentar -en sintonía con la literatura- que la dirección que eventualmente asuma una reforma al sistema de evaluación del SNI, será un terreno en disputa y probablemente dependerá de la correlación de fuerzas en las relaciones de poder que gobiernan la evaluación del Sistema. Adicionalmente, importa remarcar que estos resultados también sugieren consistencia con lo planteado por diversos autores respecto a que desde la creación del SNI hasta la actualidad, dicho sistema ha sido y sigue siendo gobernado por las normas y prácticas que rigen a la producción y difusión de conocimiento de las ciencias exactas y naturales o *ciencias duras* (ver capítulo VII). A la vez, estos resultados también sugieren consistencia con los hallazgos de investigaciones pasadas respecto a la desvalorización que hacen los sistemas de evaluación del desempeño de la investigación de los tipos y productos de la investigación que se apartan de los estándares y prácticas de producción de dichas ciencias (ver capítulo IV).

## CAPÍTULO X: Los Modelos de Medición y Estructural: Especificación, Confiabilidad y Validación

El objetivo central de este capítulo es presentar y analizar los resultados de investigación obtenidos en este trabajo para responder a la pregunta de investigación. En base a la teoría y la evidencia empírica obtenida, se validan los modelos de medida y estructural, y con ello se analizan las relaciones existentes entre las dimensiones asociadas con la evaluación del desempeño individual de investigación y los modos diferenciales de producción de conocimiento que predominantemente orientan la investigación que realizan los investigadores categorizados en el SNI mexicano. Para ello, a continuación se describe la estrategia de análisis seguida.

### X.1 Estrategia de Análisis

De acuerdo con la estrategia de investigación (capítulo VIII), la estrategia de análisis seguida en este trabajo se divide en tres grandes fases complementarias y consecutivas: i) análisis factorial exploratorio (AFE); ii) análisis factorial confirmatorio (AFC): el modelo de medida; y iii) análisis de las relaciones de dependencia estructurales para explicar las relaciones entre la evaluación de la investigación y los modos diferenciales de producción de conocimiento: el modelo teórico estructural.

Así, en la sección X.2 se implementa el AFE. Con esta técnica de análisis multivariado se explora la estructura de correlación –empírica- entre todas las variables de interés para saber si existen y cuántas son –identificar- las dimensiones conceptuales –factores- latentes capaces de explicar la estructura de correlación entre dichas variables. A partir de este análisis se busca establecer la base empírica para la construcción de una teoría de medida válida y confiable. En ese sentido y de acuerdo con los objetivos de investigación y la revisión de la literatura, con el AFE idealmente se espera identificar las cinco dimensiones definidas conceptualmente en este trabajo (*propensión modo 1, propensión modo 2, satisfacción con el proceso de evaluación, flexibilidad epistemológica del sistema de evaluación y dirección de la reforma del sistema de evaluación*) que podrían explicar la estructura de correlación entre las variables observadas. Si la solución del AFE coincide total o parcialmente con los resultados esperados, se procederá con la segunda fase para confirmarlos o para poner a prueba y validar un modelo para medir las relaciones entre la evaluación de la investigación y los modos diferenciales de producción de conocimiento.

En la segunda fase, atendiendo a los resultados obtenidos de la fase anterior y con una orientación predominantemente teórica, se ponen a prueba las hipótesis que se hacen sobre las relaciones entre las variables observadas seleccionadas para el análisis y las dimensiones conceptuales que las explican. Es decir, se pone a prueba y valida la teoría de medición, representada por el modelo de medida, para las relaciones entre la evaluación de la investigación y la producción de conocimiento. Para ello, en la sección X.3 se realiza un AFC. Con esta técnica de análisis multivariado se pone a prueba y valida el modelo de medida multifactorial de cinco factores construido para medir las relaciones entre la evaluación de la investigación del desempeño individual de los investigadores y los modos diferenciales de producción de conocimiento. Poner a prueba y validar el modelo de medición implica analizar, por un lado, su bondad de ajuste, y por otro, evaluar su validez de constructo. El resultado esperado de este análisis es lograr un modelo de medición válido y confiable. Si se alcanza este resultado esperado, se procede con la tercera fase de análisis: se estudian y validan las relaciones teóricas estructurales de dependencia entre las dimensiones

latentes para explicar la relación entre la evaluación de la investigación y la producción de conocimiento. Si dicho resultado esperado no se alcanza, es preciso regresar al inicio y revisar la teoría de medición que sustenta el trabajo, porque no se pueden extraer conclusiones rigurosas si no se basan en medidas válidas y confiables.

Finalmente, en la tercera fase, y luego de haber obtenido un modelo de medición válido y confiable, se pone a prueba y valida la teoría estructural propuesta para explicar las relaciones entre la evaluación del desempeño individual de investigación y los modos diferenciales de producción de conocimiento. La teoría estructural propuesta queda representada y formalizada por el modelo teórico estructural (secciones X.4 y X.5). A diferencia de la segunda fase, aquí el foco se pone en las hipótesis realizadas sobre las relaciones teóricas estructurales entre las variables latentes. Poner a prueba y validar la teoría que sustenta este trabajo implica, por una parte, analizar la bondad de ajuste para evaluar la aceptación o no del modelo teórico propuesto (ver figura 2 del capítulo V); y por otra parte, estudiar la naturaleza y significado conceptual de los parámetros estructurales estimados. Dichos parámetros representan las relaciones estructurales de dependencia entre los constructos que se hipotetizan. Si los resultados validan la teoría estructural que se propone, se habrá encontrado nueva evidencia -a partir del caso del SNI mexicano- que contribuye al entendimiento de las relaciones teóricas entre la evaluación de la investigación y los modos diferenciales de producción de conocimiento.

## X.2 Análisis Factorial Exploratorio: ítems, estructura de correlación y dimensionalidad

Como se dijo, en el contexto de este trabajo, el AFE se realiza justamente para explorar la estructura de correlación entre las variables seleccionadas e identificar las dimensiones conceptuales que las explican. En ese sentido y de acuerdo con los objetivos de investigación y la revisión de la literatura, idealmente se espera identificar las cinco dimensiones definidas conceptualmente en este trabajo (*propensión modo 1, propensión modo 2, satisfacción con el proceso de evaluación, flexibilidad epistemológica y dirección de la reforma del sistema de evaluación*) que podrían explicar las relaciones entre las variables observadas. Si con el AFE se obtienen los resultados esperados, se procederá con el AFC para confirmarlos. Si por el contrario, la solución del AFE devuelve parcialmente o no da el resultado esperado, se procederá a realizar el AFC para poner a prueba y validar el modelo de medida teórico-conceptual que sustenta este trabajo.

Para alcanzar el objetivo de esta sección, a continuación, en la sub-sección X.2.1 se plantean los aspectos vinculados con el diseño del AFE y en la sub-sección X.2.2 se estiman diversas soluciones del AFE para analizar el número de factores a retener, su ajuste global y la interpretación de los factores obtenidos. Paralelamente se presentan las medidas estadísticas necesarias asociadas a cada una de las etapas y los resultados empíricos obtenidos. Al final de la sección se realiza una breve discusión sobre los resultados esperados y los efectivamente obtenidos.

### X.2.1 Diseño

Para la implementación del AFE se seleccionaron únicamente las variables observadas que se espera den cuenta de las cinco dimensiones conceptuales/teóricas de este trabajo (*propensión modo 1, propensión modo 2, satisfacción con el proceso de evaluación, flexibilidad epistemológica del sistema de evaluación y dirección de la reforma del sistema de evaluación*). Se trata de 33 variables observadas, cuya representación colectiva se espera que se exprese en dichos conceptos. El número de variables seleccionadas si bien no es bajo, *a priori* se considera suficiente para mantener al menos

5 o 6 variables por cada dimensión (factor) latente. Con esto se espera reforzar la fortaleza del AFE en la búsqueda de estructuras entre grupos de variables.

Desde el punto de vista de las propiedades de medición de las 33 variables seleccionadas, 10 variables son ordinales, con escala Likert de 1 a 5, siendo 1 “Completamente en desacuerdo” y 5 “Muy de acuerdo”; y las restantes 23 son dicotómicas. Asimismo, algunas variables observadas fueron transformadas, pasando de ser originalmente nominales u ordinales a ser dicotómicas. La decisión para realizar dichas transformaciones se basó, por un lado, en los requerimientos de medición del AFE –las variables nominales deben transformarse para rescatar las características específicas en variables individuales- y por otro, en las distribuciones de los datos –las variables ordinales originales no se aproximaban a una parábola sino que se aproximaban más a una forma de U-. Adicionalmente, y teniendo presente los conceptos teóricos que se busca medir, a algunas variables observadas se les revirtió la escala para mantener la consistencia entre la medición y las definiciones conceptuales. Finalmente, las categorías de respuesta con la opción “No sabe” o “No corresponde” se trataron como Datos Faltantes, porque no aportaban información sustantiva a los objetivos de este trabajo. En la Tabla 21 se presentan las variables observadas (ítems) seleccionadas, sus propiedades de medición, y se detallan todas las transformaciones realizadas.

Tabla 21: Ítems, tamaño de muestra, datos faltantes y propiedades de medida

Ítems (Vars. Obs.)	Descripción	N	% Datos Faltantes	Escala
p_18_c_rec <sup>1</sup>	La investigación se orienta por su campo disciplinario	7825	2.86	1 – 5 Likert. Donde 1 = Completamente en desacuerdo y 5 = Muy de Acuerdo
p_18_f_rec <sup>1</sup>	El conocimiento se organiza disciplinariamente	7797	3.20	
p_18_g_rec <sup>1</sup>	Los actores involucrados en la generación de conocimiento son únicamente académicos	7831	2.78	
p_18_i_rec <sup>1</sup>	La principal fuente de reconocimiento es la publicación de artículos científicos arbitrados	7839	2.68	
p_18_j_rec <sup>1</sup>	La principal fuente de reconocimiento es el libro publicado	7809	3.05	
p_18_b_rec <sup>1</sup>	El objetivo de la generación de conocimiento es su aplicación	7848	2.57	
p_18_d_rec <sup>1</sup>	La generación de conocimiento se orienta por la búsqueda de soluciones a problemas complejos (productivos, sociales, ambientales, de política pública)	7827	2.83	
p_18_e_rec <sup>1</sup>	El conocimiento se organiza multi/inter/transdisciplinariamente	7832	2.77	
p_18_h_rec <sup>1</sup>	La generación de conocimiento depende de la interacción entre académicos y las contrapartes interesadas en la solución de problemas	7823	2.88	
p_18_k_rec <sup>1</sup>	La principal fuente de difusión del conocimiento es la solución encontrada (productos, procesos, marcas, patentes)	7827	2.83	
p_1_a_rec <sup>2</sup>	La información que solicita el SNI es la necesaria para lograr un juicio académico apropiado (ingreso o reintegro no vigente)	6894	14.41	0 – 1 Donde 0 = No y 1 = Sí
p_1_b_rec <sup>2</sup>	La información que solicita el SNI es la necesaria para lograr un juicio académico apropiado (reingreso vigente)	7030	12.73	
p_2_a_rec <sup>3</sup>	La información que solicita el SNI es excesiva	7854	2.50	0 – 1

p_2_c_rec <sup>3</sup>	La información que solicita el SNI deja afuera aportaciones importantes	7854	2.50	Donde 0 = Sí y 1 = No
p_5_r_d <sup>4</sup>	Los mecanismos de designación de los evaluadores permiten seleccionar perfiles que garanticen una evaluación académica apropiada en su área de conocimiento	7014	12.92	0 – 1 Donde 0 = No y 1 = Sí
p_7_rec <sup>2</sup>	Los requisitos para ser evaluador son pertinentes	5516	31.52	0 – 1 Donde 0 = No y 1 = Sí
p_8_rec <sup>2</sup>	La manera en que se integran las comisiones dictaminadoras es adecuada	5057	37.22	
p_9_a_rec <sup>3</sup>	La manera en que se integran las comisiones dictaminadoras no considera suficientemente el equilibrio regional	5054	37.26	0 – 1 Donde 0 = Sí y 1 = No
p_9_b_rec <sup>3</sup>	La manera en que se integran las comisiones dictaminadoras no considera suficientemente el equilibrio institucional	5054	37.26	
p_9_d_rec <sup>3</sup>	La manera en que se integran las comisiones dictaminadoras no considera otros requisitos (además de estar en el Nivel III) para dar más garantías de una evaluación objetiva.	5054	37.26	
p_3_r_d <sup>4</sup>	Los períodos de evaluación del SNI para el reingreso vigente permiten lograr resultados suficientes y relevantes en la investigación que Ud. realiza	7489	7.03	0 – 1 Donde 0 = No y 1 = Sí
p_10_rec <sup>2</sup>	Los criterios de evaluación para cada categoría reconocen los distintos perfiles de desempeño de los investigadores	6832	15.18	0 – 1 Donde 0 = No y 1 = Sí
p_11_r_d <sup>4</sup>	Los criterios de evaluación le permiten asumir estrategias de investigación innovadoras	7079	12.12	0 – 1 Donde 0 = No y 1 = Sí
p_9_c_rec <sup>3</sup>	La integración de las comisiones dictaminadoras no considera suficientemente la diversidad disciplinaria	5054	37.26	0 – 1 Donde 0 = Sí y 1 = No
p_17_a_rec <sup>3</sup>	El SNI le generó dificultades para involucrarse en investigaciones de larga duración	7400	8.13	
p_17_b_rec <sup>3</sup>	El SNI le generó dificultades para trabajar en equipo por la presión para publicar individualmente	7400	8.13	
p_17_d_rec <sup>3</sup>	El SNI generó estandarización de trayectorias	7400	8.13	
p_20_g_rec <sup>2</sup>	Dar mayor reconocimiento a la investigación colectiva	6154	23.60	0 – 1 Donde 0 = No y 1 = Sí
p_20_k_rec <sup>2</sup>	Considerar adecuadamente la investigación multi/inter/transdisciplinaria	6209	22.92	
p_20_l_rec <sup>2</sup>	Dar más peso a los resultados de la investigación aplicada	5901	26.74	
p_20_i_rec <sup>2</sup>	Reformular los criterios de evaluación	6298	21.81	
p_20_j_rec <sup>2</sup>	Modificar la composición disciplinaria de las comisiones dictaminadoras por área	5894	26.83	
p_20_e_rec <sup>2</sup>	Dar más peso a la evaluación cualitativa	6223	22.74	
<b>N con casos completos (listwise)</b>		<b>2283</b>		
<b>N con pares de casos completos (pairwise)</b>		<b>8053</b>		
<b>N en la base de datos</b>		<b>8055</b>		

Notas: (1) A estos ítems se les revirtió la escala. (2) A estos ítems se les recodificó la escala. Las respuestas en la categoría “No Sabe” se tratan como Dato Faltante. (3) A estos ítems se les transformó la variable - pasando de nominal a dicotómica-, y se les revirtió la escala. Las respuestas en la categoría “No Sabe” se tratan como Dato Faltante. (5) A estos ítems se les transformó la

variable - pasando de ordinal a dicotómica-, y se les recodificó la escala. Las respuestas en las categorías “No Sabe” y/o “No corresponde” se tratan como Dato Faltante.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

El tamaño de muestra de este trabajo es lo suficientemente grande como para asegurar la obtención de resultados confiables y robustos, ya que se cuenta con más de 240 observaciones por variable incluida en el AFE. Incluso en el escenario más exigente, si se considera sólo las observaciones con casos completos (N=2283), se alcanza un tamaño de muestra de 69 observaciones por cada variable incluida en el AFE, lo que supera largamente las reglas generales planteadas en la literatura discutidas en el capítulo VIII.<sup>69</sup> (Ver Tabla 21).

El análisis factorial utiliza la matriz de correlación para encontrar la estructura de relaciones latente entre las variables observadas. En este caso y dado que se trata de variables de escala ordinal y dicotómica (ver Tabla 21), la medida de correlación más apropiada es el cálculo de las correlaciones policóricas entre las variables observadas (variables de medida). En la Tabla 22 se presenta la matriz de correlaciones policóricas calculadas.

---

<sup>69</sup> Una forma adicional para evaluar la robustez de los resultados, es realizar un análisis específico de la potencia estadística y del tamaño de muestra mínimo requerido para este trabajo.

Tabla 22: Matriz de correlación (policórica) muestral

Ítems	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
p_18_c_rec	1.0																							
p_18_f_rec	0.5	1.0																						
p_18_g_rec	0.3	0.4	1.0																					
p_18_i_rec	0.2	0.3	0.5	1.0																				
p_18_j_rec	0.2	0.2	0.4	0.5	1.0																			
p_18_b_rec	0.4	0.2	0.1	0.1	0.1	1.0																		
p_18_d_rec	0.3	0.2	0.1	0.1	0.2	0.7	1.0																	
p_18_e_rec	0.1	0.1	0.0	0.0	0.1	0.5	0.6	1.0																
p_18_h_rec	0.2	0.2	0.1	0.1	0.1	0.5	0.6	0.5	1.0															
p_18_k_rec	0.2	0.2	0.2	0.2	0.3	0.5	0.5	0.3	0.5	1.0														
p_1_a_rec	0.1	0.2	0.2	0.4	0.2	0.1	0.1	0.0	0.1	0.1	1.0													
p_1_b_rec	0.1	0.2	0.2	0.3	0.1	0.1	0.1	-0.1	0.0	0.1	0.9	1.0												
p_2_a_rec	0.1	0.0	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.5	0.7	1.0											
p_2_c_rec	0.1	0.2	0.2	0.4	0.1	0.1	0.0	-0.1	0.0	0.1	0.9	1.0	-0.9	1.0										
p_5_r_d	0.1	0.2	0.1	0.2	0.2	0.2	0.1	0.1	0.1	0.2	0.6	0.5	0.2	0.5	1.0									
p_7_rec	0.1	0.2	0.1	0.2	0.2	0.1	0.1	0.0	0.1	0.1	0.6	0.6	0.3	0.6	0.8	1.0								
p_8_rec	0.2	0.3	0.2	0.2	0.1	0.2	0.1	0.0	0.1	0.2	0.6	0.5	0.3	0.5	0.9	0.9	1.0							
p_9_a_rec	0.1	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.1	0.2	0.1	0.4	0.3	1.0	1.0						
p_9_b_rec	0.2	0.2	0.1	0.1	0.1	0.1	-0.1	-0.1	0.0	0.0	0.2	0.2	0.1	0.1	0.5	0.4	1.0	-1.0	1.0					
p_9_d_rec	0.1	0.2	0.2	0.2	0.2	0.1	0.1	0.0	0.0	0.1	0.4	0.4	0.3	0.4	0.7	0.7	1.0	-0.9	-0.9	1.0				
p_3_r_d	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.0	0.0	0.1	0.5	0.5	0.3	0.5	0.4	0.4	0.4	0.2	0.2	0.3	1.0			
p_10_rec	0.2	0.2	0.2	0.3	0.2	0.2	0.1	0.0	0.1	0.2	0.6	0.7	0.4	0.7	0.7	0.7	0.7	0.2	0.3	0.6	0.5	1.0		
p_11_r_d	0.2	0.3	0.2	0.3	0.2	0.3	0.2	0.1	0.1	0.2	0.6	0.6	0.2	0.6	0.7	0.6	0.7	0.2	0.3	0.6	0.5	0.9	1.0	
p_9_c_rec	0.0	0.1	0.1	0.2	0.0	0.1	0.1	0.1	0.2	0.2	0.3	0.3	0.1	0.3	0.7	0.5	1.0	-1.0	-0.9	-1.0	0.3	0.5	0.4	1.0
p_17_a_rec	0.1	0.2	0.1	0.3	0.2	0.2	0.2	0.1	0.2	0.3	0.4	0.5	0.3	0.3	0.5	0.5	0.6	0.0	0.3	0.3	0.4	0.6	0.6	0.3
p_17_b_rec	0.1	0.2	0.2	0.3	0.3	0.1	0.1	0.1	0.1	0.2	0.4	0.5	0.2	0.4	0.6	0.4	0.5	0.1	0.2	0.3	0.4	0.6	0.6	0.3
p_17_d_rec	0.1	0.2	0.1	0.2	0.1	0.3	0.2	0.2	0.2	0.2	0.3	0.5	0.2	0.4	0.6	0.4	0.7	0.2	0.2	0.5	0.4	0.5	0.5	0.1
p_20_g_rec	-0.1	-0.1	-0.2	-0.3	-0.1	0.1	0.1	0.2	0.2	0.0	-0.4	-0.4	-0.2	-0.4	-0.4	-0.3	-0.4	-0.2	-0.2	-0.3	-0.4	-0.5	-0.4	-0.2
p_20_k_rec	-0.1	-0.2	-0.2	-0.3	-0.2	0.0	0.1	0.2	0.0	-0.1	-0.5	-0.6	-0.2	-0.6	-0.6	-0.5	-0.6	-0.3	-0.3	-0.5	-0.5	-0.6	-0.6	-0.4
p_20_l_rec	-0.1	-0.1	-0.2	-0.2	-0.2	0.3	0.3	0.1	0.2	0.3	-0.2	-0.3	-0.1	-0.3	-0.3	-0.3	-0.3	0.1	-0.2	-0.2	-0.3	-0.4	-0.4	-0.2
p_20_i_rec	-0.1	-0.2	-0.2	-0.3	-0.2	-0.1	0.0	0.0	0.0	-0.1	-0.6	-0.7	-0.4	-0.7	-0.6	-0.7	-0.7	-0.3	-0.4	-0.6	-0.5	-0.8	-0.7	-0.4
p_20_j_rec	-0.1	-0.2	-0.1	-0.2	-0.1	-0.1	0.0	0.0	0.0	-0.1	-0.5	-0.5	-0.3	-0.5	-0.7	-0.6	-0.7	-0.4	-0.3	-0.5	-0.5	-0.7	-0.6	-0.6
p_20_e_rec	-0.1	-0.2	-0.1	-0.3	-0.1	-0.1	0.0	0.0	0.0	-0.1	-0.4	-0.5	-0.2	-0.5	-0.5	-0.4	-0.5	-0.2	-0.2	-0.3	-0.4	-0.5	-0.5	-0.4

Notas: (1) El valor calculado del Determinante de la Matriz de correlación es 5.85047190974944E-10. (2) Los rectángulos sombreados resalta la presencia de un factor en cada factor resultante del AFE.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Antes de implementar el AFE y a efectos de determinar la relevancia dicho análisis<sup>70</sup>, es importante inspeccionar visualmente la matriz de correlaciones policóricas calculadas y observar el valor de su determinante.

Lo primero que importa señalar es que el valor calculado del Determinante de la matriz de correlación es muy cercano a cero ( $5.85E-10$ ). De acuerdo con la evidencia se concluye que existe una estructura de correlación fuerte entre las variables y por lo tanto la implementación del AFE (o cualquier otro análisis factorial) resulta pertinente.

Adicionalmente, tal como puede observarse en la Tabla 22, y de acuerdo con los objetivos de este trabajo y el marco teórico conceptual, los rectángulos sombreados resaltan la agrupación de variables que se espera encontrar en cada factor resultante de la solución del AFE. Es decir, se espera que la asociación de los ítems: i) 1 a 5 se explique por la dimensión conceptual *propensión modo 1*; ii) 6 a 10 se explique por la dimensión *propensión modo 1*; iii) 11 a 20 se explique por la dimensión *satisfacción con el proceso de evaluación*; iv) 21 a 27 se explique por la dimensión *flexibilidad epistemológica del sistema de evaluación*; y v) 28 a 33 se explique por la dimensión *dirección de la reforma del sistema de evaluación*. Si se esperan esos resultados del AFE (coincidencia perfecta entre la teoría y lo empírico), el patrón o estructura que se debería observar en la matriz de correlación (tabla 22) tiene que ser tal que los ítems que componen cada uno de esos factores correlacionen fuertemente entre sí y poco con los ítems que son explicados por los restantes factores. A continuación se realiza este análisis.

De hecho, los ítems 28 a 33 presentan una estructura de correlación relativamente muy fuerte entre sí (correlaciones entre 0.5 y 0.9) y relativamente menos fuerte con los demás ítems (ver tabla 22). Esto sugiere que probablemente la solución del AFE devuelva un factor que explique la asociación entre esas variables y que se denominaría como *dirección de la reforma del sistema de evaluación*.

De la misma manera aunque con menor fortaleza, los ítems 6 a 10 tienen una correlación relativamente fuerte entre sí (la mayoría de las correlaciones están entre 0.5 y 0.7) y relativamente más débil con los demás ítems (correlaciones entre |0.1| y |0.3|) (ver tabla 22), nuevamente aquí la evidencia sugiere que probablemente la solución del AFE devuelva un factor que explique la asociación entre esas variables y que se denominaría como *propensión modo 2*.

Algo similar ocurre con los ítems 1 a 5 aunque con una estructura de correlación menos nítida que en los casos anteriores. Si bien la mayoría de las correlaciones entre estos ítems son considerables (entre 0.3 y 0.5), algunos de éstos también presentan correlaciones considerables (entre |0.3| y |0.4|) con algunos de los restantes ítems. Por lo que en este caso no es tan claro que la solución probable del AFE devuelva un factor que explique la asociación entre esas variables y que se denominaría como *propensión modo 1*.

Los ítems 11 a 27 son los que presentan una estructura de correlación bastante menos clara en comparación con los casos anteriores. Si bien los ítems 11 a 20 exhiben –para la mayoría de los pares- correlaciones considerables y fuertes (entre |0.3| y |0.9|) entre sí, también muestran

---

<sup>70</sup> Además de los enfoques que se aplican en esta sección para determinar la relevancia de implementar el AFE, existen otros como por ejemplo: i) el análisis de las correlaciones parciales entre las variables; ii) la prueba de esfericidad de Bartlett; y iii) las medidas de adecuación muestral como el índice KMO (Kaiser-Meyer-Okin) (ver capítulo VIII).



correlaciones relativamente altas con una buena parte de los restantes ítems, en particular con los ítems 21 a 27. Simultáneamente, los ítems 21 a 27 si bien presentan correlaciones considerables y fuertes (entre 0.3 y 0.9) entre sí, también muestran correlaciones relativamente altas con una buena parte de los restantes ítems, en particular con los ítems 11 a 20. (Ver tabla 22, la figura triangular con bordes negros anchos resalta estos resultados). En este sentido, no es claro que el AFE de como solución los factores *satisfacción con el proceso de evaluación* para explicar la asociación entre los ítems 11 a 20, ni *flexibilidad epistemológica del sistema de evaluación* para explicar la asociación entre los ítems 21 a 27. Probablemente, y dada las correlaciones entre todos estos ítems, el AFE encuentre alguna otra combinación de estas variables para explicar dichas asociaciones.

Por lo argumentado hasta aquí, la implementación del AFE tiene sentido y eventualmente podría arrojar resultados empíricamente relevantes desde el punto de vista conceptual y teórico, sin embargo, también es razonable esperar que los resultados del AFE coincidirán parcialmente con el modelo teórico de medida que sustenta este trabajo. Finalmente y antes de implementar el AFE, cabe señalar que los signos de todas las correlaciones calculadas entre los ítems son consistentes con lo que se espera de acuerdo con la teoría.

A continuación se presentan los resultados obtenidos de la implementación del AFE para medir las relaciones entre la evaluación del desempeño de la investigación individual y los modos diferenciales de producción de conocimiento. La salida completa derivada de la implementación del AFE está disponible en el Anexo 1.

#### X.2.2 Estimación: selección del número de factores a retener, ajuste e interpretación

Como se planteó en el capítulo VIII, dada la escala de medición de las variables seleccionadas para el análisis, la estimación de los parámetros del AFE se realiza con mínimos cuadrados ponderados con media y varianza ajustada (WLSMV), ya que son robustos a las violaciones a la normalidad de los datos. Adicionalmente, la decisión sobre la cantidad de factores a retener depende, por un lado del número de dimensiones latentes que se espera encontrar de acuerdo con la teoría –en el contexto de este trabajo se espera encontrar una estructura multifactorial de cinco dimensiones o factores latentes-, y por otro, de la cantidad de factores que es razonable esperar de acuerdo con la evidencia empírica –según el análisis de la matriz de correlaciones realizado es razonable esperar entre 5 y 7 factores-. Además, y como otros criterios empíricos -complementarios entre sí- para decidir el número de factores a retener del AFE, a continuación se presenta: i) la Tabla 23 que contiene los valores propios derivados de la matriz de correlación muestral y el porcentaje de la varianza explicada por cada factor; ii) el *gráfico de sedimentación* (ver Gráfico 18); y iii) en la tabla 24, la comparación de las medidas de bondad de ajuste global de los modelos AFE según el número de factores retenidos.

*Tabla 23: Valores propios y porcentaje de la varianza explicada por cada factor*

Factor	Valores Propios		
	Total	% de la Varianza	% Acumulado
1	10.75	32.6%	32.6%
2	3.744	11.3%	43.9%
3	2.385	7.2%	51.1%
4	1.877	5.7%	56.8%
5	<b>1.807</b>	<b>5.5%</b>	<b>62.3%</b>

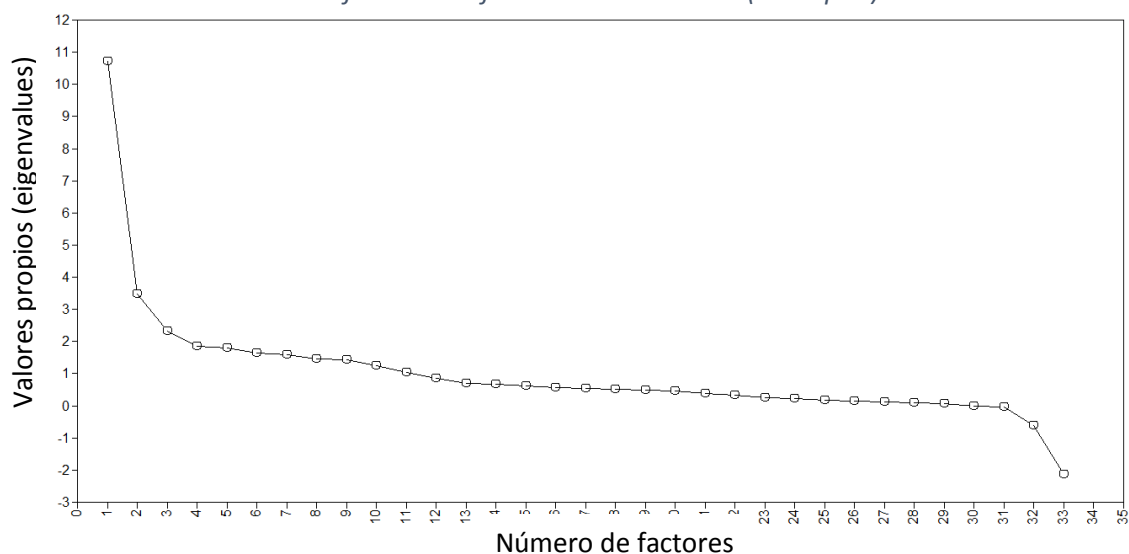
6	1.663	5.0%	67.4%
7	1.602	4.9%	72.2%
8	1.467	4.4%	76.7%
9	1.426	4.3%	81.0%
10	1.243	3.8%	84.7%
11	1.035	3.1%	87.9%
12	0.829	2.5%	90.4%
13	0.7	2.1%	92.5%
14	0.668	2.0%	94.5%
15	0.596	1.8%	96.3%
16	0.525	1.6%	97.9%
17	0.498	1.5%	99.4%
18	0.463	1.4%	100.8%
19	0.425	1.3%	102.1%
20	0.389	1.2%	103.3%
21	0.324	1.0%	104.3%
22	0.306	0.9%	105.2%
23	0.229	0.7%	105.9%
24	0.222	0.7%	106.6%
25	0.166	0.5%	107.1%
26	0.139	0.4%	107.5%
27	0.131	0.4%	107.9%
28	0.101	0.3%	108.2%
29	0.073	0.2%	108.4%
30	-0.01	0.0%	108.4%
31	-0.034	-0.1%	108.3%
32	-0.619	-1.9%	106.4%
33	-2.122	-6.4%	100.0%

Notas: Estimaciones realizadas con M-PLUS Versión 7.4.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

De acuerdo con los valores propios, el porcentaje de la varianza explicada por cada factor y las recomendaciones de la literatura, una estructura factorial con cinco factores da lugar a una varianza explicada acumulada de 62.3% (ver tabla 23), valor que resulta satisfactorio tanto desde el punto de vista conceptual como empírico.

Gráfico 18: Gráfico de sedimentación (scree plot)



De acuerdo con el gráfico de sedimentación (gráfico 18) el número óptimo de factores a retener es 3, ya que coincide con el cambio de la pendiente del gráfico de los valores propios de cada factor.

A continuación se presenta la Tabla 24 que resume algunas medidas de bondad de ajuste global derivadas de las soluciones del AFE encontradas de ajustar modelos que van de 1 a 7 factores.

Tabla 24: Medidas de Ajuste Global del AFE

Medidas <sup>71</sup>	$\chi^2$	gl	RMSEA	CFI	TLI
1 Factor	48802.109*	495	0.110	0.790	0.776
2 Factores	22445.137*	463	0.077	0.905	0.891
3 Factores	12699.639*	432	0.059	0.947	0.935
4 Factores	7009.516*	402	0.045	0.971	0.962
5 Factores	4087.423*	373	0.035	0.984	0.977
6 Factores	2250.612*	345	0.026	0.992	0.987
7 Factores	1429.704*	318	0.021	0.995	0.992

Notas: Estimaciones realizadas con M-PLUS Versión 7.4. /  $\chi^2$ = Estadístico Ji-Cuadrado de bondad de ajuste; gl=grados de libertad; RMSEA= Raíz Cuadrada del Error de Aproximación Promedio; CFI= Índice de Ajuste Comparado; TLI= Índice Tucker Lewis. /\* Indican que  $\chi^2$  son estadísticamente significativas ( $p < .001$ ).

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

La primera medida que se presenta en la segunda columna de la tabla 24 corresponde a los valores calculados del estadístico Ji-Cuadrado ( $\chi^2$ ) para cada uno de los modelos ajustados de 1 a 7 factores, y en la tercera columna pueden observarse sus respectivos grados de libertad. En todos los casos los valores calculados del estadístico  $\chi^2$  son estadísticamente significativos. Es decir, existen diferencias significativas entre las matrices de varianzas y covarianzas observada y estimada.

<sup>71</sup> En el capítulo VIII se definen todas las medidas de ajuste global y sus respectivos puntos de referencia de corte.

La evidencia empírica sugiere que el ajuste de los modelos es pobre, sin embargo no hay que olvidar la alta sensibilidad de este estadístico a los grados de libertad y al tamaño de la muestra. Dado el tamaño de la muestra (N=8053) que se utiliza en este trabajo, es esperable obtener estos resultados aunque la diferencia entre las matrices de varianza y covarianza observadas y estimadas no sean estadísticamente significativas. Es así que a efectos de no rechazar un modelo que en realidad presenta un buen ajuste, se calcularon otras medidas complementarias de bondad de ajuste, tanto absolutas como relativas. En la cuarta columna de la tabla 24 se presenta el RMSEA. De acuerdo con los valores del RMSEA calculados (ver tabla 24), los modelos estimados con 4 o más factores son los que exhiben una buena y mejor bondad de ajuste. Adicionalmente, en la quinta y sexta columna de la tabla 24, se presentan dos medidas de bondad de ajuste relativo o incremental, el CFI y el TLI.

En base a este análisis de la bondad de ajuste, se puede afirmar que, de acuerdo con los valores calculados del RMSEA, CFI y TLI, los modelos estimados con 4 factores o más son los que presentan buen y mejor ajuste (ver tabla 24).

A efectos de determinar la estructura factorial subyacente final (cantidad de factores a retener y su sentido conceptual) y observar su grado de coincidencia con lo que se espera de acuerdo con la teoría que sustenta este trabajo, a continuación en las Tablas 25, 26 y 27 se presentan las soluciones rotadas del AFE para los modelos estimados con 4, 5 y 6 factores respectivamente. La selección de la solución factorial final y la interpretación de su estructura supone la realización de tres procesos complementarios: i) análisis de las cargas factoriales; ii) análisis de la solución factorial rotada; e iii) interpretación conceptual de los factores.

#### X.2.2.1 La solución de 4 factores

Tabla 25: Estructura Factorial: Solución del AFE para el modelo con 4 factores

Nro.	Ítems	Factores (solución con rotación oblicua - GEOMIN)				Varianzas Residuales
		1	2	3	4	
1	p_18_c_rec	0.398*	0.435*	0.016	0.063*	0.675
2	p_18_f_rec	0.274*	0.466*	0.053*	0.115*	0.687
3	p_18_g_rec	0.122*	0.559*	0.113*	0.059*	0.646
4	p_18_i_rec	0.067	0.595*	0.266*	0.111*	0.504
5	p_18_j_rec	0.234*	0.576*	0.166*	-0.001	0.590
6	p_18_b_rec	0.785*	0.093*	-0.047	0.035	0.410
7	p_18_d_rec	0.848*	0.066	-0.051	-0.01	0.318
8	p_18_e_rec	0.617*	-0.015	-0.017	-0.066*	0.624
9	p_18_h_rec	0.665*	0.054	0.004	-0.066*	0.566
10	p_18_k_rec	0.599*	0.259*	0.097*	-0.026*	0.571
11	p_1_a_rec	-0.291*	0.036*	0.950*	0.013	0.180
12	p_1_b_rec	-0.341*	0.064*	1.023*	0.002	0.053
13	p_2_a_rec	-0.112*	-0.081*	0.608*	-0.063	0.687
14	p_2_c_rec	-0.382*	0.085*	1.048*	-0.028*	0.023
15	p_5_r_d	0.088*	-0.370*	0.420*	0.559*	0.188
16	p_7_rec	-0.007	-0.283*	0.554*	0.395*	0.310
17	p_8_rec	0.074*	-0.391*	0.473*	0.593*	0.061
18	p_9_a_rec	0.007	-0.135*	0.162*	0.273*	0.853
19	p_9_b_rec	-0.009	-0.169*	0.143*	0.252*	0.871

20	p_9_d_rec	0.050*	-0.241*	0.348*	0.412*	0.539
21	p_3_r_d	-0.031	0.01	0.245*	0.349*	0.749
22	p_10_rec	-0.002	-0.039*	0.583*	0.424*	0.271
23	p_11_r_d	0.131*	0.036	0.490*	0.399*	0.359
24	p_9_c_rec	0.042	-0.208*	0.229*	0.460*	0.618
25	p_17_a_rec	0.169*	-0.053	0.316*	0.332*	0.626
26	p_17_b_rec	0.067	-0.006	0.327*	0.324*	0.674
27	p_17_d_rec	0.069	-0.112*	0.312*	0.282*	0.721
28	p_20_g_rec	0.220*	-0.150*	0.138*	-0.859*	0.242
29	p_20_k_rec	0.177*	-0.145*	0.024*	-0.968*	0.010
30	p_20_l_rec	0.415*	-0.123*	0.068*	-0.658*	0.406
31	p_20_i_rec	0.114*	-0.055	-0.155*	-0.897*	0.047
32	p_20_j_rec	0.011	-0.078*	0.026	-0.826*	0.312
33	p_20_e_rec	0.064*	0.114*	-0.064	-0.894*	0.175

Notas: Estimaciones realizadas con M-PLUS Versión 7.4. /\* Indican que las cargas factoriales son estadísticamente significativas al nivel  $\alpha = 5\%$ .

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Tal como puede observarse en la Tabla 25 los ítems con mayores cargas factoriales son los que aparecen sombreados, además todos resultan significativos con una confianza del 95%. La varianza explicada acumulada para esta solución es de 56.8%, donde el primer factor explica 32.6%, el segundo 11.3%, el tercero 7.2% y el cuarto el 5.7% de la varianza total de los ítems (ver tabla 23). Al estudiar las cargas factoriales e interpretar sus resultados por factor, es posible observar que para los dos primeros factores emergen las conceptualizaciones *propensiones modo 2* y *modo 1* de producción de conocimiento respectivamente. En particular, las cargas factoriales más altas del factor 1 se concentran entre los ítems de 6 a 10, con cargas que van desde 0.599 hasta 0.848; mientras que las cargas factoriales más altas del factor 2 se concentran entre los ítems 1 a 5, con cargas relativamente menores y que van desde 0.435 hasta 0.595 (ver tabla 25). A la vez, las cargas factoriales más altas de los ítems de 1 a 5 se concentran en el factor 2, mientras que las cargas más altas de los ítems 6 a 10 se concentran en el factor 1. Estos resultados muestran que existe una asociación entre esos ítems y sus respectivos factores. Desde el punto de vista de la interpretación de esos factores, la evidencia sugiere que el factor 1 se vincula con la conceptualización dada de propensión modo 2 de producción de conocimiento, mientras que el factor 2 con la de propensión modo 1. Estos dos primeros factores miden lo que en esta tesis se conceptualizó como propensiones modo 2 y modo 1 de producción de conocimiento, y son esos conceptos –latentes– los que explican parte de la varianza de los ítems 6 a 10 y 1 a 5 respectivamente. Del estudio de las cargas factoriales de los restantes ítems (11 al 33), no emerge –con la claridad suficiente– una interpretación conceptualmente relevante de los factores. En este sentido y teniendo en cuenta los restantes estadísticos calculados, la solución del AFE con una estructura multifactorial de cuatro factores no resulta satisfactoria ni desde el punto de vista empírico ni conceptual. A continuación se presentan y discuten los resultados del AFE con una estructura multifactorial de 5 factores tal como se espera de acuerdo con la teoría.

### X.2.2.2 La solución de 5 factores

En la Tabla 26 puede observarse la solución del AFE para el modelo multifactorial de cinco factores. Los ítems con mayores cargas factoriales son los que aparecen sombreados, además todas esas cargas factoriales resultan significativas con una confianza del 95%. La varianza explicada acumulada para esta solución es de 62.3%, donde los primeros cuatro factores explican lo mismo que en el caso anterior, y el quinto -factor adicional de esta solución- contribuye con un 5.5% de varianza explicada (ver tabla 23).

Tabla 26: Estructura Factorial: Solución del AFE para el modelo con 5 factores

Nro.	Ítems	Factores (solución con rotación oblicua - GEOMIN)					Varianzas Residuales
		1	2	3	4	5	
1	p_18_c_rec	0.493*	0.212*	0.150*	-0.154*	0.015	0.644
2	p_18_f_rec	0.565*	0.053*	0.222*	-0.161*	0.01	0.618
3	p_18_g_rec	0.629*	-0.075*	0.038*	-0.004	-0.012	0.609
4	p_18_i_rec	0.639*	-0.067*	-0.050*	0.220*	-0.095*	0.476
5	p_18_j_rec	0.625*	0.067*	-0.048*	0.112*	0.022	0.567
6	p_18_b_rec	0.061*	0.760*	0.022	-0.057*	-0.051*	0.396
7	p_18_d_rec	0.013	0.849*	-0.029*	-0.031*	-0.040*	0.286
8	p_18_e_rec	-0.047*	0.630*	0.001	-0.001	0.050*	0.609
9	p_18_h_rec	0.029*	0.657*	0.007	0.001	0.052*	0.554
10	p_18_k_rec	0.258*	0.546*	-0.013	0.082*	0.026	0.561
11	p_1_a_rec	0.009	-0.026*	0.121*	0.861*	0.024	0.166
12	p_1_b_rec	-0.003	-0.018	0.046*	0.940*	-0.027*	0.055
13	p_2_a_rec	-0.141*	0.122*	0.004	0.611*	0.031	0.63
14	p_2_c_rec	0.003	-0.042*	-0.017	0.991*	-0.031	0.011
15	p_5_r_d	-0.032*	-0.006	0.920*	-0.005	-0.021	0.15
16	p_7_rec	0.001	-0.055*	0.788*	0.213*	0.089*	0.265
17	p_8_rec	-0.027*	-0.026*	0.968*	0.021	-0.038*	0.021
18	p_9_a_rec	-0.025	-0.026	0.434*	-0.029	-0.042	0.81
19	p_9_b_rec	-0.053	-0.05	0.458*	-0.055	-0.008	0.816
20	p_9_d_rec	0.004	-0.036	0.796*	-0.016	0.038	0.413
21	p_3_r_d	-0.007	0.063*	0.109*	0.227*	-0.297*	0.732
22	p_10_rec	0.068*	0.051*	0.503*	0.377*	-0.139*	0.251
23	p_11_r_d	0.146*	0.137*	0.475*	0.277*	-0.130*	0.345
24	p_9_c_rec	-0.049*	0.002	0.620*	-0.034	-0.126*	0.554
25	p_17_a_rec	0.071*	0.138*	0.483*	0.126*	-0.032	0.604
26	p_17_b_rec	0.144*	0.006	0.487*	0.124*	-0.009	0.635
27	p_17_d_rec	0.055	0.016	0.530*	0.091	0.054	0.677
28	p_20_g_rec	-0.048*	0.134*	0.112*	0.01	0.914*	0.213
29	p_20_k_rec	-0.075*	0.093*	-0.050*	-0.032*	0.931*	0.011
30	p_20_l_rec	-0.100*	0.399*	0.004	0.037	0.666*	0.383
31	p_20_i_rec	0.047*	-0.055*	-0.107*	-0.209*	0.831*	0.02
32	p_20_j_rec	0.036*	-0.125*	0.016	-0.072*	0.832*	0.27
33	p_20_e_rec	0.100*	-0.006	-0.372*	0.032*	0.716*	0.15

Notas: Estimaciones realizadas con M-PLUS Versión 7.4. /\* Indican que las cargas factoriales rotadas son estadísticamente significativas al nivel  $\alpha = 5\%$ .

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Nuevamente aquí, al estudiar las cargas factoriales e interpretar sus resultados por factor, es posible observar que para los dos primeros factores emergen las conceptualizaciones propensiones modo 1 y modo 2 de producción de conocimiento respectivamente. Las cargas factoriales más altas de los ítems de 1 a 5 se concentran en el factor 1, mientras que las cargas más altas de los ítems 6 a 10 se concentran en el factor 2. A diferencia de lo que sucedía en la solución del AFE de cuatro factores, aquí las cargas factoriales más altas del factor 1 se concentran entre los ítems de 1 a 5, con cargas más altas también, y que van desde 0.493 hasta 0.639; mientras que las cargas factoriales más altas del factor 2 se concentran entre los ítems 6 a 10, con cargas que van desde 0.546 hasta 0.849 (ver tabla 26). Estos resultados muestran que las conceptualizaciones propensión modo 1 y modo 2 de producción de conocimiento quedan mejor representadas en una estructura factorial de 5 factores.

A la vez, las cargas factoriales más altas de los ítems 15 a 20 y 22 a 27 se concentran en el factor tres, y van desde 0.434 hasta 0.968 (ver tabla 26). La mayoría de esos ítems presentan una asociación relativamente fuerte con el factor 3. Los ítems 18 y 19 son candidatos a ser eliminados del análisis debido a la magnitud de sus varianzas residuales (0.81 y 0.816 respectivamente). Es decir, más del 80% de la varianza de esos ítems no es explicada por la estructura multifactorial de cinco factores que se está analizando. La decisión final sobre la eliminación o no de esos ítems se tomará cuando se realice el AFC.

La interpretación conceptual del factor 3 no es evidente. Por un lado, podría interpretarse como el grado de acuerdo que tienen los investigadores con el sistema de evaluación, ya que los ítems que agrupa y que cuentan con mayores cargas factoriales son los vinculados a la designación de evaluadores, integración de las comisiones dictaminadoras y criterios de evaluación. También agrupa -pero con menor peso relativo- los ítems sobre algunos efectos del sistema de evaluación. Sin embargo, y por otro lado, no agrupa los ítems asociados a la relevancia de la información que solicita el sistema para la evaluación del desempeño de los investigadores, ni si los períodos entre evaluaciones son suficientes para la obtención de resultados relevantes de investigación (el ítem 21 presenta cargas factoriales significativas en 4 de los 5 factores, con cargas que van desde 0.063 a  $|0.297|$ ), ver tabla 26. Así, la conformación empírica del factor 3 no proporciona –a diferencia de los dos factores anteriores- una interpretación conceptual relevante desde el punto de vista teórico, y carece de significado ante los objetivos de investigación de este trabajo.

Adicionalmente, las cargas factoriales más altas de los ítems 11 a 14 se concentran en el factor cuatro, donde tres de los cuatro ítems tienen cargas que superan el 0.8, y el ítem que tiene la carga factorial más baja alcanza 0.611 (ver tabla 26). Estos resultados sugieren una asociación fuerte entre los ítems y el factor cuatro. Simultáneamente, el factor cuatro puede interpretarse como la relevancia de la información solicitada por el sistema para la evaluación del desempeño de investigación de los investigadores, ya que dicho factor agrupa los ítems que buscan medir la adecuación de la información solicitada por el SNI para la evaluación. Si bien la interpretación conceptual del factor cuatro es clara, su relevancia teórica y significado son limitados dados los objetivos de este trabajo.

Finalmente, las cargas factoriales más altas de los ítems 28 a 33 se concentran en el factor cinco, y van desde 0.666 hasta 0.931 (ver tabla 26). Todas esas cargas factoriales son estadísticamente significativas e importantes en magnitud, lo que sugiere una asociación fuerte entre esos ítems y el factor cinco. Desde el punto de vista de la interpretación, el factor cinco puede interpretarse como la dirección de una reforma hipotética del sistema de evaluación. Los ítems que más cargan en este factor justamente miden la percepción que tienen los investigadores sobre reformular los criterios de evaluación, dar más peso a la evaluación cualitativa, modificar la composición disciplinaria de las comisiones dictaminadoras por área, y que la evaluación considere y valore mejor la investigación colectiva, multi/inter/transdisciplinaria y aplicada. El factor cinco coincide con lo que se espera de acuerdo con la teoría que sustenta este trabajo y se denomina *dirección de la reforma del sistema de evaluación*.

En síntesis, el problema de investigación que se aborda en este trabajo es multidimensional, y una estructura factorial de cinco factores es aceptable ya que explica más del 60% de la varianza de los datos. Simultáneamente, la solución empírica obtenida del AFE -para un modelo de cinco factores- coincide parcialmente con lo que se espera de acuerdo con la teoría. En particular, tres de los cinco factores empíricos obtenidos coinciden con las conceptualizaciones teóricas dadas sobre los modos diferenciales de producción de conocimiento y con la dirección de la reforma del sistema de evaluación, tal como lo sugiere la literatura de acuerdo con las características específicas del SNI mexicano. Sin embargo, la solución empírica del AFE con cinco factores no es del todo satisfactoria, porque contiene dos factores (el tres y el cuatro) cuya relevancia teórica es limitada dados los objetivos de este trabajo. Desde el punto de vista teórico, se esperaba que los ítems que conforman los factores tres y cuatro se hubieran agrupado de manera tal que representaran dos conceptos, a saber: *satisfacción con el proceso de evaluación y flexibilidad epistemológica del sistema de evaluación*.

#### X.2.2.3 La solución de 6 factores

Debido a los resultados obtenidos en la solución del AFE con cinco factores, se exploró una solución de seis factores para observar si los ítems 11 a 27 de los factores tres y cuatro lograban re-agruparse en coincidencia con las conceptualizaciones de *satisfacción con el proceso de evaluación y flexibilidad epistemológica del sistema de evaluación*. En la tabla 27 se presenta la solución del AFE para el modelo multifactorial de seis factores.

Tal como puede observarse en la tabla 27, la solución empírica del AFE para un modelo de seis factores es menos relevante teóricamente que la solución del AFE para un modelo de cinco factores. La relevancia teórica menor de esta solución se debe a dos razones, la primera es que los ítems 11 a 27 no se re-agruparon de manera tal que fueran explicados por los conceptos de *satisfacción con el proceso de evaluación y flexibilidad epistemológica del sistema de evaluación* -tal como se esperaba teóricamente-, sino que por el contrario se mantienen con la misma estructura que en la solución del AFE con cinco factores; y la segunda razón es que los ítems 1 a 5 se des-agruparon, perdiendo así la medición del concepto propensión modo 1. En esta solución, la dimensión latente propensión modo 1 se diluye y por lo tanto no explica la variación de los ítems 1 a 5. Es así que la solución del AFE para un modelo de seis factores no resulta satisfactoria.



Tabla 27: Estructura Factorial: Solución del AFE para el modelo con 6 factores

Nro.	Ítems	Factores (solución con rotación oblicua - GEOMIN)						Varianzas Residuales
		1	2	3	4	5	6	
1	p_18_c_rec	0.621*	0.273*	0.017	0.003	0.035*	-0.008	0.501
2	p_18_f_rec	0.699*	0.095*	0.073*	0.097*	0.007	0	0.434
3	p_18_g_rec	0.384*	-0.047*	0.414*	-0.019	-0.001	-0.017	0.609
4	p_18_i_rec	0.100*	-0.071*	0.684*	-0.021	0.069*	-0.086*	0.43
5	p_18_j_rec	0.068*	0.056*	0.743*	0.007	-0.086*	0.049	0.446
6	p_18_b_rec	0.140*	0.773*	-0.037*	-0.013	0.001	-0.052*	0.383
7	p_18_d_rec	0.060*	0.845*	-0.01	-0.038*	-0.008	-0.033*	0.29
8	p_18_e_rec	-0.072*	0.618*	0.039*	0.025	-0.034	0.060*	0.602
9	p_18_h_rec	0.023	0.651*	0.040*	0.012	-0.004	0.060*	0.556
10	p_18_k_rec	-0.013	0.545*	0.351*	0.024	-0.023	0.039*	0.516
11	p_1_a_rec	-0.005	-0.035*	0.012	0.109*	0.870*	0.028	0.162
12	p_1_b_rec	0.008	-0.026*	-0.018	0.032*	0.952*	-0.028*	0.052
13	p_2_a_rec	0.027	0.119*	-0.208*	-0.033	0.696*	0.029	0.569
14	p_2_c_rec	-0.02	-0.052*	0.013	-0.029	0.996*	-0.029	0.011
15	p_5_r_d	0.018	-0.002	-0.016	0.928*	-0.016	-0.012	0.147
16	p_7_rec	0.029	-0.054*	0.001	0.788*	0.214*	0.097*	0.263
17	p_8_rec	0.039*	-0.02	-0.031*	0.973*	0.018	-0.03	0.018
18	p_9_a_rec	0.063	-0.022	-0.072	0.442*	-0.017	-0.037	0.801
19	p_9_b_rec	0.021	-0.046	-0.069	0.466*	-0.048	-0.004	0.811
20	p_9_d_rec	0.085*	-0.029	-0.043	0.803*	-0.009	0.045	0.401
21	p_3_r_d	-0.034	0.061*	0.026	0.110*	0.222*	-0.293*	0.73
22	p_10_rec	-0.002	0.049*	0.102*	0.498*	0.368*	-0.127*	0.246
23	p_11_r_d	0.060*	0.141*	0.144*	0.463*	0.270*	-0.119*	0.343
24	p_9_c_rec	-0.047	-0.002	0.014	0.654*	-0.075*	-0.114*	0.533
25	p_17_a_rec	-0.199*	0.139*	0.268*	0.477*	0.047	-0.039	0.535
26	p_17_b_rec	-0.165*	0.002	0.336*	0.476*	0.035	-0.014	0.564
27	p_17_d_rec	-0.103	0.013	0.161*	0.536*	0.041	0.055	0.654
28	p_20_g_rec	-0.037*	0.124*	-0.02	0.109*	0.015	0.914*	0.215
29	p_20_k_rec	-0.107*	0.076*	0.006	-0.046*	-0.043*	0.935*	0.002
30	p_20_l_rec	-0.073*	0.387*	-0.046*	0.009	0.038	0.670*	0.383
31	p_20_i_rec	0.073*	-0.052*	-0.021	-0.121*	-0.188*	0.826*	0.016
32	p_20_j_rec	0.021	-0.127*	0.017	0.006	-0.066*	0.828*	0.272
33	p_20_e_rec	0.053*	-0.007	0.051*	-0.395*	0.054*	0.711*	0.146

Notas: Estimaciones realizadas con M-PLUS Versión 7.4. /\* Indican que las cargas factoriales rotadas son estadísticamente significativas al nivel  $\alpha = 5\%$ .

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Finalmente y a modo de conclusión de esta sección se señala que la solución del AFE para un modelo con cinco factores es la mejor -tanto desde el punto de vista empírico como teórico- respecto de las soluciones del AFE para modelos con 4 y 6 factores. Sin embargo, a pesar de ser la mejor solución del AFE obtenida, no permite observar dos dimensiones latentes –*satisfacción con el proceso de evaluación y flexibilidad epistemológica del sistema de evaluación*- importantes y que son esperables de acuerdo con la teoría sobre la relación entre la evaluación del desempeño de investigación individual de los investigadores y la producción de conocimiento. Dado que la mejor solución empírica del AFE no permite la observación de esas dos dimensiones latentes, los resultados empíricos obtenidos pierden relevancia conceptual y no se ajustan a los objetivos de investigación de este trabajo. Dados estos resultados, a continuación, en la siguiente sección, se especifica y pone a prueba un modelo multifactorial con cinco factores para medir la relación entre la evaluación de la investigación y los modos diferenciales de producción de conocimiento. Para ello se realiza un AFC, ya que lo que se busca es confirmar (poner a prueba y validar empíricamente) la teoría de medida que sustenta este trabajo.

### X.3 Análisis Factorial Confirmatorio: El Modelo de Medida

El objetivo de esta sección es poner a prueba y validar un modelo de medida multifactorial de cinco factores para medir la relación entre la evaluación de la investigación del desempeño individual de los investigadores y los modos diferenciales de producción de conocimiento. Para ello se desarrolla un AFC que se operacionaliza para el caso del SNI de México.

A diferencia del AFE realizado en la sección anterior, aquí el análisis está orientado por la teoría. Es decir, por las hipótesis que se hacen sobre la relación entre las variables observadas de medida (ítems/indicadores) y las variables latentes (factores o constructos). En este sentido, teniendo en cuenta los resultados empíricos del análisis exploratorio de la sección anterior y el marco teórico conceptual que sustenta este trabajo (ver capítulo V), aquí se especifican *ex - ante* todos los aspectos del modelo de medida (AFC), en particular el número de variables latentes y su relación con las variables observadas de medida. En la sub-sección X.3.1 se plantean estas hipótesis, se retoman las definiciones conceptuales de las variables latentes dadas en el capítulo V y se establece su relación con las variables observadas de medida.

En el contexto del AFC, los factores representan constructos o conceptos teóricos que son inobservables directamente por su complejidad y multidimensionalidad, sin embargo son esos constructos (variables latentes) los que pueden explicar la variación de las respuestas de los investigadores a los indicadores (ítems/variables observadas de medida) individuales. En este sentido, el AFC se constituye en una herramienta analítica fundamental para la validación del modelo en general y de los constructos en particular (Brown, 2015).

Tal como se describió en el capítulo VIII, la validez de constructo implica evaluar la validez del modelo en términos de cuatro pruebas, a saber: i) validez aparente; ii) validez convergente; iii) validez discriminante; y iv) validez nomológica. Además, y a diferencia de otros métodos multivariados, el enfoque AFC permite considerar no solo los errores de la teoría (ejemplo: el término de error en un modelo de regresión lineal), sino que también considera -en el proceso de estimación y para el análisis de la validez convergente y discriminante- los errores de medición. Así los resultados del AFC proporcionan evidencia más rigurosa y más confiable sobre la validez

conceptual y empírica del modelo de medida que se está poniendo a prueba, en relación a los métodos multivariados tradicionales que no consideran el error de medición. (Brown, 2015)

Como se dijo, a continuación (X.3.1) se plantean las hipótesis que orientan el AFC y se retoman las definiciones de las variables latentes y su relación con las variables observadas de medida; en la sub-sección X.3.2 se representa gráficamente el modelo de medida conceptual a través de su diagrama de sendero y se formaliza matemáticamente. En la sub-sección X.3.3 se presentan las estimaciones obtenidas para el modelo de medida, se evalúa su bondad de ajuste y se desarrollan y analizan las pruebas de validez de constructo del modelo de medida. Los aspectos del diseño del estudio<sup>72</sup> para garantizar la obtención de resultados empíricos robustos y confiables fueron tratados en el capítulo VIII, sub-sección VIII.3.3

### X.3.1 Hipótesis del modelo de medida y definición de los constructos: las variables latentes

La teoría de medida que se está proponiendo -representada por el modelo de medida- postula en base a la teoría, los hallazgos de investigaciones pasadas (ver capítulo V) y los resultados empíricos obtenidos en la sección anterior (X.2) cinco hipótesis, a saber:

- ✓ H1: *La asociación/correlación existente entre los ítems 1 a 5 se explica por la existencia de una variable latente denominada propensión modo 1 de producción de conocimiento.*
- ✓ H2: *La asociación/correlación existente entre los ítems 6 a 10 se explica por la existencia de una variable latente denominada propensión modo 2 de producción de conocimiento.*
- ✓ H3: *La asociación/correlación existente entre los ítems 11 a 20 se explica por la existencia de una variable latente denominada satisfacción con el proceso de evaluación.*
- ✓ H4: *La asociación/correlación existente entre los ítems 21 a 27 se explica por la existencia de una variable latente denominada flexibilidad epistemológica del sistema de evaluación.*
- ✓ H5: *La asociación/correlación existente entre los ítems 28 a 33 se explica por la existencia de una variable latente denominada Dirección de la reforma del sistema de evaluación.*

La racionalidad detrás del modelo de medida que justifica las hipótesis 1 y 2 se asocia con el hecho de que las variables observadas  $y_1 - y_{10}$  (ítems 1 a 10) miden atributos específicos que son indicadores de los conceptos *propensión modo 1* y *propensión modo 2* de producción de conocimiento definidos en el capítulo V. De aquí que es esperable, desde el punto de vista teórico, que las variables latentes definidas como *propensión modo 1* y *propensión modo 2* de producción de conocimiento expliquen la variación de las respuestas de los investigadores a los indicadores individuales y su asociación  $-y_1 - y_5$  y  $y_6 - y_{10}$  respectivamente-.

Simultáneamente, la racionalidad del modelo que justifica la hipótesis 3 se vincula con el hecho de que las variables observadas  $y_{11} - y_{20}$  miden aspectos específicos que son indicadores del concepto *satisfacción con el proceso de evaluación* definido en el capítulo V. De aquí que es esperable, desde el punto de vista teórico, que la variable latente definida como *satisfacción con el proceso de evaluación* explique la variación de las respuestas de los investigadores a los indicadores individuales  $y_{11} - y_{20}$  y su asociación.

---

<sup>72</sup> Características de la muestra –representatividad, tamaño y método de muestreo-; alcance y naturaleza de los datos faltantes y el método utilizado para su tratamiento; y los aspectos vinculados a la estimación – estimador y paquete estadístico utilizado-.

A la vez, la racionalidad del modelo que justifica la hipótesis 4 tiene que ver con el hecho de que las variables observadas  $y_{21} - y_{27}$  miden aspectos específicos que son indicadores del concepto *flexibilidad epistemológica del sistema de evaluación* definido en el capítulo V. Así es esperable, desde el punto de vista teórico, que la variable latente definida como *flexibilidad epistemológica del sistema de evaluación* explique la variación de las respuestas de los investigadores a los indicadores individuales  $y_{21} - y_{27}$  y su asociación.

Finalmente, la racionalidad del modelo que justifica la hipótesis 5 tiene que ver con el hecho de que las variables observadas  $y_{28} - y_{33}$  miden aspectos específicos que son indicadores del concepto *dirección de la reforma del sistema de evaluación* definido en el capítulo V. De aquí que es esperable, desde el punto de vista teórico, que la variable latente definida como *dirección de la reforma del sistema de evaluación* explique la variación de las respuestas de los investigadores a los indicadores individuales  $y_{28} - y_{33}$  y su asociación.

En la Tabla 28 se presentan las cinco variables latentes de este trabajo, sus definiciones conceptuales y los ítems (indicadores) que las operacionalizan, junto con sus escalas de medición.

*Tabla 28: Variables Latentes o Constructos: definición conceptual, ítems y escala de medida*

Constructos/Variables Latentes	Definición Conceptual	Operacionalización			
		Nro	Notación	Ítems	Descripción de los ítems
<b>Propensión Modo 1</b>	Grado de acuerdo de los investigadores categorizados en el sistema de evaluación con aquella producción de conocimiento que se caracteriza por desarrollar procesos de investigación orientados y organizados disciplinarmente, involucrar a actores homogéneos (investigadores) y difundir el conocimiento producido por canales formales (publicación de artículos, libros y capítulos de libro), lo que representa su principal fuente de reconocimiento científico.	1	$y_1$	p_18_c_rec <sup>1,6</sup>	La investigación se orienta por disciplinario
		2	$y_2$	p_18_f_rec <sup>1,6</sup>	El conocimiento se organiza disciplinarmente
		3	$y_3$	p_18_g_rec <sup>1,6</sup>	Los actores involucrados en la generación de conocimiento son únicamente
		4	$y_4$	p_18_i_rec <sup>1,6</sup>	La principal fuente de reconocimiento es la publicación de artículos científicos
		5	$y_5$	p_18_j_rec <sup>1,6</sup>	La principal fuente de reconocimiento es el libro publicado
<b>Propensión Modo 2</b>	Grado de acuerdo de los investigadores categorizados en el sistema de evaluación con aquella producción de conocimiento que se caracteriza por desarrollar procesos de investigación orientados por la búsqueda de soluciones a problemas dependientes del contexto y organizados inter/multi/transdisciplinariamente, involucrar a actores heterogéneos (investigadores y contrapartes interesadas en la solución de problemas) y difundir el conocimiento producido por canales informales (a través de las soluciones encontradas).	6	$y_6$	p_18_b_rec <sup>1,6</sup>	El objetivo de la generación de conocimiento es su aplicación
		7	$y_7$	p_18_d_rec <sup>1,6</sup>	La generación de conocimiento se realiza por la búsqueda de soluciones a problemas complejos (productivos, sociales, ambientales, de política pública)
		8	$y_8$	p_18_e_rec <sup>1,6</sup>	El conocimiento se organiza multi/inter/transdisciplinariamente
		9	$y_9$	p_18_h_rec <sup>1,6</sup>	La generación de conocimiento se realiza por la interacción entre académicos y contrapartes interesadas en la solución de problemas
		10	$y_{10}$	p_18_k_rec <sup>1,6</sup>	La principal fuente de difusión de conocimiento es la solución encontrada (productos, procesos, marcas, patentes)
<b>Satisfacción con el proceso de evaluación</b>	Nivel de satisfacción que tienen los investigadores categorizados en el sistema de evaluación respecto a si el proceso de evaluación es capaz de reconocer el	11	$y_{11}$	p_1_a_rec <sup>2</sup>	La información que solicita el sistema es necesaria para lograr un juicio apropiado (ingreso o reingreso)

	desempeño, es decir, si cuenta con la información necesaria para elaborar juicios académicos apropiados sobre el desempeño de los investigadores e incluye todas sus contribuciones importantes, si los criterios y mecanismos de selección de los evaluadores garantizan evaluaciones académicas apropiadas en las diversas áreas de conocimiento y si la composición de las comisiones dictaminadoras respetan los equilibrios regionales e institucionales y dan garantías respecto a la objetividad de la evaluación.	12	$y_{12}$	$p_{1_b\_rec}^2$	La información que solicita el S necesaria para lograr un juicio o apropiado (reingreso vigente)
		13	$y_{13}$	$p_{2_a\_rec}^3$	La información que solicita el S
		14	$y_{14}$	$p_{2_c\_rec}^3$	La información que solicita el S aportaciones importantes
		15	$y_{15}$	$p_{5_r\_d}^4$	Los mecanismos de designación de evaluadores permiten seleccionar que garanticen una evaluación apropiada en su área de conoci
		16	$y_{16}$	$p_{7\_rec}^2$	Los requisitos para ser evaluado pertinentes
		17	$y_{17}$	$p_{8\_rec}^2$	La manera en que se integran la dictaminadores es adecuada
		18	$y_{18}$	$p_{9_a\_rec}^3$	La manera en que se integran la dictaminadoras no considera suficientemente el equilibrio re
		19	$y_{19}$	$p_{9_b\_rec}^3$	La manera en que se integran la dictaminadoras no considera suficientemente el equilibrio in
		20	$y_{20}$	$p_{9_d\_rec}^3$	La manera en que se integran la dictaminadoras no considera ot requisitos (además de estar en para dar más garantías de una objetiva.
<b>Flexibilidad epistemológica del sistema de evaluación</b>	En qué medida los investigadores categorizados en el sistema de evaluación perciben que dicho sistema de evaluación permite diversidad de perfiles y de trayectorias de investigación, procesos de investigación de largo plazo y colectivos, el desarrollo de estrategias de investigación	21	$y_{21}$	$p_{3_r\_d}^4$	Los períodos de evaluación del reingreso vigente permiten log suficientes y relevantes en la in que Ud. realiza
		22	$y_{22}$	$p_{10\_rec}^2$	Los criterios de evaluación para categoría reconocen los distinto desempeño de los investigador

	originales y novedosas, y la obtención de resultados de investigación suficientes – confiables- y científicamente relevantes.	23	$y_{23}$	$p_{11_r_d^4}$	Los criterios de evaluación le p... asumir estrategias de investiga... innovadoras
		24	$y_{24}$	$p_{9_c_rec^3}$	La integración de las comisiones... dictaminadoras no considera... suficientemente la diversidad d...
		25	$y_{25}$	$p_{17_a_rec^3}$	El SNI le generó dificultades par... involucrarse en investigaciones...
		26	$y_{26}$	$p_{17_b_rec^3}$	El SNI le generó dificultades par... equipo por la presión para publ... individualmente
		27	$y_{27}$	$p_{17_d_rec^3}$	El SNI generó estandarización d... travectorias
<b>Dirección de la reforma del sistema de evaluación</b>	El interés que tienen los investigadores categorizados en el sistema de evaluación de reformarlo en una dirección que cambia la composición disciplinaria de las comisiones dictaminadoras y los criterios de evaluación hacia una evaluación más cualitativa, con mayor reconocimiento de la investigación colectiva, mejor consideración de la investigación multi/inter/transdisciplinaria y que valore más los resultados de la investigación aplicada u orientada a la resolución de problemas.	28	$y_{28}$	$p_{20_g_rec^2}$	Dar mayor reconocimiento a la... colectiva
		29	$y_{29}$	$p_{20_k_rec^2}$	Considerar adecuadamente la i... multi/inter/transdisciplinaria
		30	$y_{30}$	$p_{20_l_rec^2}$	Dar más peso a los resultados d... investigación aplicada
		31	$y_{31}$	$p_{20_i_rec^2}$	Reformular los criterios de eval...
		32	$y_{32}$	$p_{20_j_rec^2}$	Modificar la composición discip... comisiones dictaminadoras por...
		33	$y_{33}$	$p_{20_e_rec^2}$	Dar más peso a la evaluación cu...

Notas: (1) A estos ítems se les revirtió la escala. (2) A estos ítems se les recodificó la escala. Las respuestas en la categoría Dato Faltante. (3) A estos ítems se les transformó la variable - pasando de nominal a dicotómica-, y se les revirtió la categoría "No Sabe" se tratan como Dato Faltante. (5) A estos ítems se les transformó la variable - pasando de ordinal a dicotómica-, y se les recodificó la escala. Las respuestas en las categorías "No Sabe" y/o "No corresponde" se tratan como Dato Faltante. (6) A estos ítems con valores de asimetría < 2 y curtosis < 7 (Finney y Distefano, 2006:298). Los restantes ítems son no normales (Finney y Distefano, 2006).

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT)



### X.3.2 Modelo de Medida: representación gráfica y formalización matemática

El modelo de medida que se propone, formalmente se puede escribir como:

$$y_i^* = \Lambda \eta_i + \epsilon_i \quad (1)$$

Donde:

- ✓  $y_i^*$  es un vector  $p$  dimensional de variables de respuesta latente, con  $p = 33$ . Es decir, se supone que para cada variable de medida observada  $y_j$  (con  $j = 1, 2, 3, \dots, 33$ ) existe una variable de respuesta latente  $y_j^*$ , tal que:

$$y_{ij} = \begin{cases} 1 & \text{si } y_{ij}^* \leq \tau_1 \\ 2 & \text{si } \tau_1 < y_{ij}^* \leq \tau_2 \\ 3 & \text{si } \tau_2 < y_{ij}^* \leq \tau_3 \\ 4 & \text{si } \tau_3 < y_{ij}^* \leq \tau_4 \\ 5 & \text{si } y_{ij}^* > \tau_4 \end{cases} \quad \text{Con } j = 1, 2, 3, \dots, 10 \text{ e } i = 1, 2, 3, \dots, 8053 \text{ observaciones.}$$

$$y_{ij} = \begin{cases} 0 & \text{si } y_{ij}^* \leq \tau_1 \\ 1 & \text{si } y_{ij}^* > \tau_1 \end{cases} \quad \text{Con } j = 11, 12, 13, \dots, 33 \text{ e } i = 1, 2, 3, \dots, 8053 \text{ observaciones.}$$

- ✓  $\eta$  es un vector aleatorio  $m$  dimensional de las variables latentes (constructos o factores), con media cero,  $m = 5$ . Los elementos del vector  $\eta$  son: *Propensión Modo 1, Propensión Modo 2, Satisfacción con el Proceso de Evaluación, Flexibilidad Epistemológica del Sistema de Evaluación y Dirección de la Reforma del Sistema de Evaluación*.
- ✓  $\epsilon$  es un vector aleatorio  $p$  dimensional de las varianzas residuales de  $y^*$  o errores de medición, con media cero y matriz de covarianza  $\Theta$ , con  $p = 33$ . Los errores de medida se asumen independientes entre sí y respecto a  $\eta$ .
- ✓  $\Lambda$  es una matriz de  $p \times m$  de los parámetros de pendiente de medida o cargas factoriales, con  $p = 33$  y  $m = 5$ .
- ✓  $\Theta$  es la matriz de covarianza de  $\epsilon$  de  $33 \times 33$ .

La figura 4 representa el modelo de medida, que se está poniendo a prueba, a través de un diagrama de senderos. Tal como se desprende del diagrama, se trata de un modelo de medida multifactorial con cinco constructos latentes (variables latentes ( $\eta$ ) *reflexivos* y treinta y tres variables de medida observadas ( $y$ ) a las que se les asocia 33 variables de respuesta latente ( $y^*$ ). Los constructos son reflexivos porque, como se desarrolló en la sub-sección anterior, se espera que cada uno de ellos explique la variación de las respuestas de los investigadores a los indicadores individuales (variables de medida observadas dependientes) y su asociación. La “explicación” que cada variable latente provee de las variaciones y asociación de cada una de las variables de medida observadas queda representada a través de las flechas que van desde cada uno de los constructos –representados en círculos- hacia cada una de las variables de medida observadas (ítems/indicadores) –representadas en rectángulos-. Conceptualmente, se trata de las hipótesis planteadas en la sub-sección anterior; en el marco conceptual del modelo se denominan como *relaciones de medida*, cuyos valores estimados estarán dados por las cargas factoriales ( $-\lambda_j$ - parámetros de pendiente derivados del conjunto de regresiones probit).

Además de estas relaciones de medida, en el modelo también se hipotetiza que las variables latentes correlacionan entre sí. A diferencia del AFE, en el contexto del AFC se impone la restricción de que cada una de las variables de medida observadas es explicada/causada por una única variable

latente, es decir, no se permiten cargas factoriales cruzadas. Sin embargo, sí pueden existir relaciones entre las variables de medida que están asociadas a constructos diferentes (ver tabla 22). Se predice entonces que dichas relaciones pueden ser explicadas por la correlación entre las variables latentes, las que quedan representadas en la figura 4 a través de las flechas bidireccionales que unen a cada uno de los constructos del modelo con los restantes. Así, y de acuerdo con el marco teórico conceptual (capítulo V) que sustenta este trabajo, es razonable esperar que:

- i) los modos diferenciales de producción de conocimiento sean complementarios entre sí (que exista una asociación significativa y positiva entre las *propensiones modo 1 y modo 2* de producción de conocimiento);
- ii) la *satisfacción con el proceso de evaluación* que tienen los investigadores, su percepción sobre la *flexibilidad epistemológica del sistema de evaluación* y su interés en la *dirección de la reforma del sistema de evaluación* no sean independientes de las *propensiones modo 1 y modo 2* de producción de conocimiento (que exista una asociación significativa);
- iii) el interés de los investigadores en la *dirección de la reforma del sistema de evaluación* se asocie significativamente y negativamente con la *satisfacción con el proceso de evaluación* que tienen y con su percepción sobre la *flexibilidad epistemológica del sistema de evaluación*;
- iv) la *satisfacción con el proceso de evaluación* que tienen los investigadores y su percepción sobre la *flexibilidad epistemológica del sistema de evaluación* sean complementarias entre sí (que exista una asociación significativa y positiva entre la *flexibilidad epistemológica del sistema de evaluación* percibida y la *satisfacción con el proceso de evaluación* que tienen).

Desde el punto de vista de la identificabilidad, el modelo de medida propuesto está *sobre-identificado*. De acuerdo con Muthén y Asparouhov (2002: 5), el número máximo de parámetros del modelo identificables está dado por la siguiente expresión:

$$\left(\frac{p(p-1)}{2}\right) + r$$

Dónde  $p$  es el número de variables del modelo y  $r$  es el número de umbrales. En este caso,  $p = 33$  y  $r = 63$ , por lo tanto es posible identificar un total de 591 parámetros. Sin embargo, el modelo de medida propuesto, requiere únicamente de la identificación de 106 parámetros libres a estimar, a saber:

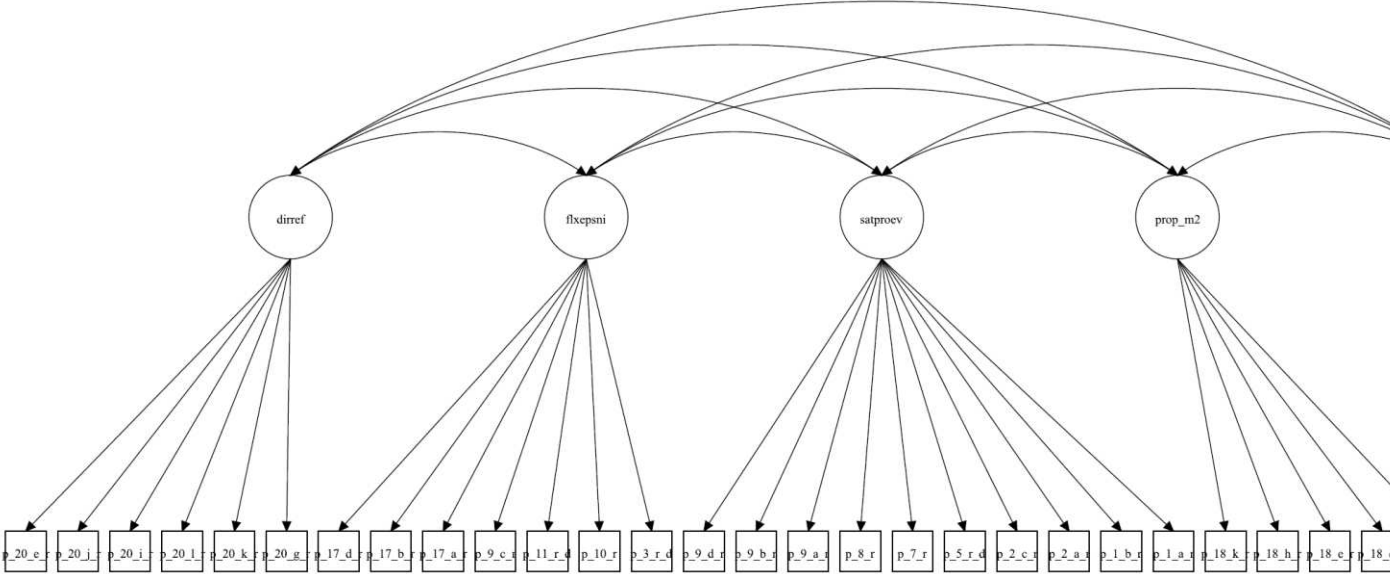
- ✓ 28 cargas factoriales ( $\lambda$ ) o relaciones de medida entre las variables de medición y las variables latentes (para fijar la escala, la carga factorial más grande por variable latente se deja constante e igual a 1);
- ✓ 63 umbrales ( $\tau$ ) -  $c - 1$  categorías por cada variable categórica observada-; y
- ✓ 15 relaciones entre las variables latentes –constructos- (5 varianzas: una por cada variable latente ( $\psi$ ), más 10 correlaciones únicas entre las variables latentes).

Por lo tanto, el modelo de medida propuesto cuenta con 485 grados de libertad ( $591-106=485$ ), dando lugar a un modelo de medida *sobre-identificado*.

Los errores de medición se calculan como residuos, fijando la varianza de  $y^*$  en 1 (*parametrización delta*).

Finalmente, vale la pena señalar que, a diferencia del modelo AFE de la sección X.2, el modelo de medida propuesto restringe y fija en cero todas las cargas factoriales cruzadas.

Figura 4: Diagrama de sendero del Modelo de Medida



Fuente: Elaboración propia

### X.3.3 Validez del Modelo de Medida

El objetivo de esta sección es evaluar la validez del modelo de medida propuesto. Para ello, por un lado se analiza y compara la teoría de medida propuesta en el modelo contra la realidad - representada por los datos muestrales-, y por otro, se estudia su validez de constructo. Se trata de responder a la pregunta: *¿el modelo de medida es exacto, mide lo que busca medir?*

Es así que la validez del modelo de medida depende de dos evaluaciones, una sobre la bondad de ajuste, y la otra, sobre la validez de constructo. Los resultados del AFC permiten confirmar si el modelo de medida teórico propuesto explica la realidad y si es válido. A continuación se presentan estos resultados. La salida completa de la estimación del AFC está disponible en el Anexo 2.

#### X.3.3.1 Evaluación de la Bondad de Ajuste del Modelo de Medida

A continuación, en la tabla 29, se presentan los resultados obtenidos para el estadístico de prueba Ji-Cuadrado de bondad de ajuste ( $X^2$ ), sus grados de libertad ( $df$ ) y el tamaño de la muestra ( $N$ ).

Tabla 29: Estadístico de prueba de bondad de ajuste Ji-Cuadrado

	$X^2$
Valor	10890.792
$df$	485
$p - value$	0.0000 <sup>1</sup>
$N$	8053
Parámetros libres estimados	106

Notas: 1. Estadísticamente significativa al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Como era esperable, el estadístico Ji-Cuadrado es significativo. La evidencia sugiere que existen diferencias significativas entre las matrices de correlación observada y estimada. Sin embargo, debido a la sensibilidad de la Ji-Cuadrado a los grados de libertad del modelo y al tamaño de la muestra, a continuación, y para concluir sobre la bondad de ajuste del modelo de medida, se presentan los resultados obtenidos para el RMSEA, TLI y CFI.

En la tabla 30 se presentan los resultados obtenidos de estas tres medidas de ajuste. Allí también pueden observarse los valores del estadístico de prueba Ji-Cuadrado de bondad de ajuste para el modelo nulo, con sus respectivos grados de libertad y  $p-value$ .

Tabla 30: Índices complementarios de Bondad de Ajuste

Índice	Valor Estimado
Medidas de Ajuste Absoluto	
RMSEA	0.052
Intervalo de Confianza 90% para el valor estimado del RMSEA	(0.051; 0.052)
Probabilidad $RMSEA \leq 0.05$	0.001
Medidas de Ajuste Incremental	
CFI	0.955
TLI	0.951
Ji-Cuadrado del Modelo Nulo	
Valor	230934.785
$df$	528
$p - value$	0.0000 <sup>1</sup>

<i>N</i>	8053
----------	------

Notas: 1. Estadísticamente significativa al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.  
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

De acuerdo con las medidas complementarias de bondad de ajuste reportadas en la tabla 30, la evidencia sugiere que el modelo de medida multifactorial de cinco factores teórico propuesto para medir la relación entre la evaluación de la investigación y la producción de conocimiento presenta un buen ajuste.

Más allá de los resultados obtenidos en el AFE en cuanto al número de factores retenidos (ver subsección X.2.2), aquí se realiza la prueba de la diferencia de la Ji-Cuadrado para modelos anidados (Brown, 2015). El objetivo es demostrar que el modelo de medida multifactorial de cinco factores propuesto ajusta significativamente mejor que un modelo de medida unidimensional (de un solo factor). Para ello, el procedimiento debe realizarse en dos etapas<sup>73</sup> porque la diferencia de los valores de la Ji-Cuadrado entre modelos anidados, que son estimados con WLSMV, no se distribuye como una Ji-Cuadrado (Brown, 2015; Muthén y Muthén, 1998-2015; Muthén, 1998-2004). A continuación se presenta la tabla 31 con todos los resultados obtenidos.<sup>74</sup> En el Anexo 3 se presenta la salida completa de la estimación de la prueba de la diferencia de la Ji-cuadrado entre el modelo multifactorial de cinco factores vs el modelo unidimensional.

*Tabla 31: Prueba de la diferencia de la Ji-Cuadrado para modelos anidados y medidas de Bondad de Ajuste del modelo de medida unifactorial*

<b>Índice</b>	<b>Valor Estimado</b>
<i>Parámetros libres estimados</i>	96
<i>N</i>	8053
<b>Prueba Ji-Cuadrado del Modelo Unifactorial (H<sub>0</sub>)</b>	
<i>Valor</i>	48802.108
<i>df</i>	495
<i>p – value</i>	0.0000 <sup>1</sup>
<b>Prueba Ji-Cuadrado para la diferencia (H<sub>0</sub> vs H<sub>1</sub>)</b>	
<i>X<sup>2</sup> para la diferencia</i>	12119.064
<i>df</i>	10
<i>p – value</i>	0.0000 <sup>1</sup>
<b>Medidas de Ajuste Absoluto del Modelo Unifactorial (H<sub>0</sub>)</b>	
<i>RMSEA</i>	0.110
Intervalo de Confianza 90% para el valor estimado del RMSEA	(0.110; 0.111)
Probabilidad RMSEA ≤ 0.05	0.000
<b>Medidas de Ajuste Incremental del Modelo Unifactorial (H<sub>0</sub>)</b>	
CFI	0.790
TLI	0.776

<sup>73</sup> En la primera etapa se estimó el modelo de medida multifactorial de cinco factores que se propone, es decir, el que cuenta con menor cantidad de restricciones (H<sub>1</sub>), y sus resultados fueron almacenados. En la segunda etapa, se estimó el modelo más restringido (H<sub>0</sub>) -el que hipotetiza que la relación entre la evaluación de la investigación y la producción de conocimiento puede ser medida por un solo factor-, y se utilizaron todos los resultados, incluidos los almacenados en la etapa 1, para calcular el valor de la prueba de la diferencia de la Ji-Cuadrado.

<sup>74</sup> Para una revisión formal, técnica y detallada de los cálculos y procedimientos para hallar el valor de la prueba de la diferencia de la Ji-Cuadrado ver Muthén (1998-2004).

Ji-Cuadrado del Modelo Nulo	
<i>Valor</i>	230934.785
<i>df</i>	528
<i>p – value</i>	0.0000 <sup>1</sup>
<i>N</i>	8053

Notas: 1. Estadísticamente significativa al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.  
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (Bensusán et al., 2014; FCCyT, 2014).

Como era esperable y se desprende de la tabla 31, la evidencia sugiere que la restricción de un único factor para explicar la relación entre la evaluación de la investigación y la producción de conocimiento, no solo degrada significativamente el ajuste del modelo de medida sino que resulta inadmisibile.

En la siguiente sub-sección se presentan los resultados obtenidos de la estimación de los parámetros (sin estandarizar y estandarizados) del modelo de medida para estudiar la aceptabilidad y utilidad de la solución encontrada, y conjuntamente se desarrollan todas las pruebas para evaluar su validez de constructo.

#### X.3.3.2 Evaluación de la Validez de Constructo

La *validez de constructo* se define como el grado en el que un conjunto de ítems de medida representan *verdaderamente* a los constructos latentes teóricos, que esos ítems buscan para medir (Hair, et al.2014: 618). Se trata de evaluar la precisión y en qué magnitud las medidas son exactas. En ese sentido, si se existe evidencia que sugiere validez de constructo, se puede afirmar con confianza que los ítems de medida –observados en una muestra representativa- representan los valores verdaderos que existen en toda la población.

Como se dijo y se describió detalladamente en el capítulo VIII, la evaluación de la validez de constructo del modelo de medida consiste en cuatro pruebas de validación: i) prueba de *validez aparente*; ii) prueba de *validez convergente* que supone el análisis conjunto y complementario de las *cargas factoriales estimadas no estandarizadas*, las *cargas factoriales estimadas estandarizadas*, la *Varianza Extraída Promedio* (AVE por sus siglas en inglés) y la *Confiabilidad Compuesta* (CR por sus siglas en inglés); iii) prueba de *validez discriminante*; y, iv) prueba de *validez nomológica*.

En el contexto de esta investigación, el panel de expertos externos –para realizar la prueba de validez aparente- está integrado por los miembros del tribunal de esta tesis, quienes en el marco del coloquio desarrollado en Enero de 2016, hicieron una primera evaluación de la validez aparente del modelo de medida propuesto. Dicha evaluación, dio como resultado que el contenido de los constructos es válido. Es decir, los evaluadores externos encontraron un alto grado de coincidencia entre las definiciones conceptuales de los constructos y los ítems que los componen.

En síntesis:

- ✓ Los constructos tienen validez aparente.

Tabla 32: Estimación del Modelo de Medida para la relación entre la evaluación de la investigación individual y la producción de conocimiento –solución sin estandarizar-

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.823	0.019	42.219	0.000
P_18_F_R	0.844	0.019	43.401	0.000
P_18_G_R	0.783	0.019	42.331	0.000
P_18_I_R	1.000	0.000	999.000	999.000
P_18_J_R	0.882	0.020	44.695	0.000
PROP_M2 BY				
P_18_B_R	0.937	0.011	83.477	0.000
P_18_D_R	1.000	0.000	999.000	999.000
P_18_E_R	0.716	0.013	56.636	0.000
P_18_H_R	0.795	0.011	69.377	0.000
P_18_K_R	0.827	0.011	73.047	0.000
SATPROEV BY				
P_1_A_R	0.897	0.007	123.275	0.000
P_1_B_R	0.994	0.008	127.770	0.000
P_2_A_R	0.481	0.025	19.038	0.000
P_2_C_R	1.000	0.000	999.000	999.000
P_5_R_D	0.914	0.008	111.814	0.000
P_7_R	0.844	0.010	82.058	0.000
P_8_R	0.996	0.008	121.154	0.000
P_9_A_R	0.386	0.028	13.738	0.000
P_9_B_R	0.339	0.031	11.018	0.000
P_9_D_R	0.686	0.016	41.716	0.000
FLXEPSNI BY				
P_3_R_D	0.588	0.016	36.599	0.000
P_10_R	1.000	0.000	999.000	999.000
P_11_R_D	0.915	0.013	70.812	0.000
P_9_C_R	0.656	0.020	32.110	0.000
P_17_A_R	0.695	0.023	30.519	0.000
P_17_B_R	0.671	0.025	27.198	0.000
P_17_D_R	0.589	0.036	16.309	0.000
DIRREF BY				
P_20_G_R	0.831	0.008	107.720	0.000
P_20_K_R	0.986	0.008	120.106	0.000
P_20_L_R	0.637	0.011	56.008	0.000
P_20_I_R	1.000	0.000	999.000	999.000
P_20_J_R	0.921	0.008	122.170	0.000
P_20_E_R	0.840	0.009	96.921	0.000

Notas: 1. Estimaciones realizadas con M-PLUS Versión 7.4.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

#### Prueba de Validez Convergente:

(a) *significancia estadística de las cargas factoriales sin estandarizar*: De acuerdo con las estimaciones realizadas, la evidencia sugiere validez convergente. Tal como puede observarse en la tabla 32, todas las cargas factoriales no estandarizadas son estadísticamente significativas. Los ítems P\_18\_I\_R, P\_18\_D\_R, P\_2\_C\_R, P\_10\_R y P\_20\_I\_R – que se corresponden con cada uno de los constructos latentes- se dejaron constantes y fijaron en 1 para establecer la escala de medida de los constructos. El criterio para fijar esos ítems indicadores se vincula con el hecho de que presentaban la carga factorial no estandarizada más alta. Las cargas factoriales son las pendientes de los modelos probit estimados para cada variable de respuesta latente ( $y_j^*$ ) asociadas a cada una de las variables categóricas (dicotómicas y ordinales) observadas.



Tabla 33: Estimación del Modelo de Medida para la relación entre la evaluación de la investigación individual y la producción de conocimiento –solución completamente estandarizada-

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.584	0.011	53.301	0.000
P_18_F_R	0.599	0.011	55.848	0.000
P_18_G_R	0.556	0.011	51.084	0.000
P_18_I_R	0.710	0.009	75.931	0.000
P_18_J_R	0.626	0.010	62.046	0.000
PROP_M2 BY				
P_18_B_R	0.771	0.007	114.790	0.000
P_18_D_R	0.823	0.006	127.568	0.000
P_18_E_R	0.589	0.010	61.525	0.000
P_18_H_R	0.654	0.008	77.680	0.000
P_18_K_R	0.680	0.008	84.493	0.000
SATPROEV BY				
P_1_A_R	0.872	0.006	143.130	0.000
P_1_B_R	0.967	0.004	222.082	0.000
P_2_A_R	0.468	0.024	19.821	0.000
P_2_C_R	0.972	0.004	246.812	0.000
P_5_R_D	0.889	0.007	127.724	0.000
P_7_R	0.821	0.009	87.335	0.000
P_8_R	0.968	0.007	141.638	0.000
P_9_A_R	0.376	0.027	13.769	0.000
P_9_B_R	0.329	0.030	11.043	0.000
P_9_D_R	0.667	0.016	42.503	0.000
FLXEPSNI BY				
P_3_R_D	0.546	0.014	38.337	0.000
P_10_R	0.928	0.008	121.983	0.000
P_11_R_D	0.849	0.009	98.690	0.000
P_9_C_R	0.609	0.018	33.088	0.000
P_17_A_R	0.645	0.020	31.618	0.000
P_17_B_R	0.623	0.022	27.783	0.000
P_17_D_R	0.547	0.033	16.414	0.000
DIRREF BY				
P_20_G_R	0.822	0.007	115.740	0.000
P_20_K_R	0.976	0.005	195.047	0.000
P_20_L_R	0.630	0.011	56.745	0.000
P_20_I_R	0.989	0.005	202.846	0.000
P_20_J_R	0.911	0.006	157.781	0.000
P_20_E_R	0.831	0.008	108.179	0.000

Notas: 1. Estimaciones realizadas con M-PLUS Versión 7.4.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Desde el punto de vista de la magnitud, 16 cargas factoriales estandarizadas se encuentran entre 0.710 y 0.989, lo que representa una magnitud muy satisfactoria porque las variables latentes explican entre el 50.4% y el 97.8% de la variación de esos ítems ( $y_j^*$ ); 14 cargas factoriales estandarizadas se encuentran entre 0.546 y 0.680, representando una magnitud aceptable (las variables latentes explican entre el 30% y el 46% de la variación de esos ítems ( $y_j^*$ ); y las restantes 3 cargas factoriales estandarizadas se encuentran entre 0.329 y 0.468, que aunque son aceptables en el contexto de las ciencias sociales, representan magnitudes relativamente bajas porque las

#### Prueba de Validez Convergente:

(b) *significancia estadística de las cargas factoriales estandarizadas, dirección y magnitud*: De acuerdo con las estimaciones realizadas, la evidencia sugiere validez convergente.

Tal como puede observarse en la tabla 33, todas las cargas factoriales estandarizadas son estadísticamente significativas y positivas. Es decir, consistentes con lo esperado de acuerdo con la definición conceptual de los constructos. Las cargas factoriales positivas de la solución completamente estandarizada (STDYX) se interpretan como un aumento en la puntuación estándar de la variable latente  $\eta_1$  se asocia con un aumento de la puntuación estándar de  $\lambda_{y_j^*1}$  en  $y_j^*$ . A modo de ejemplo: un aumento en la puntuación estándar de *propensión modo 1* de producción de conocimiento (PROP\_M1) se asocia con un aumento en la puntuación estándar de 0.584 en que la investigación se orienta por su campo disciplinario (P\_18\_C\_R). Para obtener la solución completamente estandarizada, en la parametrización *delta*, la varianza de las variables latentes se estandarizan en 1.

variables latentes solo pueden explicar entre el 11% y el 22% de la varianza de esos ítems ( $y_j^*$ ). Los ítems: la información que solicita el SNI no es excesiva (P\_2\_A\_R) y la manera en que se integran las comisiones dictaminadoras consideran suficientemente los equilibrios regional (P\_9\_A\_R) e institucional (P\_9\_B\_R) son los que tienen las cargas factoriales más bajas y son indicadores del constructo *satisfacción con el proceso de evaluación* (SATPROEV).

Tabla 34:  $R^2$  y Varianzas Residuales

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Residual Variance
P_18_C_R	0.341	0.013	26.651	0.000	0.659
P_18_F_R	0.358	0.013	27.924	0.000	0.642
P_18_G_R	0.309	0.012	25.542	0.000	0.691
P_18_I_R	0.503	0.013	37.966	0.000	0.497
P_18_J_R	0.391	0.013	31.023	0.000	0.609
P_18_B_R	0.594	0.01	57.395	0.000	0.406
P_18_D_R	0.677	0.011	63.784	0.000	0.323
P_18_E_R	0.346	0.011	30.763	0.000	0.654
P_18_H_R	0.427	0.011	38.84	0.000	0.573
P_18_K_R	0.463	0.011	42.247	0.000	0.537
P_1_A_R	0.761	0.011	71.565	0.000	0.239
P_1_B_R	0.935	0.008	111.041	0.000	0.065
P_2_A_R	0.219	0.022	9.911	0.000	0.781
P_2_C_R	0.946	0.008	123.406	0.000	0.054
P_5_R_D	0.791	0.012	63.862	0.000	0.209
P_7_R	0.673	0.015	43.667	0.000	0.327
P_8_R	0.937	0.013	70.819	0.000	0.063
P_9_A_R	0.141	0.02	6.884	0.000	0.859
P_9_B_R	0.109	0.02	5.522	0.000	0.891
P_9_D_R	0.446	0.021	21.252	0.000	0.554
P_3_R_D	0.298	0.016	19.169	0.000	0.702
P_10_R	0.861	0.014	60.991	0.000	0.139
P_11_R_D	0.721	0.015	49.345	0.000	0.279
P_9_C_R	0.371	0.022	16.544	0.000	0.629
P_17_A_R	0.417	0.026	15.809	0.000	0.583
P_17_B_R	0.388	0.028	13.891	0.000	0.612
P_17_D_R	0.299	0.036	8.207	0.000	0.701
P_20_G_R	0.675	0.012	57.87	0.000	0.325
P_20_K_R	0.952	0.01	97.524	0.000	0.048
P_20_L_R	0.396	0.014	28.372	0.000	0.604
P_20_I_R	0.978	0.01	101.423	0.000	0.022
P_20_E_R	0.691	0.013	54.089	0.000	0.309
P_20_J_R	0.83	0.011	78.891	0.000	0.17

Notas: Estimaciones realizadas con M-PLUS Versión 7.4.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCyT, 2014).

**Prueba de Validez Convergente:** (c) *Varianza Extraída Promedio (AVE)* y (d) *Confiabilidad Compuesta (CR)*

La evidencia sugiere -de acuerdo con las medidas AVE y CR- validez convergente aceptable, constructos confiables y consistentes internamente. Los valores AVE y CR calculados para cada variable latente se presentan en la tabla 35.

Tabla 35: Varianza Extraída Promedio y Confiabilidad Compuesta

<b>Variables Latentes –constructos-</b>	<b>AVE</b>	<b>CR</b>
Propensión Modo 1 (PROP_M1)	<b>0.3810</b>	<b>0.7532</b>
Propensión Modo 2 (PROP_M2)	<b>0.5018</b>	<b>0.8323</b>
Satisfacción con el proceso de evaluación (SATPROEV)	<b>0.5955</b>	<b>0.9300</b>
Flexibilidad epistemológica del sistema de evaluación (FLXEPSNI)	<b>0.4792</b>	<b>0.8608</b>
Dirección de la reforma del sistema de evaluación (DIRREF)	<b>0.7540</b>	<b>0.9474</b>

Notas: Cálculos basados en las estimaciones realizadas con M-PLUS Versión 7.4.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Debido a los altos valores de CR obtenidos es posible concluir no solamente que la convergencia es aceptable, que los constructos son confiables, sino además que las variables observadas de medida consistentemente representan al mismo constructo.

Desde el punto de vista de la varianza extraída promedio de cada constructo, los constructos con una comunalidad promedio mayor son: *dirección de la reforma del sistema de evaluación* (DIRREF), ya que en promedio, el 75.4% de la varianza de los ítems ( $y_j^*$ ) que lo componen, es varianza explicada por dicha variable latente; le siguen *satisfacción con el proceso de evaluación* (SATPROEV) y *propensión modo 2* (PROP\_M2), que en promedio explican el 59.5% y 50.2% de la variación de sus ítems ( $y_j^*$ ) respectivamente. Los constructos flexibilidad epistemológica del sistema de evaluación (FLXEPSNI) y *propensión modo 1* (PROP\_M1) son los que presentan una comunalidad menor, ya que en promedio explican el 47.9% y 38.1% de la variación de sus ítems ( $y_j^*$ ) respectivamente. Si bien la estructura teórica propuesta para la medición -dados los valores AVE calculados y los otros indicadores ya analizados- es más que aceptable, importa señalar que estos últimos dos constructos son menos precisos que los restantes porque la mayoría de la variación promedio de sus ítems queda contabilizada en los errores de medición (ver varianzas residuales en tabla 34).

En síntesis:

- ✓ la evidencia es fuerte y sugiere validez convergente del modelo de medida teórico propuesto.

**Prueba de Validez Discriminante:** *¿los constructos miden fenómenos diferentes, discriminan entre sí?*

La prueba más rigurosa para evaluar la validez discriminante de los constructos es a través de la comparación de sus valores AVE contra la correlación estimada al cuadrado entre las variables latentes. En la tabla 36 se presenta la matriz con las correlaciones estimadas entre los constructos (triángulo por debajo de la diagonal) y sus valores al cuadrado *versus* los valores AVE (triángulo por encima de la diagonal).

Tabla 36: Correlaciones entre las Variables Latentes

Corr \ Corr <sup>2</sup>	PROP_M1	PROP_M2	SATPROEV	FLXEPSNI	DIRREF
PROP_M1		0.148225< <b>0.3810</b>	0.076176< <b>0.3810</b>	0.148225< <b>0.3810</b>	0.094864< <b>0.3810</b>
PROP_M2	0.385** (0.012)		0.013225< <b>0.5018</b>	0.0576< <b>0.5018</b>	0.002116< <b>0.5018</b>
SATPROEV	0.276** (0.014)	0.115** (0.014)		0.568516< <b>0.5955</b>	0.339889< <b>0.5955</b>
FLXEPSNI	0.385** (0.014)	0.240** (0.015)	0.754** (0.009)		0.455625< <b>0.4792</b>
DIRREF	-0.308** (0.015)	0.046** (0.016)	-0.583** (0.010)	-0.675** (0.011)	

Notas: Estimaciones realizadas con M-PLUS Versión 7.4/\*\*Correlaciones significativas con una confianza del 99%/Errores Estándar entre paréntesis/En **negrita valores AVE**.

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Como se desprende de la tabla 36, todos los constructos explican más la variación de sus ítems que lo que comparten con los restantes constructos. En todos los casos, los valores AVE de cada constructo son mayores que las correlaciones al cuadrado entre los constructos. Esto significa que la evidencia sugiere que los constructos miden fenómenos distintos y discriminan entre sí, a saber:

- ✓ el constructo *propensión modo 1* mide un fenómeno distinto y por lo tanto, discrimina respecto a los fenómenos medidos por *propensión modo 2*, *satisfacción con el proceso de evaluación*, *flexibilidad epistemológica del sistema de evaluación* y *dirección de la reforma del sistema de evaluación*;
- ✓ el constructo *propensión modo 2* mide un fenómeno distinto y por lo tanto, también discrimina respecto a los fenómenos medidos por *satisfacción con el proceso de evaluación*, *flexibilidad epistemológica del sistema de evaluación* y *dirección de la reforma del sistema de evaluación*;
- ✓ el constructo *satisfacción con el proceso de evaluación* mide un fenómeno distinto y por lo tanto, también discrimina respecto a los fenómenos medidos por *flexibilidad epistemológica del sistema de evaluación* y *dirección de la reforma del sistema de evaluación*; y por último,
- ✓ el constructo *flexibilidad epistemológica del sistema de evaluación* mide un fenómeno distinto y por lo tanto, también discrimina respecto al fenómeno medido por *dirección de la reforma del sistema de evaluación*.

En síntesis:

- ✓ los constructos son únicos y capturan fenómenos distintos. Cada uno de los constructos mide lo que los otros constructos no son capaces de medir. De acuerdo con la evidencia, se concluye que los constructos discriminan entre sí y por lo tanto, el modelo de medida propuesto tiene validez discriminante.

**Prueba de Validez Nomológica:** *¿las correlaciones entre los constructos son consistentes con la teoría?*

En la sección X.3.2 se plantearon un conjunto de hipótesis –orientadas por la teoría- en relación a las correlaciones esperadas entre los constructos o variables latentes. Aquí y de acuerdo a las estimaciones realizadas –presentadas en la tabla 36-, dichas hipótesis son confirmadas empíricamente. En particular la evidencia muestra que:

- ✓ los modos diferenciales de producción de conocimiento son complementarios entre sí, de hecho la correlación entre las *propensiones modo 1 y modo 2* (PROP\_M1 y PROP\_M2) es positiva (0.385) y estadísticamente significativa;
- ✓ la *satisfacción con el proceso de evaluación* (SATPROEV) que tienen los investigadores, su percepción sobre la *flexibilidad epistemológica del sistema de evaluación* (FLXEPSNI) y el interés que tienen en la *dirección de la reforma del sistema de evaluación* (DIRREF) no son independientes de sus *propensiones modo 1 y modo 2* (PROP\_M1 y PROP\_M2). De hecho se observaron correlaciones diferenciales y significativas entre esos constructos. En particular, mientras que la *satisfacción con el proceso de evaluación* exhibe una asociación más fuerte que alcanza un valor de 0.276 con *propensión modo 1*, esa asociación cae a poco más de la mitad (0.115) cuando se trata de la *propensión modo 2* de producción de conocimiento. En base a la evidencia se puede afirmar que los investigadores que presentan una *propensión modo 2* de producción de conocimiento mayor, tienen una *satisfacción con el proceso de evaluación* positiva pero de menor magnitud que los investigadores cuya *propensión modo 1* es mayor. Así, la satisfacción que tienen los investigadores con un determinado proceso de evaluación del desempeño de la investigación está asociada con los modos diferenciales de producción de conocimiento que orientan su trabajo. Similarmente, la percepción sobre la *flexibilidad epistemológica del sistema de evaluación* exhibe una asociación estadísticamente significativa y positiva -pero diferencial- con los modos diferenciales de producción de conocimiento. Mientras que la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores exhibe una asociación más fuerte que alcanza un valor de 0.385 con *propensión modo 1*, esa asociación baja a 0.240 cuando se trata de *propensión modo 2*. Es así que los investigadores que presentan una *propensión modo 2* de producción de conocimiento mayor, perciben una *flexibilidad epistemológica del sistema de evaluación* positiva pero de menor magnitud que los investigadores cuya *propensión modo 1* es mayor. La flexibilidad que perciben los investigadores en términos de las características epistemológicas que permite un sistema de evaluación -es decir, si el sistema de evaluación permite la obtención de resultados de investigación relevantes, originales y de calidad, contempla procesos de investigación de largo plazo, colectivos y que promueve la diversidad de perfiles de investigación-, está asociada a los modos diferenciales de producción de conocimiento que orientan su trabajo. Simultáneamente y siguiendo la misma racionalidad, el interés de los investigadores en la *dirección de la reforma del sistema de evaluación* también está asociada con sus modos diferenciales de producción de conocimiento, aunque en direcciones contrarias. En particular y consistente con la teoría, mientras que la *dirección de la reforma del sistema de evaluación* exhibe una asociación significativa y negativa que alcanza un valor de -0.308 con *propensión modo 1*, esa asociación –aunque muy baja: 0.046- se convierte en positiva en

relación a *propensión modo 2*. Es decir, a mayor *propensión modo 1* de producción de conocimiento de los investigadores, menor es su interés de reformar el sistema de evaluación en una dirección que valore más la investigación colectiva, aplicada e inter/multi/transdisciplinar, que cambie los criterios de evaluación, modifique la composición disciplinaria de las comisiones dictaminadores por área y que se base más en una evaluación cualitativa. Con un nivel menor de asociación y en contraposición, a mayor *propensión modo 2* en la orientación de la producción de conocimiento que realizan los investigadores, mayor también es su interés de reformar el sistema de evaluación en la dirección planteada;

- ✓ la *satisfacción con el proceso de evaluación* que tienen los investigadores no es independiente de la *flexibilidad epistemológica del sistema de evaluación* que perciben. De hecho la asociación observada entre estas dos variables latentes es la más fuerte y su correlación alcanza un valor de 0.754 estadísticamente significativo. A mayor *flexibilidad epistemológica del sistema de evaluación* percibida por lo investigadores, mayor es su *satisfacción con el proceso de evaluación* del desempeño de la investigación al que están sometidos;
- ✓ el interés de los investigadores en la *dirección de la reforma del sistema de evaluación* no es independiente de la *satisfacción con el proceso de evaluación* que tienen ni de la *flexibilidad epistemológica del sistema de evaluación* que perciben. En particular, la asociación entre esas variables latentes es significativa y negativa (-0.583 y -0.675 respectivamente). Es decir, cuanto mayor sea la *satisfacción con el proceso de evaluación* que tienen los investigadores y/o la *flexibilidad epistemológica del sistema de evaluación* que perciben, menor es su interés en la *dirección de la reforma del sistema de evaluación* propuesta.

Dado que todas las correlaciones estimadas entre los constructos son consistentes con lo esperado de acuerdo con la teoría (ver capítulo V), se concluye que:

- ✓ la evidencia es fuerte y sugiere validez nomológica del modelo de medida teórico propuesto para las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento.

De acuerdo con la evidencia derivada de las cuatro pruebas de validez realizadas, se concluye que:

- ✓ la teoría de medición propuesta, formalizada por modelo de medida multifactorial de cinco factores, para medir las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento tiene validez de constructo. Las variables observadas de medida verdaderamente representan los constructos latentes teóricos que buscan medir.

De acuerdo con los resultados obtenidos en las evaluaciones de la bondad de ajuste (sub sección X.3.3.1) y de la validez de constructo (sub sección X.3.3.2) del modelo de medida, se concluye que:

- ✓ la evidencia es fuerte y sugiere que el modelo de medida teórico propuesto para las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento es válido, confiable y preciso.

#### X.4 El Modelo de Ecuaciones Estructurales: evaluación de la investigación y producción de conocimiento

El objetivo de esta sección es poner a prueba y validar el modelo teórico estructural propuesto para explorar las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento. A diferencia del modelo de medida (AFC), aquí el foco se pone en las relaciones de dependencia entre las variables latentes. Es decir, en las relaciones teóricas/estructurales entre los constructos hipotetizadas. En otras palabras, se pone a prueba y valida la teoría que sustenta este trabajo.

Para ello, en la siguiente sub-sección (X.4.1) se especifica el modelo teórico estructural propuesto: i) se determina la unidad de análisis; ii) se formaliza el modelo matemáticamente, y se representa gráficamente con un diagrama de senderos; y iii) se analiza su identificabilidad, el resto de los aspectos del diseño del estudio son consistentes con los planteados en el capítulo VIII, sub-sección VIII.3.3.

Posteriormente, en la sub-sección X.4.2, se evalúa la validez del modelo teórico estructural que se está poniendo a prueba. Es decir, por una parte se analizan las medidas de bondad de ajuste para evaluar la aceptación o no del modelo teórico propuesto; y por otra parte, se estudian los parámetros estructurales estimados. Estos parámetros representan las relaciones estructurales de dependencia entre los constructos (variables latentes) que se hipotetizan en el modelo.

##### X.4.1 Modelo Teórico Estructural: representación gráfica y formalización matemática

En esta sub-sección, con una teoría de medición confiable –representada en el modelo de medida- y validada con el análisis AFC (ver sección X.3), se especifica el modelo teórico estructural que se propone. Es decir, se realizan un conjunto de transformaciones teóricas (y de notación) al modelo de medida de la sección X.3 para evaluar la validez de las relaciones conceptuales que se hipotetizan para explicar teóricamente las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento en el contexto del SNI mexicano.

El análisis se desarrolla a nivel de los investigadores categorizados en el SNI mexicano. Es decir, la unidad de análisis en este trabajo es la percepción de dichos investigadores sobre diversos aspectos de la evaluación que se realiza en el SNI. Es así que en el análisis que se realiza se ponen a prueba las relaciones estructurales de dependencia entre los conceptos, que representan las percepciones individuales de los investigadores categorizados en el sistema. Se trata del análisis de un modelo estructural de un solo nivel.

La teoría estructural que se propone encuentra su fundamento en la teoría, el debate y los hallazgos de investigaciones pasadas reportados en la literatura sobre las relaciones entre la evaluación del desempeño de la investigación y el contenido, dirección y características epistemológicas de la producción de conocimiento (ver capítulo V). En esta tesis esas relaciones se estudian a través de los cinco constructos definidos en el capítulo V, cuya operacionalización se presenta en la tabla 28 de la sub-sección X.3.1.

La teoría estructural que se propone hipotetiza que la *flexibilidad epistemológica del sistema de evaluación* (FLXEPSNI) percibida por los investigadores y la *satisfacción* -que tienen- *con el proceso de evaluación* (SATPROEV) del desempeño de la investigación que allí se implementa dependen simultáneamente, significativamente y en direcciones opuestas de sus modos diferenciales de

producción de conocimiento -*propensión modo 1* (PROP\_M1) y *propensión modo 2* (PROP\_M2)-. A la vez, la *satisfacción con el proceso de evaluación* que tienen los investigadores depende significativamente y positivamente de la *flexibilidad epistemológica del sistema de evaluación* que perciben; y, además y simultáneamente, su interés en la *dirección de la reforma del sistema de evaluación* (DIRREF) depende significativamente y negativamente de la *satisfacción con el proceso de evaluación* que tienen.

Específicamente, las hipótesis que se ponen a prueba con el modelo teórico estructural son:

- ✓ H1: la *propensión Modo 1* de producción de conocimiento (PROP\_M1) se relaciona positivamente con la *flexibilidad epistemológica del sistema de evaluación del SNI* (FLXEPSNI);
- ✓ H2: la *propensión Modo 1* de producción de conocimiento (PROP\_M1) se relaciona positivamente con la *satisfacción con el proceso de evaluación del SNI* (SATPROEV);
- ✓ H3: la *propensión Modo 2* de producción de conocimiento (PROP\_M2) se relaciona negativamente con la *flexibilidad epistemológica del sistema de evaluación del SNI* (FLXEPSNI);
- ✓ H4: la *propensión Modo 2* de producción de conocimiento (PROP\_M2) se relaciona negativamente con la *satisfacción con el proceso de evaluación del SNI* (SATPROEV);
- ✓ H5: la *flexibilidad epistemológica del sistema de evaluación del SNI* (FLXEPSNI) se relaciona positivamente con la *satisfacción con el proceso de evaluación del SNI* (SATPROEV); y,
- ✓ H6: la *satisfacción con el proceso de evaluación del SNI* (SATPROEV) se relaciona negativamente con la *dirección de la reforma del sistema de evaluación* (DIRREF).

Formalmente, el modelo teórico estructural puede escribirse como:

$$\eta_i = B\eta_i + \xi_i \quad (2)$$

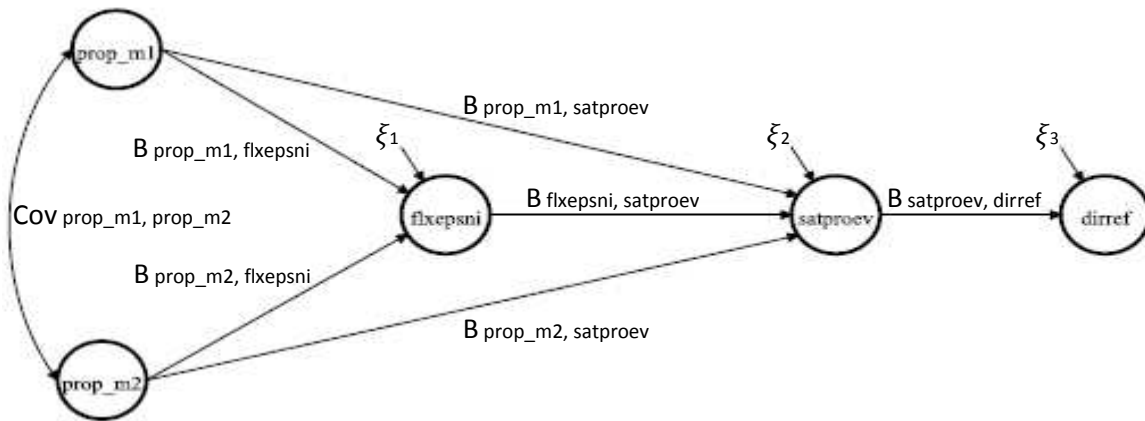
Donde:

- ✓  $B$  es una matriz  $m \times m$  con  $m = 5$ , de parámetros de pendiente estructurales para la regresión de las variables latentes sobre otras variables latentes.  $B$  tiene en la diagonal 0 y se asume que  $I - B$  es no singular.
- ✓  $\eta$  es un vector aleatorio  $m$  dimensional de las variables latentes (constructos),  $m = 5$ . Los elementos del vector  $\eta$  son: *propensión modo 1*, *propensión modo 2*, *satisfacción con el proceso de evaluación*, *flexibilidad epistemológica del sistema de evaluación* y *dirección de la reforma del sistema de evaluación*.
- ✓  $\xi$  es un vector  $m$  dimensional de varianzas residuales, con matriz de covarianzas  $\psi$ . En este caso  $m = 3$  porque el modelo contiene tres constructos endógenos.

Gráficamente, el modelo teórico estructural que se propone puede representarse a través del diagrama de senderos que se presenta en la figura 5. A efectos de no complejizar aún más el diagrama, las variables de medida asociadas a cada variable latente no se representan gráficamente.



Figura 5: Diagrama de senderos para el Modelo Teórico Estructural: evaluación y producción de conocimiento



Hipótesis	Parámetro estructural
H1: prop_m1 (+) $\rightarrow$ flxepsni	B prop_m1, flxepsni
H2: prop_m1 (+) $\rightarrow$ satproev	B prop_m1, satproev
H3: prop_m2 (-) $\rightarrow$ flxepsni	B prop_m2, flxepsni
H4: prop_m2 (-) $\rightarrow$ satproev	B prop_m2, satproev
H5: flxepsni (+) $\rightarrow$ satproev	B flxepsni, satproev
H6: satproev (-) $\rightarrow$ dirref	B satproev, dirref

Tal como se observa en la figura 5, y a diferencia de lo que se observaba en la representación gráfica del modelo de medida (figura 4 de la sub-sección X.3.2), las relaciones entre los constructos (variables latentes), que se representan con flechas (senderos) unidireccionales, indican las relaciones de dependencia que se hipotetizan y se corresponden con los parámetros libres a estimar en el modelo estructural; ya no son flechas bidireccionales ni indican correlaciones como en el caso del modelo de medida. Los constructos que no están unidos por ninguna flecha refiere a parámetros fijos (el parámetro se fija en 0) y significa que no hay una relación de dependencia directa entre esos conceptos en la teoría que se está poniendo a prueba. Adicionalmente, mientras que en el modelo de medida los cinco constructos son exógenos, en el modelo estructural algunos constructos son exógenos y otros son endógenos. De hecho, la teoría que se pone a prueba estudia el efecto de los constructos exógenos (predictores) sobre los endógenos (resultado). Simultáneamente, un constructo puede ser endógeno (resultado) en una relación de dependencia y en la siguiente puede ser exógeno (predictor). Los constructos endógenos no son explicados completamente, por lo que se les asocia un término de error o varianza residual ( $\xi$ ). Así el modelo estructural considera los errores de medida (varianzas residuales de los ítems) y también el error de predicción teórico ( $\xi$ ). Los parámetros estimados se calculan de acuerdo a los efectos directos e indirectos en el modelo y especificando relaciones de dependencia se establecen las restricciones sobre qué efectos se utilizan para la estimación.

En el contexto de esta investigación, los constructos *propensión modo 1* y *propensión modo 2* (prop\_m1 y prop\_m2 respectivamente) son exógenos en este modelo teórico. Son exógenos porque están determinados por factores que no son considerados en el modelo. Es decir, no se realiza ninguna hipótesis para explicarlos y/o predecirlos, y se utilizan únicamente como variables para explicar otros constructos, en particular *flexibilidad epistemológica del sistema de evaluación* (flxepsni) y *satisfacción con el proceso de evaluación* (satproev). *Propensión Modo 1* y *Propensión Modo 2* (prop\_m1 y prop\_m2 respectivamente) se ubican a la izquierda del diagrama de senderos (figura 5) y la flecha bidireccional que los une se especifica para capturar cualquier correlación existente entre ellos.

Por su parte, los constructos *flexibilidad epistemológica del sistema de evaluación* (flxepsni), *satisfacción con el proceso de evaluación* (satproev) y *dirección de la reforma del sistema de evaluación* (dirref) son endógenos. Cada uno de ellos está explicado por otros constructos incluidos en el modelo (ver hipótesis planteadas). Además *flexibilidad epistemológica del sistema de evaluación* (flxepsni) y *satisfacción con el proceso de evaluación* (satproev) son en algunas hipótesis resultado y en otras son predictores. Todas las hipótesis son puestas a prueba simultáneamente en el mismo modelo teórico estructural.

Finalmente y desde el punto de vista del diseño, queda por analizar la identificabilidad del modelo teórico estructural propuesto. Los grados de libertad del modelo teórico estructural se calculan de la misma manera que en el modelo de medida (ver sección X.3.2). Sin embargo, en el modelo estructural que se propone el número de parámetros libres a estimar es menor que en el modelo de medida, porque no se incluyen todas las relaciones estructurales entre los constructos. En particular, se fijan en cero tres relaciones estructurales, a saber: i) la relación entre prop\_m1 y dirref; ii) la relación entre prop\_m2 y dirref; y, iii) la relación entre flxepsni y dirref. Es así que en lugar de estimar 106 parámetros -como se hizo en el modelo de medida-, ahora se estiman 103 parámetros, a saber:

- ✓ 28 cargas factoriales ( $\lambda$ ) o relaciones de medida entre las variables de medición y las variables latentes (para fijar la escala, la carga factorial más grande por variable latente se deja constante e igual a 1);
- ✓ 63 umbrales ( $\tau$ ) -c - 1 categorías por cada variable categórica observada-; y,
- ✓ 12 relaciones entre las variables latentes o constructos (5 varianzas: una por cada variable latente, más una correlación entre las variables latentes prop\_m1 y prop\_m2, más 6 relaciones de dependencia estructurales entre los constructos).

El modelo teórico estructural que se propone está anidado dentro del modelo de medida (AFC) y es más parsimonioso porque contiene menor cantidad de parámetros libres a estimar con el mismo número de constructos (5) y de variables observadas de medida (33).

Además se trata de un modelo teórico estructural *recursivo*, porque todos los senderos entre constructos provienen solo del constructo predictor hacia otro constructo de resultado, no hay retroalimentación entre constructos (Hair et al., 2014). De acuerdo con estos autores, los modelos estructurales recursivos nunca pueden tener menor cantidad de grados de libertad que el AFC cuando se utilizan la misma cantidad de constructos y variables observadas de medida.

Con datos de corte transversal, como los que se tienen disponibles para este estudio, no es plausible desarrollar un modelo estructural no recursivo porque es difícil sino imposible re-crear el conjunto de condiciones necesarias para establecer relaciones recíprocas entre constructos, porque para ello es preciso considerar la secuencia temporal de los eventos (se requieren datos longitudinales) (Kline, 2011; Hair et al., 2014).

Dado que el modelo teórico estructural propuesto es un modelo recursivo, que no contiene términos de interacción, donde el tamaño de la muestra es lo suficientemente grande (N=8053), que cuenta con al menos tres variables observadas de medida por constructo, donde el modelo de medida está identificado y que se estima un número menor de parámetros que en el AFC, es posible concluir que el modelo teórico estructural propuesto también está identificado. De hecho, y al igual que el modelo de medida, está sobre-identificado.

Si en el modelo teórico estructural se especificaran la misma cantidad de relaciones estructurales que relaciones de correlación entre constructos del modelo de medida, se obtendría el *modelo estructural saturado*. De acuerdo con Hair et al. (2014) los modelos estructurales saturados generalmente no representan demasiado interés porque usualmente no revelan más información que la obtenida con el análisis AFC. Las medidas de bondad de ajuste del modelo estructural saturado deberían ser las mismas que las obtenidas en el modelo de medida (AFC). En la siguiente sub-sección, entre otras cosas se presentan y comparan las medidas de bondad de ajuste del modelo teórico estructural saturado contra las obtenidas con el AFC para corroborar que no se cometió ningún error en la transformación desde el modelo de medida hacia el estructural.

A continuación se evalúa la validez del modelo teórico estructural propuesto.

#### X.4.2 Validez del Modelo Teórico Estructural: La relación entre la evaluación y la producción de conocimiento

En esta sub-sección se pone a prueba y se realiza la validación del modelo estructural para explicar teóricamente la relación entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento. Se trata de poner a prueba las seis hipótesis planteadas en la sub-sección anterior, es decir, validar la teoría estructural que se representa con las relaciones de dependencia estructurales entre los constructos especificadas en el modelo.

Para ello, y dado que a esta altura del análisis ya se logró una teoría de medición confiable y válida -representada en el modelo de medida (AFC) de la sección X.3-, a continuación y en primer lugar, se analiza la bondad de ajuste del modelo estructural (X.4.2.1) para evaluar el grado de aceptación del modelo teórico propuesto. En segundo lugar, en la sub-sección X.4.2.2, se estudian los parámetros estructurales estimados (los senderos o relaciones hipotetizadas). Es decir, se analiza su significancia estadística y dirección, como también su significado práctico o implicancia conceptual. La salida completa de la estimación del modelo de ecuaciones estructurales se presenta en el Anexo 4.

Si el modelo estructural presenta buen ajuste y los parámetros estructurales son significativos y en las direcciones hipotetizadas, entonces se puede concluir que el modelo teórico estructural es sostenido por la evidencia. Estos resultados junto con la validez del modelo de medida representan evidencia fuerte para validar la teoría estructural que se propone para explicar las relaciones entre la evaluación del desempeño de investigación y los modos diferenciales de producción de conocimiento en el contexto del SNI mexicano. A continuación se desarrolla este análisis.

*X.4.2.1 Evaluación de la Bondad de Ajuste del Modelo Teórico Estructural: evaluación y producción de conocimiento*

El modelo estructural –recursivo- tiene mayor cantidad de grados de libertad que el modelo de medida –porque se estiman menor cantidad de relaciones entre los constructos- y dado que todo lo demás se deja constante (el tamaño de la muestra, la cantidad de constructos y de ítems), el valor del estadístico Ji-Cuadrado nunca puede ser menor que el valor obtenido en el modelo de medida. Es decir, el modelo estructural no va a presentar un ajuste mejor que el modelo de medida.

En este sentido, la estrategia para evaluar la bondad de ajuste del modelo estructural implica comparar las medidas de bondad de ajuste del modelo estructural contra las obtenidas en el modelo de medida para analizar en qué medida el ajuste se degrada debido a la especificación de las relaciones estructurales. Así, las medidas de bondad de ajuste que se analizan aquí son las mismas que se analizaron en el modelo de medida (Ji-Cuadrado, RMSEA, CFI, TLI) con los mismos valores de referencia de corte. Si el ajuste del modelo estructural es sustantivamente peor que el del modelo de medida, entonces se puede concluir que la teoría estructural carece de validez. (Hair et al., 2014).

Antes de presentar los resultados de esa comparación, vale señalar que Hair et al. (2014) proponen una medida adicional para evaluar la bondad de ajuste del modelo estructural respecto al ajuste del modelo de medida, a saber: la diferencia del estadístico Ji-Cuadrado entre esos dos modelos. El criterio indica que si esa diferencia es significativa, entonces el ajuste empeora significativamente. Sin embargo, Kline (2011:219) señala que con modelos no jerárquicos, de un solo nivel como el que se propone en este estudio, esa comparación se puede realizar pero la prueba estadística de la diferencia de la Ji-Cuadrado no puede interpretarse estadísticamente. Además de esto, y dado el efecto del tamaño de la muestra de este estudio sobre el valor de la Ji-Cuadrado, es probable que dicha diferencia sea significativa, cuando en realidad el ajuste es aceptable. A continuación, en la tabla 37, se presentan los resultados.

*Tabla 37: Medidas de Bondad de Ajuste del Modelo Estructural comparadas con el Modelo de Medida*

	<b>Modelo Estructural</b>	<b>Modelo de Medida</b>	<b>Modelo Estructural Saturado</b>
<b>Medidas</b>	<b>Valores estimados</b>		
Parámetros libres estimados	103	106	106
N	8053	8053	8053
<b>Medidas de Ajuste Absoluto</b>			
Prueba Ji-Cuadrado			
$X^2$	11855.593*	10890.792*	10890.791*
$df$	488	485	485
$p - value$	0.0000 <sup>1</sup>	0.0000 <sup>1</sup>	0.0000 <sup>1</sup>
RMSEA	0.054	0.052	0.052
90% IC para el valor estimado del RMSEA	(0.053; 0.055)	(0.051; 0.052)	(0.051; 0.052)
Probabilidad RMSEA $\leq$ 0.05	0.000	0.001	0.001
<b>Medidas de Ajuste Incremental</b>			
CFI	0.951	0.955	0.955
TLI	0.947	0.951	0.951

Ji-Cuadrado del Modelo Nulo			
$X^2$	230934.785	230934.785	230934.785
$df$	528	528	528
$p - value$	0.0000 <sup>1</sup>	0.0000 <sup>1</sup>	0.0000 <sup>1</sup>
Prueba de la diferencia de la Ji-Cuadrado: Modelo Estructural (H <sub>0</sub> ) vs Modelo de Medida (H <sub>1</sub> )			
<i>Diferencia de la X<sup>2</sup></i>	398.058	---	
$df$	3	---	
$p - value$	0.0000 <sup>1</sup>	---	

Notas: 1. Estadísticamente significativa al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.  
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Tal como puede observarse en la tabla 37 y como era esperable, el valor del estadístico Ji-Cuadrado de bondad de ajuste del modelo teórico estructural es mayor que el del modelo de medida, y por lo tanto, igualmente significativo. Tampoco sorprende que la prueba estadística de la diferencia de la Ji-Cuadrado -derivada de la comparación entre el modelo estructural y el modelo de medida-, sea significativa. Ambos resultados sugieren un ajuste pobre del modelo estructural, y el último resultado, implica además –si no se tiene en cuenta la advertencia de Kline (2011)- que las relaciones estructurales especificadas degradan significativamente el ajuste del modelo. Sin embargo, tal como se planteó, este resultado no es concluyente debido a la sensibilidad de la Ji-Cuadrado al tamaño de la muestra, la complejidad del modelo propuesto y que además se trata de un modelo estructural no jerárquico.

Interesa entonces analizar las otras medidas de bondad de ajuste, tales como: RMSEA, CFI y TLI. Lo primero que hay que señalar es que los valores estimados de esas medidas de bondad de ajuste para el modelo estructural y para el modelo de medida son similares (ver tabla 37). Como era esperable y de acuerdo con esas medidas de bondad de ajuste, el ajuste del modelo estructural empeoró, pero relativamente poco en relación al ajuste del modelo de medida. Lo segundo que hay que señalar es que los valores estimados del RMSEA, CFI y TLI del modelo estructural, aunque un poco peores que los estimados para el modelo de medida, siguen estando dentro de los márgenes establecidos en la literatura que sugieren un ajuste satisfactorio del modelo estructural. En síntesis:

- ✓ hasta este punto del análisis no se ha encontrado evidencia que sugiera que el modelo teórico estructural propuesto carece de validez.

Adicionalmente, tal como se muestra en la tabla 37, las medidas de bondad de ajuste estimadas para el modelo estructural saturado son exactamente iguales a las estimadas para el modelo de medida. Este resultado es útil porque permite corroborar que no se ha cometido ningún error en la transformación del modelo de medida al estructural. La salida completa de la estimación del modelo de ecuaciones estructurales saturado se presenta en el Anexo 5.

A continuación se estudia la naturaleza y significado de los parámetros estructurales estimados.

#### X.4.2.2 Las relaciones de dependencia hipotetizadas: naturaleza y significado de los parámetros estructurales

La validez de la teoría estructural que se propone depende de los resultados obtenidos en el análisis de: i) la validez y confiabilidad del modelo de medida; ii) la bondad de ajuste del modelo estructural; y iii) la naturaleza y significado de los parámetros estructurales estimados.

Hasta este punto del proceso de análisis no se ha encontrado evidencia que sugiera que el modelo teórico estructural que se propone carezca de validez, ya que se ha construido una teoría de medición válida y confiable, y se ha constatado un ajuste razonable y satisfactorio del modelo teórico estructural a los datos. Sin embargo, esos resultados son insuficientes para sostener la teoría estructural que se propone; se requiere además el estudio de los parámetros estructurales estimados. Es preciso analizar la aceptación o rechazo de cada una de las hipótesis -sobre las relaciones de dependencia entre los constructos- a la luz de los parámetros estructurales estimados. Si los parámetros estructurales estimados son estadísticamente significativos y en las direcciones hipotetizadas (relaciones positivas o negativas), entonces existe evidencia que apoya el modelo teórico propuesto (la teoría estructural) y además aumenta su validez. Aquí también y a efectos de analizar la capacidad explicativa del modelo teórico estructural que se propone, se examina la varianza explicada de los constructos endógenos ( $R^2$ ).

A continuación y antes de analizar los parámetros estructurales, en la tabla 38 se presenta la comparación entre los parámetros de medida estimados en la estimación del modelo teórico estructural propuesto y las cargas factoriales obtenidas en la estimación del modelo de medida. El objetivo es analizar si existe evidencia que sugiera *confusión interpretativa* (“*interpretational confounding*”) (Kline, 2011; Hair et al., 2014). Si hay evidencia que sugiera confusión interpretativa, significa que las estimaciones de las relaciones de medida en el modelo estructural para los distintos constructos son afectadas sustantivamente más por las relaciones estructurales especificadas en el modelo estructural que por las variables de medida específicas. Esto representa un problema y atenta contra la estabilidad y confiabilidad de la medición (genera confusión o dudas respecto al significado de los constructos), básicamente porque las cargas factoriales de los constructos – estimadas en el modelo de medida- y los parámetros de medición –estimados en el modelo estructural- deberían ser similares y no variar demasiado por los cambios que se realicen en la especificación del modelo estructural. De acuerdo con Hair et al. (2014) variaciones de 0.05 en valor absoluto o menores en la comparación de esos parámetros, son aceptables y no representan un problema de confusión interpretativa.

Como se observa en la tabla 38, las estimaciones de los parámetros de medida no cambiaron demasiado respecto a las cargas factoriales estimadas en el modelo de medida. De hecho, la variación mayor alcanza –en valor absoluto- un 0.023. Este resultado provee evidencia que sugiere que no hay problemas de confusión interpretativa y además representa nueva evidencia sobre la estabilidad de las variables observadas de medida, reforzando la validez de la teoría de medición y su modelo respectivo. Adicionalmente, y como es esperable dada la estabilidad en las estimaciones de los parámetros de medida, la confiabilidad compuesta (CR) permanece casi constante.

Tabla 38: Comparación entre los parámetros de medida del Modelo Estructural y las cargas factoriales del Modelo de Medida

		Modelo Estructural	Modelo de Medida	Variación
Variable Latente	Ítems	Parámetros de medida estandarizados <sup>1</sup>	Cargas Factoriales estandarizadas <sup>1</sup>	
PROP_M1	P_18_C_R	0.587	0.584	0.003
	P_18_F_R	0.600	0.599	0.001
	P_18_G_R	0.555	0.556	-0.001
	P_18_I_R	0.705	0.710	-0.005
	P_18_J_R	0.628	0.626	0.002
PROP_M2	P_18_B_R	0.772	0.771	0.001
	P_18_D_R	0.823	0.823	0.000
	P_18_E_R	0.586	0.589	-0.003
	P_18_H_R	0.651	0.654	-0.003
	P_18_K_R	0.682	0.680	0.002
SATPROEV	P_1_A_R	0.860	0.872	-0.012
	P_1_B_R	0.962	0.967	-0.005
	P_2_A_R	0.447	0.468	-0.021
	P_2_C_R	0.970	0.972	-0.002
	P_5_R_D	0.874	0.889	-0.015
	P_7_R	0.803	0.821	-0.018
	P_8_R	0.953	0.968	-0.015
	P_9_A_R	0.360	0.376	-0.016
	P_9_B_R	0.317	0.329	-0.012
P_9_D_R	0.644	0.667	-0.023	
FLXEPSNI	P_3_R_D	0.542	0.546	-0.004
	P_10_R	0.927	0.928	-0.001
	P_11_R_D	0.852	0.849	0.003
	P_9_C_R	0.612	0.609	0.003
	P_17_A_R	0.643	0.645	-0.002
	P_17_B_R	0.621	0.623	-0.002
DIRREF	P_17_D_R	0.547	0.547	0.000
	P_20_G_R	0.820	0.822	-0.002
	P_20_K_R	0.975	0.976	-0.001
	P_20_L_R	0.617	0.630	-0.013
	P_20_I_R	0.991	0.989	0.002
	P_20_J_R	0.912	0.911	0.001
	P_20_E_R	0.833	0.831	0.002
		<b>CR: Confiabilidad Compuesta</b>		
PROP_M1		0.7533	0.7532	----
PROP_M2		0.8320	0.8323	----
SATPROEV		0.9243	0.9300	----
FLXEPSNI		0.8605	0.8608	----
DIRREF		0.9469	0.9474	----

Notas: 1. Estadísticamente significativos al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.  
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

En la tabla 39 se presentan los parámetros estructurales estandarizados y sin estandarizar estimados. Como se observa allí, todos los parámetros estructurales son estadísticamente significativos y la amplia mayoría con la dirección esperada. Es decir, la evidencia empírica confirma todas las hipótesis planteadas, excepto la direccionalidad de la hipótesis H3.

Tabla 39: Parámetros estructurales del Modelo Teórico Estructural: Evaluación y producción de conocimiento

Relación Estructural	Parámetro sin estandarizar	S.E	Est./S.E.	Two-Tailed P-Value	Parámetro estandarizado
PROP_M1 ---> FLXEPSNI	0.451	0.023	19.575	0.000	0.343
PROP_M2 ---> FLXEPSNI	0.121	0.019	6.235	0.000	0.108
PROP_M1 ---> SATPROEV	0.096	0.021	4.571	0.000	0.07
PROP_M2 ---> SATPROEV	-0.194	0.017	-11.582	0.000	-0.165
FLXEPSNI ---> SATPROEV	0.884	0.014	61.563	0.000	0.845
SATPROEV ---> DIRREF	-0.675	0.011	-61.580	0.000	-0.661
PROP_M1 correlacionado PROP_M2	0.223	0.008	28.738	0.000	0.385

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014). Estimaciones realizadas con M-PLUS Versión 7.4.

El único parámetro estructural que, si bien es significativo, va en la dirección contraria a la hipotetizada (H3), es el que representa la relación estructural de dependencia entre propensión modo 2 de producción de conocimiento (PROP\_M2) y flexibilidad epistemológica del sistema de evaluación (FLXEPSNI).

De acuerdo con la teoría, se esperaba una relación negativa entre propensión modo 2 y la flexibilidad epistemológica del sistema de evaluación, es decir, a mayor puntuación propensión modo 2, menor la flexibilidad epistemológica del sistema de evaluación que perciben los investigadores. Esta relación no se corrobora para el caso mexicano, quizás por la propia institucionalización de la evaluación en los sistemas de evaluación de las propias instituciones, que refuerza los efectos del sistema de evaluación reduciendo la diversidad cognitiva (Whitley, 2007). Esto se debe a las estrategias adaptativas de los investigadores (Bensusán et al., 2014; Gläser y Laudel, 2007b) para mantenerse y/o ascender en el sistema.

Sin embargo, resulta interesante y consistente con la racionalidad conceptual de la hipótesis (H3) planteada, que dicho parámetro estructural estandarizado (0.108) es tres veces menor que el observado para la relación estructural entre propensión modo 1 y flexibilidad epistemológica del sistema de evaluación (H1), cuyo parámetro estandarizado es 0.343. Este resultado es, por lo tanto, consistente con la lógica o racionalidad conceptual subyacente al modelo. Se puede concluir que:

- ✓ La magnitud de la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI depende significativamente de sus modos diferenciales de producción de conocimiento (*propensión modo 1 y propensión modo 2*).
- ✓ Los investigadores con mayor *propensión modo 1* perciben una flexibilidad epistemológica del sistema de evaluación tres veces mayor que sus pares con mayor propensión modo 2.

Los resultados empíricos obtenidos en las estimaciones de las relaciones estructurales entre los modos diferenciales de producción de conocimiento y satisfacción con el proceso de evaluación (PROP\_M1 y SATPROEV y PROP\_M2 y SATPROEV) confirman las hipótesis H2 y H4. Mientras que los investigadores con puntuaciones mayores en propensión modo 1 están significativamente satisfechos con el proceso de evaluación, sucede lo contrario con los investigadores con mayores



puntuaciones en propensión modo 2, cuya satisfacción disminuye significativamente. Se puede concluir que:

- ✓ La *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI depende significativamente y en direcciones opuestas de sus modos diferenciales de producción de conocimiento (*propensión modo 1* y *propensión modo 2*).
- ✓ Conforme aumenta la *propensión modo 1* de producción de conocimiento que orienta el trabajo de los investigadores categorizados en el sistema, también aumenta su *satisfacción con el proceso de evaluación* del desempeño de la investigación que allí se realiza. Por el contrario, conforme aumenta la *propensión modo 2* de producción de conocimiento que orienta el trabajo de los investigadores categorizados en el sistema, su *satisfacción con el proceso de evaluación* del desempeño de la investigación que allí se realiza disminuye significativamente.

Antes de continuar con los resultados para los restantes parámetros, es preciso señalar que la magnitud del efecto directo de *propensión modo 1* sobre *satisfacción con el proceso de evaluación* es bastante pequeña (0.07), lo que sugiere una alta probabilidad de estar cometiendo un “Error Tipo I” y con ello, se cuestionaría la validez de la relación teórica estructural de dependencia entre ambas variables. Para minimizar la probabilidad de estar cometiendo un Error Tipo I, es decir, para minimizar la probabilidad de que relaciones que en realidad son insignificantes aparezcan como estadísticamente significativas, en la sección X.5 se especifica, estima y presentan los resultados para el modelo teórico estructural condicional. Esto permite aumentar la eficiencia del diseño, reducir ciertos sesgos, y con ello, alcanzar una evaluación más precisa y confiable de todas las relaciones estructurales (hipótesis) que se están poniendo a prueba en esta investigación.

Hecha la salvedad y continuando con los resultados, la *satisfacción con el proceso de evaluación* que tienen los investigadores además de depender significativamente de sus modos diferenciales de producción de conocimiento, depende también positiva y significativamente de la *flexibilidad epistemológica del sistema de evaluación* que perciben. Este resultado confirma la H5 planteada. Un aumento de una unidad estándar en la puntuación de la *flexibilidad epistemológica del sistema de evaluación* percibida tiene un efecto directo de 0.845 unidades estándar de aumento en la *satisfacción con el proceso de evaluación* que tienen los investigadores. De esta forma, es la variable latente o dimensión conceptual más importante para explicar la satisfacción que tienen los investigadores categorizados en el SNI con el proceso de evaluación del desempeño de la investigación que allí se implementa. Se puede concluir que:

- ✓ La *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI depende significativamente y positivamente de su percepción sobre la *flexibilidad epistemológica del sistema de evaluación*.
- ✓ Conforme aumenta la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI, también se incrementa significativamente su *satisfacción con el proceso de evaluación* del desempeño de la investigación que allí se realiza.

Este resultado pone de relieve la importancia de pensar cuidadosamente, cuando se diseñan los instrumentos de evaluación del desempeño de la investigación, en su flexibilidad desde el punto de vista de su relación con las características epistemológicas de los procesos de producción de

conocimiento. Es decir, de desarrollar instrumentos de evaluación que cuiden –limiten- sus efectos constitutivos sobre las características epistemológicas de la producción de conocimiento; cuidando la calidad, originalidad, relevancia conceptual y confiabilidad de los resultados de investigación, su diversidad cognitiva (orientación y tipos de investigación), los procesos de producción de conocimiento colectivos (ya sea entre investigadores o entre investigadores y otros actores externos al ámbito de la academia) y la investigación de largo plazo.

Simultáneamente, los resultados empíricos obtenidos en la estimación de la relación estructural entre *satisfacción con el proceso de evaluación* (SATPROEV) y *dirección de la reforma del sistema de evaluación* (DIRREF) confirman la H6. El parámetro estructural estimado que relaciona la satisfacción con el proceso de evaluación que tienen los investigadores y su interés en la dirección de la reforma del sistema de evaluación es estadísticamente significativo, negativo y su valor estandarizado es de -0.661. Se puede concluir que:

- ✓ El interés de los investigadores categorizados en el SNI en la *dirección de la reforma del sistema de evaluación* planteada depende significativamente y negativamente de la *satisfacción con el proceso de evaluación* del desempeño de la investigación que dichos investigadores tienen.
- ✓ Conforme aumenta la *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI, porque el mismo es capaz de reconocer su desempeño, ya que cuenta con la información necesaria e incluye todas sus contribuciones importantes, los criterios y mecanismos de selección de los evaluadores respetan los equilibrios regionales e institucionales y dan garantías de una evaluación objetiva y la elaboración de juicios académicos apropiados en su área de conocimiento; disminuye significativamente su interés en la *dirección de la reforma del sistema de evaluación*, porque cambia los criterios actuales de evaluación dándole mayor peso a la evaluación cualitativa –*versus* una evaluación cuantitativa y centrada en indicadores bibliométricos-, modifica la composición disciplinaria de las comisiones dictaminadoras, y valora y pondera mejor la investigación colectiva, no disciplinar y orientada por la investigación aplicada o por la búsqueda de soluciones a problemas dependientes del contexto -ya sean sociales, productivos y/o de política pública-.

La evidencia empírica encontrada apoya el modelo teórico estructural propuesto y aumentan su validez. Los resultados obtenidos sugieren que el sistema de evaluación del SNI mexicano no es neutro, sino que responde a un cierto modelo de evaluación, cuyas medidas para ponderar la productividad, calidad e impacto de la producción de conocimiento están sesgadas a favor de un tipo de producción. En particular, los resultados sugieren que el sistema de evaluación del SNI al tiempo que es capaz reconocer y valorar mejor el desempeño de los investigadores que predominantemente orientan su producción de conocimiento a través del Modo 1 y facilitar más las características epistemológicas de dicha producción, desconoce o tiene dificultades para valorar y reconocer el desempeño de la investigación realizada por los investigadores que predominantemente trabajan en el Modo 2 de producción de conocimiento y permitir las características epistemológicas asociadas a esa producción. De aquí que éstos últimos perciban una menor flexibilidad epistemológica del sistema de evaluación y tengan una menor satisfacción –negativa- con el proceso de evaluación, que los que trabajan predominantemente en el Modo 1 de producción de conocimiento. Consistentemente, el interés de los investigadores en reformar el

sistema de evaluación en la dirección planteada (más orientada a reconocer y valorar también la producción de conocimiento modo 2) depende negativamente de la satisfacción que tienen los investigadores con el proceso de evaluación, satisfacción que además depende positivamente de la flexibilidad epistemológica del sistema de evaluación que perciben.

Así, los resultados obtenidos sugieren que el SNI, dados los criterios de evaluación que utiliza, introduce sesgos que pueden conducir a efectos constitutivos sobre la producción de conocimiento modo 2. Es decir, en la medida que la evaluación no reconoce adecuadamente las especificidades de este modo de producción de conocimiento, por omisión lo castiga y los investigadores perciben menos incentivos a trabajar en este modo de producción. Así se reduce el impacto potencial de la producción de conocimiento en términos de su contribución en la búsqueda de soluciones a los principales problemas que afectan al desarrollo social y económico de México.

Finalmente, y a efectos de observar la capacidad explicativa del modelo teórico propuesto, en la tabla 40 se presentan las varianzas explicadas ( $R^2$ ) de los constructos endógenos.

*Tabla 40: Varianza Explicada ( $R^2$ ) de los constructos endógenos*

Latent Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SATPROEV	0.715	0.013	53.688	0.000
FLXEPSNI	0.158	0.011	14.555	0.000
DIRREF	0.437	0.013	33.728	0.000

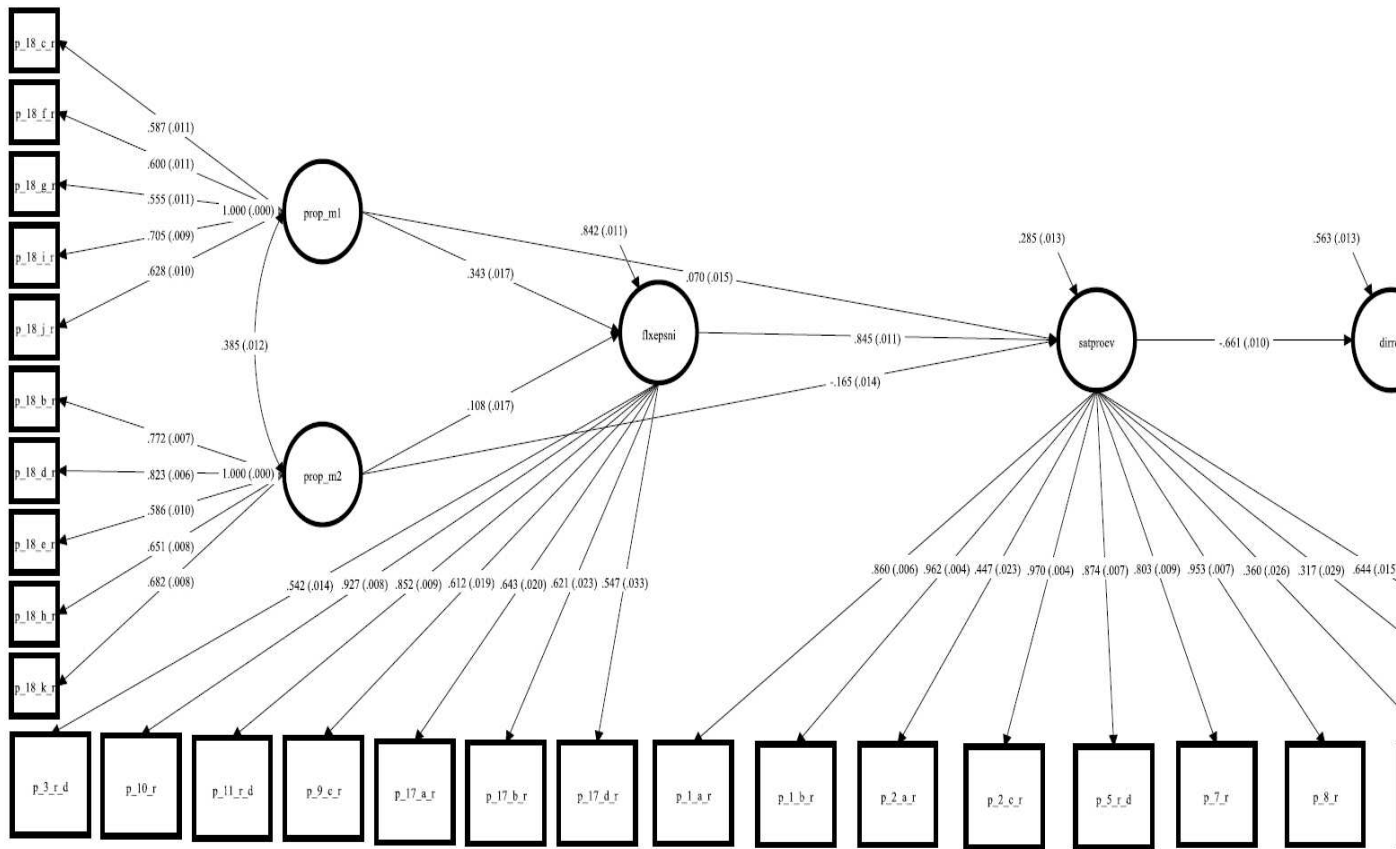
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014). Estimaciones realizadas con M-PLUS Versión 7.4.

Tal como se observa en la tabla 40, las varianzas explicadas de los constructos endógenos son todas significativas. El modelo teórico propuesto explica el 71.5% de la varianza de satisfacción con el proceso de evaluación, el 43.7% de la varianza de la *dirección de la reforma del sistema de evaluación* y un 15.8% de la varianza de *flexibilidad epistemológica del sistema de evaluación*. Este último resultado sugiere que además de los modos diferenciales de producción de conocimiento, existen otras dimensiones conceptuales no consideradas en el modelo importantes para explicar la variación en la percepción que tienen los investigadores sobre la *flexibilidad epistemológica del sistema de evaluación*.

En la figura 6 se presenta el diagrama de senderos del modelo teórico estructural completo con todos los parámetros estimados estandarizados y entre paréntesis sus errores estándar.

En la siguiente sección (X.5) se presentan y analizan los resultados obtenidos de la estimación del modelo teórico estructural condicional. La salida completa de la estimación del modelo de ecuaciones estructurales condicional se presenta en el Anexo 6.

Figura 6: Modelo teórico estructural: evaluación de la investigación y producción de conocimiento



Notas: Parámetros estimados estandarizados, errores estándar entre paréntesis. Estimaciones realizadas con M-PLUS Ver 6.23.  
 Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

## X.5 El modelo teórico estructural condicional

De acuerdo con lo planteado en el capítulo VIII sobre: i) el control de las variables perturbadoras; y ii) el alcance, naturaleza y tratamiento de los datos faltantes; en esta sección se estima el modelo teórico estructural condicional y se analizan sus resultados.

Aquí, la intención es doble. Por un lado se trata de controlar los resultados obtenidos en la subsección anterior, aislando el efecto potencial que podrían tener -sobre las variables explicadas- ciertas variables perturbadoras. Por otro lado, se busca aplicar un tratamiento a los datos faltantes que contribuya a la obtención de estimaciones de los parámetros eficientes e insesgadas. Así se espera mejorar la eficiencia y minimizar la probabilidad de cometer errores tipo 1.

A diferencia del modelo teórico estructural desarrollado en la sección X.4, aquí y siguiendo las recomendaciones de Enders (2010) para asegurar estimaciones de los parámetros eficientes e insesgadas, se aplica la estimación por máxima verosimilitud con información completa (FILM por sus siglas en inglés) (Asparouhov y Muthén, 2010) para el tratamiento de los datos faltantes.

Es así que se incluyen en el modelo un conjunto de covariables  $X$  de control útiles para: i) aumentar la eficiencia reduciendo los errores aleatorios y disminuyendo el sesgo de las variables perturbadoras, cuyos efectos pueden confundirse -si no se controlan- con los efectos de las variables explicativas; y, ii) predecir la probabilidad de datos faltantes, haciendo plausible que los datos faltantes sigan un mecanismo MARX (Asparouhov y Muthén, 2010). Las covariables  $X$  de control, que condicionan el modelo, identificadas son: *edad, género, nivel SNI, área de conocimiento y tipo de institución*. (Ver Anexo 7).

El objetivo en última instancia es aumentar la validez de la teoría que se contrasta y representa a través del modelo teórico estructural condicional.

A continuación, en la sección X.5.1 se especifica formalmente el modelo condicional y se analizan los aspectos vinculados con su identificabilidad. En la sección X.5.2 se examina su bondad de ajuste, en la sección X.5.3 se analizan y comparan los parámetros estructurales del modelo teórico estructural condicional en relación a lo no condicional, y se concluye sobre la validez de la teoría que se pone a prueba.

### X.5.1 Especificación matemática

Formalmente, el modelo teórico estructural condicional puede escribirse como:

$$\eta_i = B\eta_i + \Gamma x_i + \xi_i \quad (1);$$

y, su modelo de medida asociado queda especificado formalmente como:

$$y_i^* = \Lambda \eta_i + Kx_i + \epsilon_i \quad (2)$$

Donde:

- ✓  $y_i^*$  es un vector  $p$  dimensional de variables de respuesta latente, con  $p = 33$ . Es decir, se supone que para cada variable de medida observada  $y_j$  (con  $j = 1, 2, 3, \dots, 33$ ) existe una variable de respuesta latente  $y_j^*$ , tal que:

$$y_{ij} = \begin{cases} 1 & \text{si } y_{ij}^* \leq \tau_1 \\ 2 & \text{si } \tau_1 < y_{ij}^* \leq \tau_2 \\ 3 & \text{si } \tau_2 < y_{ij}^* \leq \tau_3 \\ 4 & \text{si } \tau_3 < y_{ij}^* \leq \tau_4 \\ 5 & \text{si } y_{ij}^* > \tau_4 \end{cases} \quad \text{Con } j = 1, 2, 3, \dots, 10 \text{ e } i = 1, 2, 3, \dots, 8053 \text{ observaciones.}$$

$$y_{ij} = \begin{cases} 0 & \text{si } y_{ij}^* \leq \tau_1 \\ 1 & \text{si } y_{ij}^* > \tau_1 \end{cases} \quad \text{Con } j = 11, 12, 13, \dots, 33 \text{ e } i = 1, 2, 3, \dots, 8053 \text{ observaciones.}$$

- ✓  $\eta$  es un vector aleatorio  $m$  dimensional de las variables latentes (constructos o factores), con media cero,  $m = 5$ . Los elementos del vector  $\eta$  son: *propensión modo 1, propensión modo 2, satisfacción con el proceso de evaluación, flexibilidad epistemológica del sistema de evaluación y dirección de la Reforma del sistema de evaluación.*
- ✓  $\epsilon$  es un vector aleatorio  $p$  dimensional de las varianzas residuales de  $y^*$  o errores de medición, con media cero y matriz de covarianza  $\Theta$ , con  $p = 33$ . Los errores de medida se asumen independientes entre sí y respecto a  $\eta$ .
- ✓  $\Lambda$  es una matriz de  $p \times m$  de los parámetros de pendiente de medida o cargas factoriales, con  $p = 33$  y  $m = 5$ .
- ✓  $\Theta$  es la matriz de covarianza de  $\epsilon$  de  $33 \times 33$ .
- ✓  $x$  es un vector  $q$  dimensional de variables independientes, con  $q = 19$ . En la tabla 41 se describen las covariables  $x$ .
- ✓  $K$  es una matriz  $p \times q$  de parámetros de pendiente de regresión. Cuando dichos parámetros son distintos de cero, indica que la variable de respuesta latente  $y^*$  está influenciada directamente por una o más variables  $x$ .
- ✓  $B$  es una matriz  $m \times m$  con  $m = 5$ , de parámetros de pendiente estructurales para la regresión de las variables latentes sobre otras variables latentes.  $B$  tiene en la diagonal 0 y se asume que  $I - B$  es no singular.
- ✓  $\Gamma$  es una matriz  $m \times q$  de parámetros de pendiente para las regresiones de las variables latentes (*propensión modo 1 y propensión modo 2*) sobre las variables independientes  $x$ .
- ✓  $\xi$  es un vector  $m$  dimensional de varianzas residuales, con matriz de covarianzas  $\psi$ . En este caso  $m = 5$  porque el modelo condicional, ahora contiene 5 constructos endógenos.

En la tabla 41 se describen las covariables  $x_i$  (variables independientes)

*Tabla 41: Variables independientes: Características de los investigadores categorizados en el SNI*

Variabes independientes	$x_i$	Descripción	Escala
Edad	$x_1$	Edad de los investigadores	Años
Cand	$x_2$	Categorizado en Candidato	0 – 1 1 = Candidato; 0 = En otro caso
niv_I	$x_3$	Categorizado en Nivel I	0 – 1 1 = Nivel I; 0 = En otro caso
niv_II	$x_4$	Categorizado en Nivel II	0 – 1 1 = Nivel II; 0 = En otro caso
f_m_ct	$x_5$	Trabaja en el área de Física, Matemáticas y Ciencias de la Tierra	0 – 1 1 = Física Matemáticas y Ciencias de la Tierra; 0 = En otro caso
b_q	$x_6$	Trabaja en el área de Biología y Química	0 – 1

			1 = Biología y Química; 0 = En otro caso
m_cs	$x_7$	Trabaja en el área de Medicina y Ciencias de la Salud	0 – 1 1 = Medicina y Ciencias de la Salud; 0 = En otro caso
Cs	$x_8$	Trabaja en el área de Ciencias Sociales	0 – 1 1 = Ciencias Sociales; 0 = En otro caso
b_ca	$x_9$	Trabaja en el área de Biotecnología y Ciencias Agropecuarias	0 – 1 1 = Biotecnología y Ciencias Agropecuarias; 0 = En otro caso
Ing	$x_{10}$	Trabaja en el área de Ingenierías	0 – 1 1 = Ingenierías; 0 = En otro caso
Mujer	$x_{11}$	Es Mujer	0 – 1 1 = Mujer; 0 = En otro caso
cpi_e_n	$x_{12}$	Desempeña actividades en Centro Público de Investigación de Ciencias Exactas y Naturales	0 – 1 1 = CPI Exactas y Naturales; 0 = En otro caso
cpi_cs_h	$x_{13}$	Desempeña actividades en Centro Público de Investigación de Ciencias Sociales y Humanidades	0 – 1 1 = CPI Ciencias Sociales y Humanidades; 0 = En otro caso
cpi_dt_s	$x_{14}$	Desempeña actividades en Centro Público de Investigación de Desarrollo Tecnológico y Servicios	0 – 1 1 = CPI Desarrollo Tecnológico y Servicios; 0 = En otro caso
cpi_sec	$x_{15}$	Desempeña actividades en Centro Público de Investigación Sectorizado	0 – 1 1 = CPI Sectorizado; 0 = En otro caso
Gobierno	$x_{16}$	Desempeña actividades en Gobierno	0 – 1 1 = Gobierno; 0 = En otro caso
inst_n_s	$x_{17}$	Desempeña actividades en Institutos Nacionales de Salud	0 – 1 1 = Institutos Nacionales de Salud; 0 = En otro caso
inst_ut	$x_{18}$	Desempeña actividades en Institutos y Universidades Tecnológicas	0 – 1 1 = Institutos y Universidades Tecnológicas; 0 = En otro caso
Univ	$x_{19}$	Desempeña actividades en Universidades	0 – 1 1 = Universidades; 0 = En otro caso

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

En el modelo teórico estructural condicionado se estiman 140 parámetros libres, a saber:

- ✓ 28 cargas factoriales ( $\lambda$ ) o relaciones de medida entre las variables de medición y las variables latentes (para fijar la escala, la carga factorial más grande por variable latente se deja constante e igual a 1);
- ✓ 63 umbrales ( $\tau$ ) -c – 1 categorías por cada variable categórica observada-;
- ✓ 38 parámetros de pendiente de regresión (uno para cada covariable predictora de propensión modo 1 y propensión modo 2); y,
- ✓ 11 relaciones entre las variables latentes o constructos (5 varianzas residuales: una por cada variable latente, más 6 relaciones de dependencia estructurales entre los constructos).

Debido a que el modelo teórico estructural condicional propuesto es un modelo recursivo, que no contiene términos de interacción, donde el tamaño de la muestra es lo suficientemente grande ( $N=8053$ ), cuenta con al menos tres variables observadas de medida por constructo, y su modelo de medida está identificado, es posible concluir que el modelo teórico estructural condicional también está identificado. De hecho, y al igual que el modelo de medida, está sobre-identificado.

#### X.5.2 Evaluación de la Bondad de Ajuste del Modelo Teórico Estructural Condicional

Antes de analizar los resultados obtenidos sobre la bondad de ajuste del modelo teórico estructural condicional, interesa señalar -más allá de la especificación realizada en la sub-sección anterior-, las implicancias derivadas de condicionar el modelo teórico estructural de la sección X.4.

El modelo estructural sin condicionar contaba con cinco variables latentes (dos exógenas y tres endógenas) y con 33 variables observadas de medida; dando lugar a un modelo con 488 grados de libertad y dada la especificación realizada, se estimaron 103 parámetros. Ahora, el modelo teórico estructural condicionado cuenta con las mismas 33 variables observadas y las cinco variables latentes, y además incluye 19 variables independientes. Esto determina: i) un aumento en los grados de libertad del modelo; ii) un aumento del número de parámetros libres a estimar; y iii) un aumento en la cantidad de constructos endógenos (ahora son 5).

Dado que el  $N$  es constante, el aumento en la complejidad del modelo implica un aumento en el valor estimado del estadístico Ji-Cuadrado de Bondad de Ajuste, por lo que no es razonable esperar que dicho estadístico sea no significativo. Es decir, si la bondad de ajuste del modelo estructural sin condicionar, medida por el estadístico Ji-Cuadrado, fue significativa, también debería serlo para el modelo teórico estructural condicional.

A continuación en la tabla 42 se presentan las medidas de bondad de ajuste. Tal como puede observarse y como era esperable, el valor del estadístico Ji-Cuadrado de bondad de ajuste del modelo teórico estructural condicional es mayor que el del modelo teórico estructural sin condicionar, y por lo tanto, igualmente significativo. Este resultado sugiere un ajuste pobre del modelo estructural condicional. Sin embargo, tal como se detalló este resultado no es concluyente debido a la sensibilidad de la Ji-Cuadrado al tamaño de la muestra y a la complejidad del modelo propuesto.

Interesa entonces analizar las otras medidas de bondad de ajuste, tales como: RMSEA, CFI y TLI. Tal como se observa en la tabla 42, condicionar el modelo estructural determinó una mejora sustantiva del RMSEA pasando de 0.054 a 0.041; esto sugiere un ajuste muy satisfactorio del modelo teórico estructural condicional. Por otra parte y a pesar de que las medidas de CFI y TLI del modelo estructural condicional empeoraron un poco en relación a las estimadas para el mismo modelo sin condicionar, siguen estando dentro de los márgenes establecidos en la literatura que sugieren un ajuste satisfactorio del modelo estructural.

En síntesis:

- ✓ estos resultados implican que hasta este punto del análisis no se ha encontrado evidencia que sugiera que el modelo teórico estructural condicional carece de validez.



Tabla 42: Medidas de Bondad de Ajuste del Modelo Estructural Condicional comparadas con las del Modelo Estructural sin condicionar

	Modelo Estructural Condicional	Modelo Estructural
Parámetros libres estimados	140	103
N	8053	8053
<b>Medidas de Ajuste Absoluto</b>		
Prueba Ji-Cuadrado		
$\chi^2$	15784.388*	11855.593*
$df$	1078	488
$p - value$	0.0000 <sup>1</sup>	0.0000 <sup>1</sup>
RMSEA	0.041	0.054
90% IC para el valor estimado del RMSEA	(0.041; 0.042)	(0.053; 0.055)
Probabilidad RMSEA $\leq$ 0.05	1.000	0.000
<b>Medidas de Ajuste Incremental</b>		
CFI	0.925	0.951
TLI	0.920	0.947
<b>Ji-Cuadrado del Modelo Nulo</b>		
$\chi^2$	197262.432	230934.785
$df$	1155	528
$p - value$	0.0000 <sup>1</sup>	0.0000 <sup>1</sup>

Notas: 1. Estadísticamente significativa al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.  
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

### X.5.3 Modelo teórico estructural condicional *versus* el modelo sin condicionar

Antes de analizar los parámetros estructurales del modelo condicional en relación al sin condicionar, en la tabla 43 se comparan los parámetros de medida estimados de cada uno de estos modelos, como también las varianzas promedio extraídas y la confiabilidad compuesta de los constructos. El objetivo es analizar la estabilidad de las estimaciones.

Tabla 43: Comparación entre los parámetros de medida del Modelo Estructural condicional *versus* el modelo sin condicionar

		Modelo Estructural Condicional	Modelo Estructural
Variable Latente	Ítems	Parámetros de medida estandarizados <sup>1</sup>	Parámetros de medida estandarizados <sup>1</sup>
<b>PROP_M1</b>	P_18_C_R	0.510	0.587
	P_18_F_R	0.580	0.600
	P_18_G_R	0.580	0.555
	P_18_I_R	0.788	0.705
	P_18_J_R	0.670	0.628
<b>PROP_M2</b>	P_18_B_R	0.771	0.772
	P_18_D_R	0.843	0.823
	P_18_E_R	0.616	0.586
	P_18_H_R	0.670	0.651
	P_18_K_R	0.636	0.682
	P_1_A_R	0.859	0.860

SATPROEV	P_1_B_R	0.964	0.962
	P_2_A_R	0.468	0.447
	P_2_C_R	0.970	0.970
	P_5_R_D	0.870	0.874
	P_7_R	0.802	0.803
	P_8_R	0.949	0.953
	P_9_A_R	0.363	0.360
	P_9_B_R	0.331	0.317
	P_9_D_R	0.653	0.644
FLXEPSNI	P_3_R_D	0.572	0.542
	P_10_R	0.928	0.927
	P_11_R_D	0.857	0.852
	P_9_C_R	0.597	0.612
	P_17_A_R	0.648	0.643
	P_17_B_R	0.621	0.621
DIRREF	P_17_D_R	0.528	0.547
	P_20_G_R	0.821	0.820
	P_20_K_R	0.976	0.975
	P_20_L_R	0.645	0.617
	P_20_I_R	0.991	0.991
	P_20_J_R	0.908	0.912
	P_20_E_R	0.834	0.833
<b>AVE: Varianza Promedio Extraída</b>			
PROP_M1		0.4005	0.3808
PROP_M2		0.5076	0.5011
SATPROEV		0.5796	0.5767
FLXEPSNI		0.4805	0.4788
DIRREF		0.7575	0.7519
<b>CR: Confiabilidad Compuesta</b>			
PROP_M1		0.7618	0.7533
PROP_M2		0.8289	0.8320
SATPROEV		0.9254	0.9243
FLXEPSNI		0.8606	0.8605
DIRREF		0.9484	0.9469

Notas: 1. Estadísticamente significativos al nivel  $\alpha = 5\%$ . Estimaciones realizadas con M-PLUS Versión 7.4.  
Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014).

Tal como se desprende de la tabla precedente, las estimaciones de los parámetros de medida permanecen estables. Del mismo modo, se mantienen relativamente constantes la varianza promedio extraída y la confiabilidad de los constructos.

En la tabla 44 se presentan los parámetros estructurales estandarizados y sin estandarizar estimados del modelo condicionado *versus* el modelo sin condicionar.

Tabla 44: Parámetros estructurales del Modelo Teórico Estructural Condicionado versus el Modelo sin condicionar: conocimiento

Relación Estructural	Modelo Teórico Estructural Condicionado				Modelo Teórico Estructural sin Condicionar			
	Parámetro sin estandarizar	S.E	Est./S.E.	Two-Tailed P-Value	Parámetro estandarizado	Parámetro sin estandarizar	S.E	Est./S.E.
PROP_M1 ---> FLXEPSNI	0.474	0.018	26.137	0.000	0.407	0.451	0.023	19.575
PROP_M2 ---> FLXEPSNI	0.272	0.017	16.217	0.000	0.254	0.121	0.019	6.235
PROP_M1 ---> SATPROEV	0.007	0.019	0.393	0.694	0.006	0.096	0.021	4.571
PROP_M2 ---> SATPROEV	-0.154	0.016	-9.544	0.000	-0.138	-0.194	0.017	-11.582
FLXEPSNI ---> SATPROEV	0.889	0.016	54.284	0.000	0.852	0.884	0.014	61.563
SATPROEV ---> DIRREF	-0.668	0.011	-59.898	0.000	-0.655	-0.675	0.011	-61.580
Cov (PROP_M1; PROP_M2)	----	----	----	----	----	0.223	0.008	28.738

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014). E-PLUS Versión 7.4.

Tabla 45: Varianza Explicada ( $R^2$ ) de los constructos endógenos Modelo Teórico Estructural Condicionado versus el Modelo sin Condicionar

Latent Variable	Modelo Teórico Estructural Condicionado				Modelo Teórico Estructural sin Condicionar			
	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1	0.053	0.006	8.689	0.000	---	---	---	---
PROP_M2	0.095	0.007	13.621	0.000	---	---	---	---
SATPROEV	0.690	0.013	51.434	0.000	0.715	0.013	53.688	0.000
FLXEPSNI	0.229	0.015	15.410	0.000	0.158	0.011	14.555	0.000
DIRREF	0.429	0.013	32.695	0.000	0.437	0.013	33.728	0.000

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014). E-PLUS Versión 7.4.

Como se observa en la tabla 44, todos los resultados –excepto uno- son conceptualmente idénticos a los obtenidos en el modelo teórico estructural sin condicionar de la sección (X.4). Si bien existen algunas diferencias de magnitud de los parámetros estructurales estimados, no se observan cambios respecto a su dirección y significancia estadística en cinco de los seis parámetros estructurales estimados. En este sentido y en términos generales, la discusión teórica y sus implicancias -o significado conceptual-, realizada en la sección anterior, es igualmente válida. A efectos de no repetir lo ya dicho, a continuación sólo se discuten algunos matices conceptuales importantes que se derivan de la relación estructural entre propensión modo 1 de producción de conocimiento y satisfacción con el proceso de evaluación, que en el modelo teórico estructural condicional resultó ser no significativa como se esperaba –dada la pequeña magnitud que alcanzó en la estimación del modelo estructural sin condicionar-.

Al controlar el efecto que tienen variables tales como la edad, el nivel de categorización, el área cognitiva, el tipo de institución y el género de los investigadores sobre su percepción en relación a los ámbitos y motivaciones de la producción de conocimiento (propensión modo 1 y propensión modo 2), la flexibilidad epistemológica del sistema de evaluación, la satisfacción que tienen con el proceso de evaluación y el interés que tienen en la dirección de la reforma al sistema de evaluación, el único parámetro estructural que deja de ser significativo es el que mide la relación entre propensión modo 1 de producción de conocimiento y satisfacción con el proceso de evaluación. Es decir, se rechaza la hipótesis que sostiene que existe una relación de dependencia positiva y significativa entre esas dos variables latentes (PROP\_M1 ---> SATPROEV) (H2). Ahora con los resultados controlados, la evidencia sugiere que no hay una la relación directa significativa entre propensión modo 1 y satisfacción con el proceso de evaluación.

También es posible observar –cuando se condiciona el modelo teórico estructural- un aumento en las magnitudes de los parámetros estructurales que miden las relaciones directas entre: i) propensión modo 1 y flexibilidad epistemológica del sistema de evaluación, y ii) flexibilidad epistemológica del sistema de evaluación y satisfacción con el proceso de evaluación (ver tabla 44). Esos últimos dos resultados sugieren que el efecto de *propensión modo 1* sobre la *satisfacción con el proceso de evaluación* está mediado –o es transmitido- por una tercera variable, a saber: la *flexibilidad epistemológica del sistema de evaluación*. De hecho así es, en la tabla 46 se presentan los resultados obtenidos de la estimación del efecto indirecto de propensión modo 1 sobre satisfacción con el proceso de evaluación, que es mediado por la flexibilidad epistemológica del sistema de evaluación.

*Tabla 46: Efecto total, directo e indirecto de propensión modo 1 sobre satisfacción con el proceso de evaluación (resultados estandarizados)*

Efecto	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Indirecto: PROP_M1 ---> FLXEPSNI ---> SATPROEV	0.347	0.014	24.019	0.000
Directo: PROP_M1 ---> SATPROEV	0.006	0.016	0.394	0.694
Total	0.353	0.014	25.659	0.000

Fuente: Elaboración propia en base a microdatos de la encuesta "Percepciones sobre la evaluación académica" (FCCyT, 2014). Estimaciones realizadas con M-PLUS Versión 7.4.

Tal como se observa en la tabla precedente, el efecto indirecto de propensión modo 1 sobre satisfacción con el proceso de evaluación es positivo y significativo estadísticamente, y alcanza un valor estandarizado de 0.347. En otros términos, casi la totalidad (98.3%) del efecto total de

propensión modo 1 sobre satisfacción con el proceso de evaluación es transmitido indirectamente a través de la flexibilidad epistemológica del sistema de evaluación. Se puede concluir que:

- ✓ El efecto mediador de *flexibilidad epistemológica del sistema de evaluación* es muy importante para explicar la *satisfacción con el proceso de evaluación* que tienen los investigadores con mayor puntuación en *propensión modo 1* de producción de conocimiento.

Este resultado representa un hallazgo importante en la medida que refuerza la relevancia de la flexibilidad epistemológica del sistema de evaluación para explicar la satisfacción con el proceso de evaluación que tienen los investigadores. La relevancia de esa variable latente para explicar la satisfacción con el proceso de evaluación que tienen los investigadores ya había sido constada en la estimación del modelo teórico estructural sin condicionar. Sin embargo aquí, en la estimación del modelo estructural condicional, la relevancia de dicha variable latente se refuerza por dos razones, por un lado porque el parámetro estructural estandarizado que mide esa relación aumentó, pasando de 0.845 –en el modelo sin condicionar- a 0.852 –en el modelo condicionado- y, por otro lado, por su papel mediador para explicar la satisfacción con el proceso de evaluación que tienen los investigadores que exhiben una mayor propensión modo 1.

Es decir, la satisfacción con el proceso de evaluación que tienen los investigadores que exhiben mayor propensión modo 1 de producción de conocimiento depende de la flexibilidad epistemológica que perciben del sistema de evaluación. Dicho de otro modo, para estos investigadores lo que realmente importa para explicar su satisfacción con el proceso de evaluación es que el sistema de evaluación permita diversidad de perfiles y de trayectorias de investigación, procesos de investigación de largo plazo y colectivos, el desarrollo de estrategias de investigación originales y novedosas, y la obtención de resultados de investigación suficientes –confiables- y científicamente relevantes, y no directamente de que el proceso de evaluación sea capaz de reconocer su desempeño. La explicación más razonable para este resultado es que el proceso de evaluación sí cuenta con la información necesaria para elaborar juicios académicos apropiados sobre el desempeño de los investigadores que exhiben mayor propensión modo 1 e incluye todas sus contribuciones importantes que, de acuerdo con la definición conceptual de propensión modo 1, se corresponden con la publicación de artículos, libros y capítulos de libro, y que a la vez representan su principal fuente de reconocimiento científico.

Este resultado cobra relevancia al analizarlo conjuntamente con la *satisfacción con el proceso de evaluación* que tienen los investigadores con mayor *propensión modo 2*, ya que para estos investigadores sí es posible observar un efecto directo, negativo y significativo entre *satisfacción con el proceso de evaluación* y *propensión modo 2* (ver tabla 44). Es decir, conforme aumenta la *propensión modo 2* de producción de conocimiento que orienta el trabajo de los investigadores categorizados en el sistema, su *satisfacción con el proceso de evaluación* del desempeño de la investigación que allí se realiza disminuye significativamente. Esto es evidencia que demuestra que el proceso de evaluación que se implementa en el SNI, a diferencia de lo indicado para los investigadores con mayor propensión modo 1, no está siendo capaz de reconocer el desempeño de los investigadores que exhiben mayor propensión modo 2, es decir que no cuenta con la información necesaria para elaborar juicios académicos apropiados sobre su desempeño, ni incluye todas sus contribuciones importantes y los criterios y mecanismos de selección de los evaluadores no

garantizan evaluaciones apropiadas de acuerdo a la especificidad de este tipo de producción de conocimiento.

En síntesis:

- ✓ Esos resultados son más precisos que los obtenidos para el modelo estructural sin condicionar y proveen evidencia específica de que el SNI no es neutro, por el contrario, introduce sesgos en su evaluación a favor de la producción de conocimiento modo 1 y en contra de la producción de conocimiento modo 2.

Así y en coincidencia con los hallazgos de investigaciones pasadas, estos resultados sugieren que el SNI, está alterando el contenido y dirección de la producción de conocimiento. En particular hacia una producción de conocimiento más caracterizada por los atributos del Modo 1 que los que caracterizan al Modo 2. Los criterios de evaluación que utiliza el SNI, introducen sesgos en su evaluación que desalientan la producción de conocimiento Modo 2. Debido a la importancia del SNI en términos del prestigio académico que provee, el incremento sustantivo de los ingresos económicos mensuales de los investigadores y el mejor acceso a recursos externos competitivos para financiar la investigación que realizan, los investigadores adaptan su producción de conocimiento para cumplir con las exigencias que impone el sistema de evaluación, es decir, producen lo que el sistema de evaluación reconoce y valora. De ese modo se limita el progreso de la ciencia e impacto de la producción de conocimiento, en particular porque se limita la diversidad de la producción de conocimiento y no reconoce aquella producción orientada a atender los desafíos que impone el desarrollo económico y social del país, es decir, la producción de conocimiento contextualizada dirigida por la búsqueda de soluciones a problemas específicos dependientes del contexto, ya sean sociales, productivos y/o de política pública. De ese modo, la evaluación que se implementa en el SNI, más que ayudar al cumplimiento integral de sus objetivos de política, los contradice. En particular, compromete el alcance de aquellos objetivos de política del SNI orientados a incentivar la producción de conocimiento científico y tecnológico de calidad para contribuir al desarrollo social y económico del país.

La principal ventaja del modelo condicional sobre el no condicional, es que la introducción de las variables de control (covariables  $x$ ) permitió un aumento de la eficiencia de los parámetros estimados y la reducción de dos fuentes de sesgo en los resultados. Una a través del aislamiento de ciertas variables perturbadoras, y la otra, a través de la aplicación del enfoque de estimación por máxima verosimilitud con información completa (FILM) para el tratamiento de los datos faltantes – en el modelo sin condicionar, el tratamiento de los datos faltantes siguió el enfoque *pairwise deletion*-. Así:

- ✓ el modelo teórico estructural condicional es más preciso y confiable; y además, dada la consistencia en los resultados obtenidos, su validez teórico-conceptual se ve reforzada.

Tal como se observa en la tabla 45, las varianzas explicadas de los constructos endógenos son todas significativas, y para los casos que corresponde, si bien se redujeron un poco las varianzas explicadas de los constructos satisfacción con el proceso de evaluación y dirección de la reforma del sistema de evaluación -en relación al modelo estructural sin condicionar-, se incrementó bastante la varianza explicada de flexibilidad epistemológica del sistema de evaluación.

El modelo teórico estructural condicional explica el 69% de la varianza de la satisfacción con el proceso de evaluación que tienen los investigadores, el 42.9% de la varianza de su interés en la dirección de la reforma del sistema de evaluación, un 22.9% de la varianza de la flexibilidad epistemológica del sistema de evaluación que perciben, y un 5.3% y un 9.5% de las varianzas de sus propensiones modo 1 y modo 2 respectivamente. Si bien la capacidad del modelo para explicar la varianza de la flexibilidad epistemológica del sistema de evaluación que perciben los investigadores mejoró, ésta sigue siendo relativamente baja, lo que sugiere la existencia de otras dimensiones conceptuales también importantes para explicarla. Estudiar la flexibilidad epistemológica de los sistemas de evaluación del desempeño de la investigación, incluidos sus determinantes, no fue el objetivo de esta tesis, pero a partir de los resultados obtenidos en este trabajo, emerge la falta de conocimiento en relación a ese tema.

## X.6 Conclusión: principales resultados

A continuación se sintetizan los principales resultados obtenidos de cada una de las tres fases de análisis.

El AFE provee evidencia que justifica la pertinencia de implementar un AFC para poner a prueba un modelo de medida multifactorial de cinco factores.

Los resultados derivados de la implementación del AFC permiten concluir que el modelo multifactorial de cinco factores propuesto para medir las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento es válido, confiable y preciso. Así:

- ✓ *propensión modo 1* es un concepto multidimensional, cuya operacionalización representa una aproximación válida y confiable para medir, de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento de los investigadores categorizados en el SNI, el fenómeno Modo 1 de producción de conocimiento conceptualizado por la teoría y permite inferir en relación al Modo 1 como un tipo ideal de producción de conocimiento;
- ✓ *propensión modo 2* es un concepto multidimensional, cuya operacionalización representa una aproximación válida y confiable para medir, de acuerdo con los ámbitos y motivaciones que predominantemente orientan la generación de conocimiento de los investigadores categorizados en el SNI, el fenómeno Modo 2 de producción de conocimiento conceptualizado por la teoría y permite inferir en relación al Modo 2 como un tipo ideal de producción de conocimiento;
- ✓ *flexibilidad epistemológica del sistema de evaluación* es un concepto multidimensional, cuya operacionalización representa una aproximación válida y confiable para medir, de acuerdo con la percepción de los investigadores categorizados en el SNI, en qué medida el sistema de evaluación permite procesos de producción de conocimiento de calidad, originales, de largo plazo, colectivos y con diversidad de perfiles y trayectorias de investigación, y permite inferir sobre la relación entre la evaluación y las características epistemológicas de la producción de conocimiento.
- ✓ *satisfacción con el proceso de evaluación* es un concepto multidimensional, cuya operacionalización representa una aproximación válida y confiable para medir el nivel de satisfacción que tienen los investigadores categorizados en el SNI con el proceso de evaluación

al que están sometidos, respecto a si dicho proceso de evaluación es capaz de reconocer integralmente y objetivamente su desempeño, y permite inferir sobre la relación entre la percepción que tienen los investigadores sobre el proceso de evaluación y su desempeño de investigación.

- ✓ *dirección de la reforma del sistema de evaluación* es un concepto multidimensional, cuya operacionalización representa una aproximación válida y confiable para medir el interés que tienen los investigadores categorizados en el SNI de reformarlo en una dirección cuyo proceso de evaluación sea más cualitativo, que respete las diferencias disciplinares en sus prácticas de producción de conocimiento y que valore y reconozca la investigación colectiva, multi/inter/transdisciplinaria y los resultados de la investigación aplicada u orientada a la resolución de problemas.

De acuerdo con las estimaciones realizadas del modelo teórico estructural propuesto para explorar la relación entre la evaluación del desempeño individual de investigación y los modos diferenciales de producción de conocimiento. La evidencia empírica encontrada es fuerte y sostiene la validez del modelo teórico estructural, así como la estabilidad y confiabilidad de los resultados obtenidos. La evidencia muestra que:

- ✓ La *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI depende significativamente y positivamente de su *propensión modo 1* de producción de conocimiento. Un aumento de una unidad estándar en la puntuación de *propensión modo 1* tiene un efecto directo de 0.407 unidades estándar de aumento en la percepción sobre la *flexibilidad epistemológica del sistema de evaluación*. Este resultado confirma la H1.
- ✓ La *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI depende significativamente y positivamente de su *propensión modo 2* de producción de conocimiento. Un aumento de una unidad estándar en la puntuación de *propensión modo 2* tiene un efecto directo de 0.254 unidades estándar de aumento en la percepción sobre la *flexibilidad epistemológica del sistema de evaluación*. Este resultado confirma parcialmente la H3, ya que si bien existe una relación significativa entre ambas variables latentes, su signo va en la dirección contraria a la hipotetizada. Sin embargo, interesa remarcar que los investigadores con mayor *propensión modo 1* perciben una *flexibilidad epistemológica del sistema de evaluación* 1.6 veces mayor que sus pares con mayor *propensión modo 2*.
- ✓ Se rechaza H2. Es decir, no se encontró evidencia que permita sostener una relación directa significativa y positiva entre *satisfacción con el proceso de evaluación* y *propensión modo 1*. Sin embargo, importa resaltar que sí se encontró evidencia de una relación indirecta positiva y significativa entre esas dos variables. Es decir, la *satisfacción con el proceso de evaluación* que tienen los investigadores que exhiben mayor *propensión modo 1*, está mediada o es transmitida por una tercera variable, a saber: la *flexibilidad epistemológica del sistema de evaluación* que perciben. El efecto total de *propensión modo 1* sobre *satisfacción con el proceso de evaluación* alcanza un valor estandarizado de 0.353, cuyo efecto indirecto es positivo y significativo, y alcanza un valor estandarizado de 0.347. Así, un aumento de una unidad estándar en la puntuación de *propensión modo 1* tiene un efecto indirecto de 0.347 unidades estándar de aumento en la *satisfacción con el proceso de evaluación* que es transmitido por la *flexibilidad epistemológica del sistema de evaluación* que perciben esos investigadores.



- ✓ La *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI depende significativamente y negativamente de su *propensión modo 2*. Un aumento de una unidad estándar en la puntuación de *propensión modo 2* tiene un efecto directo de -0.138 unidades estándar de descenso en la *satisfacción con el proceso de evaluación*. Este resultado confirma H4.
- ✓ La *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI depende significativamente y positivamente de la *flexibilidad epistemológica del sistema de evaluación*. Un aumento de una unidad estándar en la puntuación de *flexibilidad epistemológica del sistema de evaluación* tiene un efecto directo de 0.852 unidades estándar de aumento en la *satisfacción con el proceso de evaluación*. Este resultado confirma H5.
- ✓ El interés en la *dirección de la reforma del sistema de evaluación* que tienen los investigadores categorizados en el SNI depende significativamente y negativamente de su *satisfacción con el proceso de evaluación*. Un aumento de una unidad estándar en la puntuación de *satisfacción con el proceso de evaluación* tiene un efecto directo de -0.655 unidades estándar de descenso en la *dirección de la reforma del sistema de evaluación*. Este resultado confirma H5.

## CAPÍTULO XI: Conclusiones

El estudio de la relación entre la evaluación del desempeño de la investigación y la producción de conocimiento ha cobrado vigor, relevancia y pertinencia a partir de la implementación de los sistemas de evaluación del desempeño de la investigación en distintos sistemas públicos de ciencia de diversos países –desarrollados y en desarrollo- a inicios de la década de 1980. La implementación de esos sistemas de evaluación ha contribuido a cambios profundos en la gobernanza de la ciencia, en términos de qué conocimiento se produce, cómo se produce, en qué contextos, cómo se organiza, cómo se evalúa y se recompensa. Distintos estudios, internacionales y nacionales, realizados en diversos campos de las ciencias sociales –en sociología de la ciencia, gobernanza, evaluación, política científica, educación superior e innovación, entre otros- han aportado reflexión y evidencia sobre esos cambios y sus consecuencias probables sobre la producción de conocimiento.

Así, distintos trabajos han postulado un conjunto de relaciones teóricas entre los *sistemas de evaluación fuertes* del desempeño de la investigación y la *dirección, contenido y características epistemológicas* de la producción de conocimiento (capítulo III). Diversos estudios empíricos han reportado evidencia de distintos países –principalmente desarrollados- en los que se han implementado dichos sistemas de evaluación, que es consistente con las relaciones postuladas por la teoría. A la vez, también muestran la transformación en las prácticas de producción de conocimiento de los investigadores de los distintos campos científicos como consecuencia probable de esos sistemas de evaluación. En particular, de aquellos investigadores que pertenecen a los campos que gozan de menor prestigio en el sistema de ciencia, haciendo verosímil el mecanismo que conecta a la evaluación de la investigación con el contenido, dirección y características epistemológicas de la producción de conocimiento postulado por la teoría.

Los hallazgos de investigaciones pasadas reportados en la literatura son consistentes en documentar, como consecuencia probable de la implementación de evaluaciones cuantitativas o centrada en indicadores bibliométricos, cierta homogeneización o estandarización de los tipos de investigación, los productos que genera y sus medios de difusión. Destacan que esas evaluaciones han favorecido una producción de conocimiento especializada disciplinariamente, internacionalizada y acelerada, donde la producción de artículos científicos y su publicación en revistas con alto factor de impacto e indexadas a bases de datos internacionales representan los productos mejor valorados en esas evaluaciones. Y además, que otros tipos de investigación, en particular la producción de conocimiento contextualizada o preocupada por atender los desafíos que le impone su contexto social, cultural, político y económico es desvalorizada, ya que los productos que genera – por ejemplo: reportes para informar la toma de decisiones de política pública, reportes técnicos, desarrollos tecnológicos, desarrollo de técnicas e instrumentos, libros y capítulos de libro entre otros- son invisibilizados ante esos indicadores de desempeño. Es decir, las evaluaciones del desempeño de la investigación informadas por indicadores bibliométricos no son capaces de valorar y reconocer la diversidad de objetivos y productos derivados de la investigación. Adicionalmente, diversos estudios han sugerido que la aplicación de ciertos indicadores bibliométricos (cantidad de publicaciones, de citas y factores de impacto) en la evaluación introducen sesgos contra ciertas características epistemológicas de los procesos de producción conocimiento, tales como: la originalidad, relevancia y confiabilidad de los resultados de la investigación, la investigación multi/inter/transdisciplinaria, colectiva –en particular si involucra a actores externos al ámbito académico- y de la investigación de largo plazo. Los estudios también documentan que el

mecanismo más plausible por el cual ocurren todos esos efectos –constitutivos- probablemente derivados de las evaluaciones cuantitativas sobre el contenido, dirección y características epistemológicas de la producción de conocimiento, se vincula con las estrategias adaptativas de los investigadores sobre sus decisiones de producción de conocimiento para cumplir con las exigencias que les impone el sistema de evaluación, ya que de eso depende el prestigio y reconocimiento académico que alcanzan y a su vez, el mejor acceso a fuentes externas –competitivas- para financiar la investigación que realizan, o para competir en mejores condiciones por plazas en las universidades u otras organizaciones, o –como en el caso de México y Uruguay- para poder acceder a un aumento de sus ingresos mensuales –como complemento extrasalarial-. Por lo argumentado, diversos estudios señalan que dichos sistemas de evaluación más que contribuir al progreso de la ciencia, la limitan y reducen el impacto potencial de la producción de conocimiento en términos de su contribución al desarrollo social y económico, contradiciendo con ello, los objetivos de política pública de CTI que se plantean. (capítulo IV)

Además, el debate internacional sobre el uso de indicadores bibliométricos como criterios centrales de evaluación del desempeño de la investigación y el nacional sobre el SNI, son consistentes con la literatura especializada en el diagnóstico que hacen sobre los desafíos o problemas que exhiben esos sistemas de evaluación para el progreso de la ciencia. Además, hacen visible la insatisfacción que genera el uso de los indicadores bibliométricos como criterios centrales de evaluación en la comunidad académica –internacional y nacional-, porque los mismos, al no reconocer los diversos perfiles, tipos y productos derivados de la investigación y el impacto real de la producción de conocimiento en lo social, político y/o económico –no solamente en términos de la incidencia internacional de las publicaciones-, proveen una imagen distorsionada del desempeño de investigadores, grupos u organizaciones. Es decir, dichos indicadores son cuestionados por la escasa validez que tienen para reflejar la productividad, calidad e impacto de la producción de conocimiento.

Adicionalmente, de esos debates también emergen un conjunto de recomendaciones para mejorar las prácticas de evaluación del desempeño de la investigación, orientadas a valorar y reconocer la calidad e impacto de los diversos productos derivados de la investigación por sus propios méritos, atendiendo a las especificidades de las diversas áreas de conocimiento, los objetivos de investigación que se plantean, los contextos en que se produce el conocimiento y la valoración que hacen las audiencias para las cuales ese conocimiento fue creado. Así, la recomendación de fondo es que la evaluación del desempeño de la investigación tiene que estar informada por el juicio cualitativo realizados por los pares expertos en el campo objeto de evaluación y los indicadores bibliométricos pueden ser utilizados –siempre que se consideren todas sus limitaciones- como insumos adicionales o complementarios de información, pero nunca pueden ser el insumo central ni más importante para informar el juicio experto de los pares. (capítulos IV y VII).

Por otro lado, otro cuerpo de literatura discute los cambios en el rol del conocimiento, organización, en sus formas de producción, legitimación y difusión, observables desde la segunda mitad del SXX en los países altamente industrializados. En particular, las contribuciones teóricas sobre los modos diferenciales de producción de conocimiento (Modo 1 y Modo 2), que además de proponer un conjunto de atributos que caracterizan teóricamente a esos modos de producción de conocimiento como formas diferenciales de producción, advierten sobre la necesidad de desarrollar formas flexibles, adaptativas y contextualizadas para evaluar/controlar la calidad de la producción de

conocimiento realizada en esos modos diferenciales de producción. Destacan que si la evaluación de la investigación no considera las especificidades de los modos diferenciales de producción de conocimiento, no la reconoce y, por lo tanto, puede castigar a las *nuevas formas de producción de conocimiento* caracterizadas por su orientación hacia la atención de problemas complejos dependientes del contexto, transdisciplinarios, que involucran la participación de actores diversos y que ocurre en los *contextos de aplicación*. Además, estas contribuciones destacan que si alguno de esos modos no es adecuadamente valorado y reconocido se pone en riesgo el avance de la ciencia, ya que los modos 1 y 2 interactúan entre sí, se complementan debido a que uno puede ser el punto de partida para la producción de conocimiento del otro y *viceversa*. (capítulo II)

Sin embargo y a pesar de las conexiones entre las contribuciones teóricas sobre los modos diferenciales de producción de conocimiento y la teoría que vincula a la evaluación con el contenido, dirección y características epistemológicas de la producción de conocimiento<sup>75</sup>, los estudios empíricos no han explorado las relaciones entre los sistemas de evaluación fuertes del desempeño de la investigación y los modos diferenciales de producción de conocimiento, ni han explorado las relaciones entre las evaluaciones cuantitativas del desempeño de la investigación y los modos 1 y 2 de producción de conocimiento.

Esta tesis se insertó en esos espacios de falta de conocimiento, aportando evidencia nueva y un análisis riguroso, válido, confiable, original, relevante conceptualmente y pertinente por sus implicancias prácticas para la política, sobre las relaciones teóricas entre un sistema de evaluación fuerte como el SNI -e informado principalmente por indicadores bibliométricos- del desempeño de la investigación y los modos diferenciales de producción de conocimiento –Modo 1 y Modo 2- que predominantemente orientan el trabajo de los investigadores categorizados en ese sistema.

Esta tesis, a través del desarrollo y validación de un modelo de ecuaciones estructurales, abordó la pregunta sobre ¿qué relaciones hay entre un sistema de evaluación fuerte del desempeño individual de investigación como el SNI y los Modos 1 y 2 de producción de conocimiento que predominantemente orientan la investigación que realizan los investigadores categorizados en ese sistema de evaluación y cuál es la naturaleza de esas relaciones?

Para poder explorar la naturaleza de esas relaciones y así dar respuesta a la pregunta de investigación planteada, en esta tesis se desarrollaron un conjunto de operaciones teóricas-conceptuales, metodológicas y empíricas que se sintetizan a continuación.

Con base en la teoría, el debate internacional y los hallazgos de investigaciones pasadas, se elaboró el marco teórico conceptual que justifica y provee los fundamentos teóricos sobre los que se desarrolló el análisis que se hizo en esta tesis (capítulo V). Así se definieron conceptualmente los cinco constructos relevantes para el análisis y se establecieron las seis hipótesis sobre las relaciones entre ellos y que se pusieron a prueba en esta investigación. A saber:

- ✓ H1: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *flexibilidad epistemológica del sistema de evaluación*.

---

<sup>75</sup> Ver la integración que se hace en esta tesis de esos dos cuerpos teóricos de literatura en la sub-sección III.5 del capítulo III.

- ✓ H2: la *propensión modo 1* de producción de conocimiento se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.
- ✓ H3: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *flexibilidad epistemológica del sistema de evaluación*.
- ✓ H4: la *propensión modo 2* de producción de conocimiento se relaciona significativamente y negativamente con la *satisfacción con el proceso de evaluación*.
- ✓ H5: la *flexibilidad epistemológica del sistema de evaluación* se relaciona significativamente y positivamente con la *satisfacción con el proceso de evaluación*.
- ✓ H6: la *satisfacción con el proceso de evaluación* se relaciona significativamente y negativamente con la *dirección de la reforma del sistema de evaluación*.

Estos cinco constructos y las relaciones entre ellos hipotetizadas, conjuntamente representan la teoría estructural sobre las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento desarrollada y validada en esta tesis.

Con base en la definición conceptual de los cinco constructos, y en los datos contenidos en la encuesta utilizada en esta investigación, se desarrolló y validó una teoría de medición que relaciona a las variables observadas de medida con los constructos que explican su variación (capítulo X). La teoría de medida propuesta fue expresada formalmente y puesta a prueba a través de un modelo de medida multifactorial de cinco factores. Para contrastar el modelo de medida con la realidad – representada por los datos- y evaluar su validez, se implementó un AFC. Con este análisis se evaluó por un lado la bondad de ajuste del modelo de medida y por otro, su validez de constructo. De aquí se concluye, con base en la fortaleza de la evidencia encontrada, que el modelo para medir las relaciones entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento es válido, confiable y preciso.

De acuerdo con las estimaciones realizadas del modelo teórico estructural propuesto para explorar la relación entre la evaluación del desempeño individual de investigación y los modos diferenciales de producción de conocimiento, la evidencia empírica encontrada es fuerte y sostiene la validez del modelo teórico estructural, así como la estabilidad y confiabilidad de los resultados obtenidos.

Adicionalmente, importa resaltar que los resultados obtenidos también son consistentes con el conjunto de trabajos -realizados por especialistas de diversas disciplinas, que reflexionan sobre el desempeño del SNI, sus problemas y los desafíos que enfrenta. Estos trabajos fueron discutidos en las secciones VII.2 a VII.5 del capítulo VII y representan fuentes adicionales de información que refuerzan la validez del fundamento teórico y la justificación de la coherencia, adecuación y pertinencia del marco conceptual desarrollado en esta tesis, sintetizado en la figura 2 y formalizado a través del modelo teórico estructural para entender las relaciones entre la evaluación que se realiza en el SNI y los modos diferenciales de producción de conocimiento observados en México. Así, en esta tesis se hicieron dos grandes esfuerzos de revisión, integración, sistematización y presentación de ambas fuentes de información –una asociada a los datos contenidos en la encuesta y la otra asociada a los datos generados por especialistas sobre el desempeño del SNI-, que a pesar de las diferencias en términos de sus estructuras metódicas, ambas convergen en la consistencia de los resultados que arrojan. La tabla 47 sintetiza los resultados y a continuación se describen brevemente.

Tabla 47: Síntesis de los resultados. Estimaciones estandarizadas del modelo teórico estructural condicional

		Variables Latentes Dependientes		
		Flexibilidad epistemológica del sistema de evaluación	Satisfacción con el proceso de evaluación	Dirección de la reforma del sistema de evaluación
Variables Latentes Independientes	Propensión Modo 1	H1: Confirmada (0.407)	H2: No Confirmada	-----
	Propensión Modo 2	H3: Confirmada con dirección contraria (0.254)	H4: Confirmada (-0.138)	-----
	Flexibilidad epistemológica del sistema de evaluación	-----	H5: Confirmada (0.852)	-----
	Satisfacción con el proceso de evaluación	-----	-----	H6: Confirmada (-0.655)

Fuente: Elaboración propia

La evidencia encontrada en esta tesis muestra que la *flexibilidad epistemológica del sistema de evaluación* que perciben los investigadores categorizados en el SNI depende significativamente y positivamente de sus modos diferenciales de producción de conocimiento (*propensión modo 1* y *propensión modo 2*).

Más específicamente, un aumento de una unidad estándar en la puntuación de *propensión modo 1* tiene un efecto directo de 0.407 unidades estándar de aumento en la *flexibilidad epistemológica del sistema de evaluación*, confirmando la H1 (ver tabla 47). A la vez, un aumento de una unidad estándar en la puntuación de *propensión modo 2* tiene un efecto directo de 0.254 unidades estándar de aumento en la *flexibilidad epistemológica del sistema de evaluación* (ver tabla 47). Este resultado confirma parcialmente la H3, ya que si bien existe una relación significativa entre ambas variables latentes, su signo va en la dirección contraria a la hipotetizada. Sin embargo, interesa remarcar que los investigadores con mayor *propensión modo 1* perciben una *flexibilidad epistemológica del sistema de evaluación* 1.6 veces mayor que sus pares con mayor *propensión modo 2*. Es decir, los investigadores que predominantemente orientan su producción de conocimiento en modo 2 perciben una menor flexibilidad del sistema de evaluación en términos de las características epistemológicas de la producción de conocimiento que realizan. Para estos investigadores, el sistema de evaluación del SNI dificulta –o permite menos- la diversidad de perfiles y de trayectorias de investigación, los procesos de investigación de largo plazo y colectivos, el desarrollo de estrategias de investigación originales y novedosas, y la obtención de resultados de investigación suficientes –confiables- y científicamente relevantes.

Estos resultados, en sintonía con la teoría y los hallazgos de investigaciones pasadas reportados en la literatura, muestran que el sistema de evaluación no es igualmente flexible desde el punto de vista de las características epistemológicas que permite de los procesos de producción de conocimiento, por el contrario, la percepción sobre la flexibilidad epistemológica del sistema de evaluación depende del modo de producción de conocimiento en que predominantemente trabajen los investigadores.

Simultáneamente, la evidencia muestra que la *satisfacción con el proceso de evaluación* que tienen los investigadores categorizados en el SNI depende:

- a. significativamente, positivamente e indirectamente de su *propensión modo 1*. Es decir, no se encontró evidencia que permita confirmar una relación directa significativa y positiva entre *satisfacción con el proceso de evaluación* y *propensión modo 1*, en ese sentido H2 fue rechazada (ver tabla 47). Sin embargo, la *satisfacción con el proceso de evaluación* que tienen los investigadores que exhiben mayor *propensión modo 1*, está mediada o es transmitida por una tercera variable, a saber: la *flexibilidad epistemológica del sistema de evaluación* que perciben. El efecto total de *propensión modo 1* sobre *satisfacción con el proceso de evaluación* alcanza un valor estandarizado de 0.353, cuyo efecto indirecto es positivo y significativo, y alcanza un valor estandarizado de 0.347. Así, un aumento de una unidad estándar en la puntuación de *propensión modo 1* tiene un efecto indirecto de 0.347 unidades estándar de aumento en la *satisfacción con el proceso de evaluación* que es transmitido por la *flexibilidad epistemológica del sistema de evaluación* que perciben esos investigadores;
- b. significativamente y negativamente de su *propensión modo 2*. Un aumento de una unidad estándar en la puntuación de *propensión modo 2* tiene un efecto directo de -0.138 unidades estándar de descenso en la *satisfacción con el proceso de evaluación*. Este resultado confirma H4 (ver tabla 47); y,
- c. significativamente y positivamente de la *flexibilidad epistemológica del sistema de evaluación* que perciben. Un aumento de una unidad estándar en la puntuación de *flexibilidad epistemológica del sistema de evaluación* tiene un efecto directo de 0.852 unidades estándar de aumento en la *satisfacción con el proceso de evaluación*. Este resultado confirma H5 (ver tabla 47).

De los resultados anteriores se puede afirmar que la variable latente o dimensión conceptual más importante para explicar la satisfacción que tienen los investigadores categorizados en el SNI con el proceso de evaluación del desempeño de la investigación que allí se implementa, es la *flexibilidad epistemológica del sistema de evaluación* que perciben.

Este resultado pone de relieve la importancia de pensar cuidadosamente, cuando se diseñan los instrumentos de evaluación del desempeño de la investigación, en su flexibilidad desde el punto de vista de su relación con las características epistemológicas de los procesos de producción de conocimiento. En ese sentido y como recomendación de política, se resalta la importancia que tiene el diseño de instrumentos de evaluación válidos, que limiten sus efectos constitutivos sobre las características epistemológicas de la producción de conocimiento; para proteger la calidad, originalidad, relevancia conceptual y confiabilidad de los resultados de investigación, su diversidad cognitiva (orientación y tipos de investigación), los procesos de producción de conocimiento colectivos (ya sea entre investigadores o entre investigadores y otros actores externos al ámbito de la academia) y la investigación de largo plazo.

A la vez, la relevancia de la *flexibilidad epistemológica del sistema de evaluación* se refuerza debido a su papel mediador para explicar también la *satisfacción con el proceso de evaluación* que tienen los investigadores que exhiben una mayor *propensión modo 1*. Es decir, la satisfacción con el proceso de evaluación que tienen los investigadores que exhiben mayor *propensión modo 1* de producción

de conocimiento depende de la flexibilidad epistemológica que perciben del sistema de evaluación. Dicho de otro modo, para estos investigadores lo que realmente importa para explicar su satisfacción con el proceso de evaluación es que el sistema de evaluación permita diversidad de perfiles y de trayectorias de investigación, procesos de investigación de largo plazo y colectivos, el desarrollo de estrategias de investigación originales y novedosas, y la obtención de resultados de investigación suficientes –confiables- y científicamente relevantes, y no directamente de que el proceso de evaluación sea capaz de reconocer su desempeño.

Ese resultado cobra importancia al analizarlo conjuntamente con la *satisfacción con el proceso de evaluación* que tienen los investigadores con mayor *propensión modo 2*, ya que para estos investigadores sí es posible observar un efecto directo, negativo y significativo entre *satisfacción con el proceso de evaluación* y *propensión modo 2*.

Conjuntamente, esos dos resultados proveen evidencia que demuestra que el proceso de evaluación que se implementa en el SNI, a diferencia de lo indicado para los investigadores con mayor propensión modo 1, no está siendo capaz de reconocer objetivamente e integralmente el desempeño de los investigadores que exhiben mayor propensión modo 2. Es decir, para estos investigadores, el proceso de evaluación del desempeño de la investigación que se implementa en el SNI no reconoce –o tiene dificultades- para reconocer su desempeño porque no cuenta con la información necesaria para elaborar juicios académicos apropiados sobre su desempeño, ni incluye todas sus contribuciones importantes y los criterios y mecanismos de selección de los evaluadores no garantizan evaluaciones apropiadas de acuerdo a la especificidad de este tipo de producción de conocimiento. Así y como otra recomendación de política, los instrumentos de evaluación además de que deben ser válidos, deben considerar la diversidad de productos que se derivan de los diversos tipos de producción de conocimiento científico, respetando sus objetivos, especificidades y prácticas de producción.

Simultáneamente, la evidencia muestra que el interés en la *dirección de la reforma del sistema de evaluación* que tienen los investigadores categorizados en el SNI depende significativamente y negativamente de su *satisfacción con el proceso de evaluación*. Un aumento de una unidad estándar en la puntuación de *satisfacción con el proceso de evaluación* tiene un efecto directo de -0.655 unidades estándar de descenso en la *dirección de la reforma del sistema de evaluación*. Este resultado confirma H5 (ver tabla 47). Este resultado es consistente, ya que el interés de los investigadores en reformar el sistema de evaluación en una la dirección más orientada a reconocer y valorar también la producción de conocimiento modo 2, depende negativamente de la satisfacción que tienen los investigadores con el proceso de evaluación.

Los resultados obtenidos proveen evidencia específica de que el SNI no es neutro, por el contrario, introduce sesgos en su evaluación a favor de la producción de conocimiento modo 1 y en contra de la producción de conocimiento modo 2. En ese sentido, los resultados obtenidos sugieren que el SNI introduce sesgos que pueden conducir a efectos constitutivos sobre la producción de conocimiento modo 2. Es decir, en la medida que la evaluación no reconoce adecuadamente las especificidades de este modo de producción de conocimiento, por acción u omisión lo castiga y los investigadores perciben menos incentivos a trabajar en este modo de producción. Así se reduce el impacto el impacto potencial de la producción de conocimiento en términos de su contribución en la búsqueda de soluciones a los principales problemas que afectan al desarrollo social y económico de México.



De ese modo y en coincidencia con los hallazgos de investigaciones pasadas, estos resultados sugieren que el SNI está alterando el contenido y dirección de la producción de conocimiento. En particular hacia una producción de conocimiento más caracterizada por los atributos del Modo 1 que los que caracterizan al Modo 2. Los criterios de evaluación que utiliza el SNI, introducen sesgos en su evaluación que desalientan la producción de conocimiento Modo 2. Debido a la importancia del SNI en términos del prestigio académico que provee, el incremento sustantivo de los ingresos económicos mensuales de los investigadores y el mejor acceso a recursos externos competitivos para financiar la investigación que realizan, los investigadores adaptan su producción de conocimiento para cumplir con las exigencias que impone el sistema de evaluación, es decir, producen lo que el sistema de evaluación reconoce y valora. Lo que se refuerza aún más, debido al uso de criterios de evaluación similares en las organizaciones que los emplean.

En ese sentido, se puede argumentar que el SNI está limitando el progreso de la ciencia e impacto de la producción de conocimiento, en particular porque limita la diversidad de la producción de conocimiento ya que no reconoce aquella producción más orientada a atender los desafíos que impone el desarrollo económico y social del país, es decir, la producción de conocimiento contextualizada dirigida por la búsqueda de soluciones a problemas específicos dependientes del contexto, ya sean sociales, productivos y/o de política pública. De ese modo y desde el punto de vista de las implicancias de política de los resultados de investigación, la evaluación que se implementa en el SNI, más que ayudar al cumplimiento integral de sus objetivos de política, los contradice. En particular, compromete el alcance de aquellos objetivos de política del SNI orientados a incentivar la producción de conocimiento científico y tecnológico de calidad para contribuir al desarrollo social y económico del país.

Esta tesis ha aportado un análisis original, válido y confiable sobre la relación entre la evaluación del desempeño de la investigación y los modos diferenciales de producción de conocimiento y empíricamente ha contribuido con evidencia nueva sobre esa relación en el contexto de un sistema de evaluación fuerte como el SNI de un país en desarrollo como lo es México. En particular, esta tesis representa uno de los primeros trabajos de investigación que muestra de forma sistemática, a partir de las percepciones de los investigadores, la no neutralidad epistémica de un sistema de evaluación del desempeño de la investigación sobre dos formas diferenciales de producción de conocimiento científico.

Desde el punto de vista de las contribuciones conceptuales, esta tesis introduce por un lado, el concepto de *flexibilidad epistemológica de los sistemas de evaluación* del desempeño de la investigación. Este concepto si bien introduce una perspectiva crítica de los modelos de evaluación del desempeño de la investigación dominantes, también supone que es posible alcanzar un modelo de evaluación flexible y neutro. Importa remarcar que la relevancia y pertinencia del concepto *flexibilidad epistemológica de los sistemas de evaluación* no está en desarrollar sistemas de evaluación flexibles y neutrales, que es cuestionable y que contradice la noción de efectos constitutivos, ya que toda evaluación por definición tiene efectos -performativos y constitutivos-; sino que su relevancia y pertinencia radica en las posibilidades que ofrece como instrumento de evaluación y monitoreo de los efectos epistémicos del sistema de evaluación sobre la producción de conocimiento científico. Es decir, la relevancia y pertinencia del concepto *flexibilidad epistemológica del sistema de evaluación* está en que puede utilizarse como un indicador adicional -que no sustituye el juicio de los expertos, sino que lo complementa- para la evaluación, monitoreo

y, si corresponde, re-diseño de los instrumentos de política pública orientados al estímulo de la investigación de calidad.

Por otro lado y como otra contribución conceptual de este trabajo, esta tesis introduce los modos 1 y 2 como dimensiones relevantes que también están siendo afectadas por los sistemas de evaluación fuertes, en particular por aquellos sistemas que son principalmente informados por indicadores bibliométricos como el SNI. Importa insistir en que el estudio específico de las relaciones entre un sistema de evaluación fuerte –informado principalmente por indicadores bibliométricos- y los modos 1 y 2 de producción de conocimiento no había sido abordado con anterioridad a este trabajo de investigación. Así, esta tesis aporta evidencia sistemática sobre efectos epistémicos del sistema de evaluación sobre la dirección y contenido de la investigación.

Desde el punto de vista de la contribución metodológica de este trabajo, en esta investigación se realizaron esfuerzos importantes para garantizar la validez de constructo de todas las medidas que se proveen. La relevancia de estos esfuerzos está en términos de la reflexión metodológica y técnica que aporta esta investigación al campo de los estudios sobre evaluación del desempeño de la investigación y al ámbito de la política pública de estímulo a la investigación y sus instrumentos de evaluación. De aquí también la relevancia conceptual y pertinencia práctica de esta investigación. Por otro lado, adicionalmente y hasta donde se sabe, en la literatura especializada -nacional e internacional- no existe ningún estudio que haya aproximado empíricamente los conceptos Modo 1 y Modo 2 con características similares a como se hace en este trabajo. De aquí la originalidad y relevancia empírica de los constructos *propensión modo 1* y *propensión modo 2* desarrollados y validados en esta investigación.

Desde el punto de vista del alcance de esta investigación, los resultados obtenidos en este trabajo son representativos de la población de investigadores categorizados en el SNI de México. Desde el punto de vista de sus limitaciones, las mismas se vinculan con el hecho de que los resultados empíricos obtenidos están basados en datos sobre las *percepciones* de los investigadores y no en sus *prácticas*.

Así, la información obtenida –como se señaló- no es sobre la *flexibilidad epistemológica del sistema de evaluación* sino sobre la percepción que tienen los investigadores categorizados en el SNI de la *flexibilidad epistemológica* de ese sistema de evaluación. En ese sentido, ciertas preguntas del cuestionario pueden ser criticadas, ya que las percepciones de los investigadores para responder a esas preguntas son cuestionables desde el punto de vista de su validez. A modo de ejemplo: “El SNI generó estandarización de trayectorias” es una pregunta en la que las percepciones de los investigadores son de dudosa validez debido a que no es evidente ni obvio que los encuestados tengan una visión clara y única de la noción de trayectorias científicas. Otro ejemplo de esta limitación son las proporciones tan elevadas de investigadores con *propensiones modo 1* y *modo 2* alta o muy alta. En ese sentido, hubiera sido preferible contar con medidas que dieran cuenta de las prácticas o del *modo* en que “efectivamente” trabajan los investigadores –del estilo: cuántas horas semanales dedica a...- para minimizar la probabilidad de respuestas “políticamente correctas” que cuestionan la validez de dichas respuestas.

Estudiar la flexibilidad epistemológica de los sistemas de evaluación del desempeño de la investigación –en particular desde las prácticas-, incluidos sus determinantes, es una línea nueva de investigación que se abre a partir de los hallazgos obtenidos en este trabajo. Avanzar en los

esfuerzos futuros de producción de conocimiento en esta línea de investigación tiene relevancia científica y pertinencia práctica en la medida que podría arrojar conocimiento nuevo y útil para el diseño, implementación y monitoreo de la política pública en CTI, en particular, la asociada con los sistemas de estímulos a la investigación.

Finalmente, se espera que el análisis y los resultados de esta investigación contribuyan al debate público sobre el SNI y la importancia de estimular también una agenda de producción de conocimiento propia, contextualizada, de calidad, diversa y plural en México. Conjuntamente con ello, también se espera contribuir a la discusión informada en el ámbito de la política pública de CTI. En particular, sobre las dimensiones y parámetros a re-considerar (criterios de evaluación y su validez) en el re-diseño e implementación de su sistema de reconocimiento y recompensa de la ciencia, para que la producción de conocimiento que se realiza en México pueda contribuir también al alcance de sus objetivos de desarrollo nacional.

## Bibliografía

- Aboites, J. y Díaz, C. (2012) El SNI y la movilidad de los académicos en la globalización. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Alberts, B. (2013) Impact factor distortions. Editorial de la revista *Science*, Vol. 340, 17 de Mayo de 2013. Disponible en [www.sciencemag.org](http://www.sciencemag.org)
- Albornoz, M. (2007) Los problemas de la ciencia y el poder. Revista Iberoamericana Ciencia, tecnología y sociedad, Vol. 3, N°8, Buenos Aires, Abril de 2007
- Álvarez Gallegos, J. (2012) Retos y perspectivas del Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Amaro, M. R; Corona, J.M y Soria, M.L. (2008). Incentivos y colaboraciones universidad-empresa: un estudio en el sector biotecnológico mexicano. MT7, SinncO 2008, México.
- AFC (2011) Du bon usage de la bibliométrie pour l'évaluation individuelle des chercheurs. L'Académie des sciences de l'Institut de France, 2011. Disponible en [http://www.sauvonsluniversite.com/IMG/pdf/avis\\_170111.pdf](http://www.sauvonsluniversite.com/IMG/pdf/avis_170111.pdf)
- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides y R. E. Schumacker (Eds.), *Advanced Structural Equation Modeling: Issues and Techniques* (pp. 243-277). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Arocena, R.; Bortagaray, I. y Sutz, J. (2008) Reforma Universitaria y Desarrollo. Proyecto UniDev: «Universidades para el Desarrollo. El rol cambiante de las Instituciones Académicas en los Sistemas de Innovación y Desarrollo», Montevideo, UniDev-Udelar.
- Arocena, R. y Sutz, J. (2012) Research and innovation policies for social inclusion: An opportunity for developing countries, *Innovation and Development*, Vol. 2, Nro.1, pp. 147-158, abril.
- Asomoza Palacio, R. (2005) El impacto del SNI y su futuro en la ciencia nacional. (Mesa Redonda II). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Asparouhov, T. y Muthén, B. (2010) Weighted Least Squares Estimation with Missing Data, August 14, 2010. Disponible en <https://www.statmodel.com/download/GstrucMissingRevision.pdf>

- Ávalos, I. (1997). "El CONICYT: Casa de pares e impares (o cómo no hay ideas equivocadas sino extemporáneas)", en Sutz, J. (ed.): *Innovación y Desarrollo en América Latina*, CLACSO, Caracas. Editorial Nueva Sociedad - Agencia Española de Cooperación Internacional, 1997, p. 151-162.
- Balankin, A. (2005) Reflexiones sobre el Sistema Nacional de Investigadores. En *Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación*, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Barré, R. (2010) Towards socially robust S&T indicators: indicators are debatable devices, enabling collective learning. *Research Evaluation* 19(3): 227–231.
- Becerra Rodríguez, N. (2014) "Vinculación Universidad-Empresa y productividad científica: el caso de los investigadores mexicanos", Doctorado en Ciencias Sociales, UAM-X
- Benninghoff, M. y Braun, D. (2010). Research Funding, Authority Relations, and Scientific Production in Switzerland. En *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*, eds. Whitley, R., Gläser, J. y Engwall, L. 81–108. New York: Oxford University Press Inc.
- Bensusán, G., Ahumada, e Inclán, D. (2013) Sobre la propuesta de jubilación de los miembros del SNI. FCCyT, Abril de 2013, Ciudad de México.
- Bensusán, G., Gras, N., Inclán, D., Rodríguez, E., Valenti, G. y Varela, G. (2014) Reflexiones sobre la evaluación a los investigadores: una mirada desde diferentes perspectivas. FCCyT, Junio de 2014, Ciudad de México. Disponible en [http://www.foroconsultivo.org.mx/libros\\_editados/evaluacion\\_de\\_la\\_evaluacion\\_subgrupos\\_individuos.pdf](http://www.foroconsultivo.org.mx/libros_editados/evaluacion_de_la_evaluacion_subgrupos_individuos.pdf)
- Bianco, M.; Gras, N. y Sutz, J. (2016) Academic Evaluation: Universal Instrument? Tool for Development? *Minerva* (2016) 54:399–421. DOI 10.1007/s11024-016-9306-9
- Bianco, M.; Gras, N. y Sutz, J. (2014) Reflexiones sobre la práctica de la evaluación académica. En Bianco, M. y Sutz, J. (Coord.) *Análisis y reflexiones sobre 20 años de políticas de investigación en la Universidad de la República: aciertos, dudas y aprendizajes*, CSIC-UdelaR, Trilce, Montevideo.
- Bocco, G., Espejel, I., Hualde, A., Liedo, P., Olivé, L., Reyes, C., Robles, E., Suárez, R. (2014) Evaluación de proyectos multi/inter/transdisciplinarios. Reporte de investigación, Marzo, 2014. FCCyT, Distrito Federal, México.
- Böhme, G., Van den Daele, W., Hohlfeld, R., Krohn, W. y Schäfer, W. (1983) *Finalization in Science: The Social Orientation of Scientific Progress*. Riedel, Dordrecht.

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NYS: John Wiley & Sons.
- Bollen, K. A., y Long, J. S. (Eds.). (1993). *Testing structural equation models*. Newbury Park, CA: Sage.
- Bortagaray, I. y Gras, N. (2014) Science, Technology and Innovation Policies for Inclusive Development: Shifting Trends in South America. En Crespi, G. y Dutrénit, G (Eds.) Science, Technology and Innovation Policies for Inclusive Development: The Latin American Experience, Springer, New York
- Brown, T. A. (2015) Confirmatory Factor Analysis for Applied Research. Second Edition, The Guilford Press, New York
- Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36, 111-150.
- Butler, L. (2003) Modifying publication practices in response to funding formulas. *Research Evaluation*, Vol. 12, Nro. 1, Abril 2003, pp. 39-46.
- Cantú, J.M. (2005) Comentarios alusivos al vigésimo aniversario del SNI. Reflexiones acerca del Sistema Nacional de Investigadores. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Chavarro, D., Tang, P. y Rafols, I. (2017) Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling. *Research Policy* 46 (2017) 1666-1680. Doi: <http://dx.doi.org/10.1016/j.respol.2017.08.002>
- Cimoli, M. y Primi, A. (2008) Technology and intellectual property: a taxonomy of contemporary markets for knowledge and their implications for development. LEM Papers Series 2008/06, March 2008, Pisa, Italy.
- CONACYT (2014) Informe General del estado de la Ciencia, la Tecnología, la Innovación. México 2014, Consejo Nacional de Ciencia y Tecnología, México, D.F.
- Cordero, G.; Galaz Fontes, J.F. y Sevilla, J.J. (2003) *La evaluación de la diversidad en el trabajo académico: Los programas de estímulo de la UABC 1990-2002*. México: Asociación Nacional de Universidades e Instituciones de Educación Superior y Universidad Autónoma de Baja California.
- Dahler-Larsen, P. (2014) Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. *Public Management Review* 16(7): 969–986.
- Das, J., Do, Q-T., Shaines, K. y Srinivasan, S. (2009) U.S. and Them. The Geography of Academic Research. Policy Research Working Paper 5152. The World Bank Development Research

Group Poverty and Inequality Team & Human Development and Public Services Team, December 2009.

- De Ibarrola, M. (2012) Consolidación del Sistema Nacional de Investigadores, profesionalización de la investigación en México. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México
- De Ibarrola, M. (2005) El Sistema Nacional de Investigadores a 20 años de su creación. Borrador para discusión en Mesa de Diálogo sobre deshomologación salarial y carrera académica en México del Segundo Encuentro de Auto-estudio de las Universidades Públicas mexicanas, CIICH/CESU/UNAM, México 10 de febrero de 2005.
- de Jong, S. P. L., van Arensbergen, P., Daemen, F., van der Meulen, B. y van den Besselaar, P. (2011) Evaluation of research in context: an approach and two cases. *Research Evaluation*, 20(1), March 2011, pages 61–72.
- De Rijcke, S., Wouters, P.F., Rushforth, A.D., Franssen, T.P. y Hammarfelt, B. (2016) Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, Vol. 25, Nro. 2, 2016, 161–169 doi:10.1093/reseval/rvv038
- Delanty, G. (1998) The idea of the university in the global era: From knowledge as an end to the end of knowledge? *Social Epistemology: A Journal of Knowledge, Culture and Policy*, 12:1, 3-25
- Díaz-Barriga, A. (2005) El profesor de educación superior frente a las demandas de los nuevos debates educativos. *Perfiles Educativos*, tercera época, año/vol. XXVII, Nro. 108, pp. 9-30
- Didou, S. (2002) Las políticas de educación superior en los institutos tecnológicos federales: una reforma inconclusa. *Revista Mexicana de Investigación Educativa*, enero-abril 2002, Vol. 7, Nro. 14, pp. 51-73
- Didou, S. y Gérard, E. (2010) El Sistema Nacional de Investigadores, veinticinco años después. La comunidad científica, entre distinción e internacionalización. (Primera Edición) ANUIES, México.
- Didou, S. y Gérard, E. (2011) El Sistema Nacional de Investigadores en 2009. ¿Un vector para la internacionalización de las élites científicas? *Perfiles Educativos*, Vol. XXXIII, Nro. 132, 2011. IISUE-UNAM
- Dillon, W., y Goldstein, M. (1984). *Multivariate analysis: Methods and applications*. New York: Wiley.
- DORA (2012) *Declaration on Research Assessment: Putting science into the assessment of research*. San Francisco. Disponible en <http://am.ascb.org/dora/>
- Drucker Colin, R. (2005) El impacto del SNI y su futuro en la ciencia nacional. (Mesa Redonda II). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-

FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017

Drucker Colin, R. (2005b) Palabras de bienvenida. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017

Dutrénit, G., Capdevielle, M., Corona, J.M., Puchet, M., Santiago, F. y Vera-Cruz, A. (2010) El Sistema Nacional de Innovación Mexicano: políticas, desempeño y desafíos. UAM, México.

Dutrénit, G., De Fuentes, C., Santiago, F., Torres, A. y Gras, N. (2013) Innovation and Productivity in the Service Sector: the case of Mexico. Inter-American Development Bank, Discussion Paper, Nro. IDB-DP-293, June 2013

Dutrénit, G. y Sutz, J (2014) Introduction to national innovation systems, social inclusion and development. En Dutrénit, G. y Sutz, J. (Eds.) National Innovation Systems, Social Inclusion and Development: The Latin American Experience, Edward Elgar, Cheltenham

De Fuentes, C., Dutrénit, G., Santiago, F. y Gras, N. (2015) Determinants of Innovation and Productivity in the Service Sector in Mexico. Emerging Markets Finance and Trade, 51:3, 578-592, DOI: 10.1080/1540496X.2015.1026693

Elzinga, A. (1988) The consequences of evaluation for academic research. Science Studies, 1, 5-14.

Enders, C. (2010) Applied Missing Data Analysis. The Guilford Press, New York

Esteinou Madrid, J. (2012) Hacia un nuevo Sistema Nacional de Investigadores que contribuya al desarrollo equilibrado del país. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México

Etzkowitz, H. (1998) The norms of entrepreneurial science: cognitive effects of the new university-industry linkages. Research Policy 27 (1998) 823-833.

FCCyT (2016) El Sistema Nacional de Investigadores en números, FCCyT: CDMX.

FCCyT (2014) Encuesta "Percepciones sobre la evaluación académica" Foro Consultivo Científico y Tecnológico, México.

FCCyT y AMC (2005) Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación. Presentan Academia Mexicana de Ciencias y Foro Consultivo Científico y Tecnológico, Febrero 2005, FCCyT, México, D.F. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017



- Finney, S. y Distefano, C. (2006) Nonnormal and Categorical Data in Structural Equation Modeling. En Hancock, G. y Mueller, R. (Eds.) *Structural Equation Modeling. A second course*, Information Age Publishing, US.
- Flores Valdés, J. (2012) Los orígenes del Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México
- Fortes Besprosvani, M. (2012) El Sistema Nacional de Investigadores como termómetro de la inteligencia colectiva. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México
- Freeman, C. (1987) *Technology Policy and Economic Performance. Lessons from Japan*. Pinter Publisher, Londres.
- Galaz Fontes, J. F. y Gil Antón, M. (2009) La profesión académica en México: Un oficio en proceso de reconfiguración. *Revista Electrónica de Investigación Educativa*, 11 (2). Disponible en: <http://redie.uabc.mx/vol11no2/contenido-galaz2.html>. Consultado el 9 de Septiembre de 2013.
- Galaz Fontes, J.F.; De la Cruz Santana, A.L.; Rodríguez García, R.; Cedillo Nakay, R.A. y Villaseñor Amézquita, M.G. (2012) El académico mexicano miembro del Sistema Nacional de Investigadores: Una primera exploración con base en los resultados de la encuesta "La Reconfiguración de la Profesión Académica en México". En *El Futuro de la Profesión Académica. Desafíos para los países emergentes*. Eds. Norberto Fernández Lamarra y Mónica Marquina, Buenos Aires: Universidad Nacional de Tres de Febrero, 2012. 344-355. Disponible en <http://works.bepress.com/galazfontes/23>
- Galaz Fontes, J.F. (2008) Los académicos mexicanos a principios del siglo XXI: Una primera exploración sobre quiénes son y cómo perciben su trabajo, sus instituciones y algunas políticas públicas. Documento de trabajo en la XXXII Sesión Ordinaria del Consejo de Universidades Públicas e Instituciones Afines (CUPIA) de la Asociación Nacional de Universidades e Instituciones de Educación Superior. Villahermosa, Tabasco, 21 de Noviembre de 2008. Disponible en <http://works.bepress.com/galazfontes/21>
- Galaz Fontes, J.F.; Padilla, L.E.; Gil, M. y Sevilla, J.J. (2008) Los dilemas del profesorado en la educación superior Mexicana. *Calidad en la Educación*, número 28, julio 2008, pp. 53-69
- Garfield, E (1955) Citation indexes for science. A new dimension in documentation through association of ideas. *Science*, Vol.122, 108-111, July, 1955.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. y Trow, M. (1994) *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. SAGE, London.

- Gil Antón, M. (2002) Amor de Ciudad Grande: una visión general del espacio para el trabajo académico en México. *Sociológica*, año 17, número 49, mayo-agosto de 2002, pp. 93-130.
- Gläser, J. (2012) How does Governance change research content? On the possibility of a sociological middle-range theory linking science policy to the sociology of scientific knowledge. TUTS-WP-1-2012, <https://www.ts.tu-berlin.de/fileadmin/fg226/TUTS/TUTS-WP-1-2012.pdf>
- Gläser, J. y Laudel, G. (2007b) Evaluation Without Evaluators: The Impact of Funding Formulae on Australian University Research. En Whitley, R. y Gläser, J. (Eds.), *The changing governance of the sciences. The Advent of Research Evaluation Systems*. Springer, Netherlands
- Gläser, J. y Laudel, G. (2007a) The social construction of bibliometric evaluation. En Whitley, R. y Gläser, J. (Eds.), *The changing governance of the sciences. The Advent of Research Evaluation Systems*. Springer, Netherlands.
- Gläser, J. y Laudel, G. (2016). Governing Science. *European Journal of Sociology*, 57, pp. 117-168. doi:10.1017/S0003975616000047
- Gläser, J. (2017) A fight on epistemological quicksand: Comment on the dispute between van den Besselaar et al. and Butler. *Journal of Informetrics* 11 (2017) 927–932.
- Gonzalez-Brambila, C. (2005) Exploring Academic Scientific Productivity for the Design of Public Policies. Tesis doctorado. Carnegie Mellon University, Pittsburgh, Pennsylvania.
- Gonzalez-Brambila, C. y Veloso, F.M. (2007) The determinants of research output and impact: A study of Mexican researchers. *Research Policy* 36 (2007) 1035–1051. doi:10.1016/j.respol.2007.03.005
- González Valenzuela, J. (2005) Sobre el Sistema Nacional de Investigadores. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Grediaga Kuri, M.R. (1998) Cambios en el sistema de recompensa y reconocimiento en la profesión académica en México. *Revista de la Educación Superior*, vol. XXVII (4), núm. 108, octubre-diciembre, pp. 125-205.
- Hammarfelt, B. y De Rijcke, S. (2014) Accountability in context: effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University. *Research Evaluation* (2014) pp. 1–15. doi:10.1093/reseval/rvu029
- Hair, J. F.; Black, W. C.; Babin, B. J. y Anderson, R. E. (2014) *Multivariate Data Analysis*, Seventh Edition, Pearson Education Limited, London.

- Hemlin, S. y Barlebo Rasmussen, S. (2006) *The Shift in Academic Quality Control Science, Technology, & Human Values*, Vol. 31, No. 2 (Mar., 2006), pp. 173-198
- Heras, L. (2005) *La política de Educación Superior en México: los programas de estímulos a profesores e investigadores*. EDUCERE Foro universitario Artículos arbitrados, Año 9, Nro. 29, Abril-Mayo-Junio, 2005, pp. 207-215.
- Hernández-Calderón, I. (2005) *El SNI veinte años después. Recuento de algunas virtudes y defectos de nuestra propia comunidad. Reflexiones acerca del Sistema Nacional de Investigadores (III)*. En *Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación*, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Hess, D. (2007) *Alternative Pathways in Science and Industry. Activism, Innovation, and the Environment in an Era of Globalization*. Cambridge: The MIT Press
- Hessels, L. y van Lente, H. (2008) *Re-thinking new knowledge production: A literature review and a research agenda*. *Research Policy* 37 (2008) 740–760.
- Hicks, D. (2004) *The Four Literatures of Social Science Handbook of Quantitative Science and Technology Research*, ed. Henk Moed, Kluwer Academic.
- Hicks, D. (2006) *The Dangers of Partial Bibliometric Evaluation in the Social Sciences*. *Economia Política*, XXIII, n. 2 (August, 2006) pp. 145-162
- Hicks, D. (2012) *Performance-based university research funding systems*. *Research Policy* 41 (2012) 251– 261. doi:10.1016/j.respol.2011.09.007
- Hicks, D. (2013) *One size doesn't fit all: On the co-evolution of national evaluation systems and social science publishing*; en *Confero. Essays on Education Philosophy and Politics. Managing by measuring: Academic knowledge production under the ranks Volume 1, Number 1, March 2013*. Disponible en [www.confero.ep.liu.se](http://www.confero.ep.liu.se)
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. y Rafols, I. (2015). *The Leiden Manifesto for research metrics*. *Nature* 520 (April): 429-431.
- Hirsch, J.E. (2005) *An index to quantify an individual's scientific research output*. *PNAS*, November 15, 2005, Vol. 102, Nro. 46, pp. 16569-16572
- Hood, C. (1991) *A public management for all seasons?*. *Public Administration*, Vol. 69, Issue 1, pp. 3-19.
- Ibarra, E. (1998) *La Universidad en México hoy: gubernamentalidad y modernización*. Tesis de Doctorado en Sociología, Facultad de Ciencias Políticas y Sociales, Universidad Nacional Autónoma de México, México.

- Ibarra, E. (1999) Evaluación, productividad y conocimiento: barreras institucionales al desarrollo académico. Ponencia en el Primer Congreso de Ciencias Sociales, México.
- Irvine, J. y Martin, B.R. (1984) *Foresight in Science: Picking the Winners*. Frances Pinter, London.
- Izquierdo, I. (2006) La formación de investigadores y el ejercicio profesional de la investigación: el caso de los ingenieros y físicos de la UAEM. *Revista de Educación Superior*, Vol. XXXV (4), Nro. 140, Octubre-Diciembre de 2006, pp. 7-28.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Kish, L. (2004, [1987]) *Statistical Design for Research*. Wiley-Interscience, New Jersey
- Kline, R. B. (2011) *Principles and Practice of Structural Equation Modeling*. Third Edition, The Guilford Press, New York
- Kostoff, R. N. (1997) Use and Misuse of Metrics in Research Evaluation. *Science and Engineering Ethics*, Vol. 3, Issue 2, pp. 109-120.
- Leisyte, L. y Westerheijden, D.F. (2014) Research evaluation and its implications for academic research in the United Kingdom and the Netherlands. Discussion Paper Nro. 1-2014. Technische Universität Dortmund
- Lundvall, B. Å. (1992) *National System of Innovation —Toward a Theory of Innovation and Innovative Learning*, Londres, Pinter.
- MacCallum, R. C., Browne, M. W., y Suagwara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-149.
- Malo Álvarez, S. (2005) Creación, evolución y evaluación de los investigadores (Mesa Redonda I). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Mårtensson, P., Fors, U., Wallin, S-B., Zander, U. y Nilsson, G (2016) Evaluating research: A multidisciplinary approach to assessing research practice and quality. *Research Policy* 45 (2016), pp. 593–603.
- Martin, B. (2016) Editors' JIF-boosting stratagems – Which are appropriate and which not? (Editorial). *Research Policy* 45 (2016) 1–7
- Martin, B. y Whitley, R. (2010) The UK Research Assessment Exercise: A Case of Regulatory Capture? En *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences*

*and their Consequences for Intellectual Innovation*, ed. R. Whitley, J. Gläser and L. Engwall. New York: Oxford University Press Inc.

- Martínez Assad, C. (2005) Parece que fue ayer. Reflexiones acerca del Sistema Nacional de Investigadores. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Martínez Palomo, A. (2005) El impacto del SNI y su futuro en la ciencia nacional. (Mesa Redonda II). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Mateos Gómez, J.L. (2005) Reflexiones acerca del Sistema Nacional de Investigadores. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Meier, F. y Schimank, U. (2010) Mission Now Possible: Profile Building and Leadership in German Universities. En *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*, ed. R. Whitley, J. Gläser and L. Engwall. New York: Oxford University Press Inc.
- Membrillo Hernández, J. (2005) La importancia de pertenecer al SNI para un investigador joven (Conferencias). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Merton, R. (1936) The Unanticipated Consequences of Purposive Social Action, *American Sociological Review*, 1 (6): 894-904.
- Merton, R. (1942) The Normative Structure of Science, re-impreso en Merton, *The Sociology of Science*, Norman W. Storer, ed. Chicago: University of Chicago Press, 1973.
- Merton, R. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property, *ISIS*, 79(4): 606-623.
- Moed, H. F.; (2009) New developments in the use of citation analysis in research evaluation *Variation-Scientometrics*, 2009, 57, 13–18.
- Mulaik, Stanley A. (2010). *Foundations of Factor Analysis*. Boca Raton, London, New York: CRC Press, A Taylor & Francis Group
- Mulás del Pozo, P. (2005) Comentarios sobre el XX aniversario del Sistema Nacional de Investigadores. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de

su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017

- Müller, R. y De Rijcke, S. (2017) Exploring the epistemic impacts of academic performance indicators in the life sciences. *Research Evaluation*, 26(3), 2017, 157–168 doi: 10.1093/reseval/rvx023
- Muthén, B. (1993) Goodness of fit with categorical and other non-normal variables. En Bollen, K.A. y Long, J.S. (Eds.) *Testing Structural Equation Models*, Newbury Park, CA: Sage (#45)
- Muthén, B. (1998-2004). *Mplus Technical Appendices*. Los Angeles, CA: Muthén & Muthén
- Muthén, B. y Asparouhov, T. (2002) Latent Variable Analysis With Categorical Outcomes: Multiple-Group And Growth Modeling In Mplus. *Mplus Web Notes*: No. 4, Version 5, December 9, 2002.
- Muthén, B., du Toit, S. y Spisic, D. (1997) *Robust Inference using Weighted Least Squares and Quadratic Estimating Equations in Latent Variable Modeling with Categorical and Continuous Outcomes*. November 18, 1997. Disponible en [https://www.statmodel.com/download/Article\\_075.pdf](https://www.statmodel.com/download/Article_075.pdf)
- Muthén, L., y Muthén, B. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 4, 599-620.
- Muthén, L.K. y Muthén, B.O. (1998-2015) *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén
- Nowotny, H., Scott, P. y Gibbons, M. (2001) *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Polity Press, Cambridge.
- Nunnally, J. C., y Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw Hill.
- OECD (2017) *Main Science and Technology Indicators*, Vol. 2016, Issue 2.
- Ocampo, E. y Rueda, J. (2015) El Sistema Nacional de Investigadores en la Universidad Veracruzana: análisis exploratorio de cómo se experimenta el reconocimiento académico. *Revista Interamericana de Educación de Adultos*, año 37, nro. 1, enero-junio de 2015, pp. 65-85
- Padilla, L. (2010) El Académico Mexicano Miembro del Sistema Nacional de Investigadores: Su contexto institucional, uso del tiempo, productividad académica, e implicaciones salariales. I Congreso de los Miembros del Sistema Nacional de Investigadores Querétaro, Querétaro, Mayo 5-8 de 2010-04-27
- PECiTI (2014-2018) *Programa Especial de Ciencia, Tecnología e Innovación 2014-2018*, Consejo Nacional de Ciencia y Tecnología (CONACYT) Gobierno Federal, México

- PECiTI (2008-2012) *Programa Especial de Ciencia, Tecnología e Innovación 2008-2012*, Consejo Nacional de Ciencia y Tecnología (CONACYT), México. Publicado en el Diario Oficial de la Federación como Programa Especial de Ciencia y Tecnología 2008-2012.
- PECyT (2001-2006) Programa Especial de Ciencia y Tecnología 2001-2006, Consejo Nacional de Ciencia y Tecnología (CONACYT), México.
- Pendlebury, D. A. (2009) The use and misuse of journal metrics and other citation indicators *Variation in Scientometrics*, 2009, 57, 1–11.
- Peña, A. (2005b) ¿Es el SNI un sistema modelo de evaluación de los investigadores en México? En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Peña Díaz, A. (2005) Creación, evolución y evaluación de los investigadores (Mesa Redonda I). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Pérez Tamayo, R. (2005) Reflexiones sobre el Sistema Nacional de Investigadores. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Pescador Osuna, J. A. (2012) Modificaciones al Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Plan Nacional de Desarrollo (2013-2018), Gobierno de la República. Disponible en [http://www.snieg.mx/contenidos/espanol/normatividad/MarcoJuridico/PND\\_2013-2018.pdf](http://www.snieg.mx/contenidos/espanol/normatividad/MarcoJuridico/PND_2013-2018.pdf)
- Polanyi, M. (1962) *The Republic of Science: Its Political and Economic Theory*. Minerva, 1, pp. 54-74.
- Porter, V.L. (2012) La etapa de consolidación del Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Power, M. (1996) *The Audit Explosion*. Demos, 1996, London
- Pradilla Cobos, E. (2012) El Sistema Nacional de Investigadores y las condiciones de vida de los investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.

- Rafols, I.; Leydesdorff, L.; O'Hare, A.; Nightingale, P. y Stirling, A. (2012) How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. *Research Policy* 41 (2012) 1262– 1282
- Ramírez Amador, V. A (2012) Desafíos del Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Regeer, B., Hoes, A., van Amstel, M., Caron-Flinterman Saane, F. y Bunders, J. (2009) Six Guiding Principles for Evaluating Mode-2 Strategies for Sustainable Development. *American Journal of Evaluation* 30: 515–537.
- Reglamentos del SNI 2006, 2008, 2012 y 2017 Disponibles en DOF 26/09/2006; 21/03/2008, 26/12/2012 y 27/01/2017 respectivamente
- Reséndiz Núñez, D. (2005) Creación, evolución y evaluación de los investigadores (Mesa Redonda I). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Reyes, G. y Suriñach, J. (2015) Análisis sobre la evolución del Sistema Nacional de Investigadores (SNI) de México. *Investigación Administrativa*, Nro. 115, enero-junio, 2015, pp. 55-69
- Reyes, G. y Suriñach, J. (2012) SNI Internal Evaluations: Coherence or Coincidences. *Secuencia*, Nro. 83, mayo-agosto 2012, pp. 179-217
- RICYT Red de Indicadores de Ciencia y Tecnología.
- Rip, A. (2004) Strategic research, post-modern universities and research training. *Higher Education Policy* 17(2), 153–166.
- Rivera-Huerta, R., Dutrénit, G., Ekboir, J.M., Sampedro, J.L. y Vera-Cruz, A.O. (2011) Do linkages between farmers and academic researchers influence researcher productivity? The Mexican case. *Research Policy* 40 (2011) 932– 942.
- Rosales, A.M. (2012) La consolidación del Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Rothwell, R. (1994) Industrial innovation: success, strategy, trends, en Dodgson, M. y Rothwell, R. (Eds.), *The handbook of industrial innovation*, Edward Elgar, Aldershot.
- Rothwell, R. y Zegveld, W. (1985) *Reindustrialization and technology*, Longman, US.
- Rubin, D. B. (1976) Inference and missing data. *Biometrika*, 63, pp. 581–592.



- Ruiz Gutiérrez, R. (2012) El Sistema Nacional de Investigadores. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México.
- Ruiz Herrera, J. (2005) Creación, evolución y evaluación de los investigadores (Mesa Redonda I). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Russell, J. y Liberman, S. (2002) Desarrollo de las bases de un modelo de comunicación de la producción científica de la Universidad Nacional Autónoma de México. *Revista Española de Documentación Científica*, 25, 4, 2002, pp. 361-370
- Sahel, J.A. (2011) Quality versus quantity: assessing individual research performance *Sci Transl Med* 25 May 2011: Vol. 3, Issue 84.
- Sarukhán Kermez, J. (2005) Creación, evolución y evaluación de los investigadores (Mesa Redonda I). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Sen, A. (2000) *Social exclusion: Concept, application and scrutiny*, Manila, Asian Development Bank.
- Serra Puche, M.C (2005) El impacto del SNI y su futuro en la ciencia nacional. (Mesa Redonda II). En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Slaughter, S. y Leslie, L.L. (1997) *Academic Capitalism: Politics, Policies, and the Entrepreneurial University*. The John Hopkins University Press, Baltimore
- Spaapen, J., Dijstelbloem, H. y Wamelink, F. (2007) *Evaluating Research in Context: A Method for Comprehensive Assessment*. Consultative Committee of Sector Councils for Research and Development (COS), The Netherlands; 2nd edition (2007).
- Tamez, S. y Pérez, J. (2009) El trabajador universitario: entre el malestar y la lucha. *Educ. Soc.*, Campinas, Vol. 30, Nro. 107, pp. 373-387
- Terrones Maldonado, H. (2005) El SNI, la repatriación y la descentralización: experiencias de un joven investigador. En Una reflexión sobre el Sistema Nacional de Investigadores a 20 años de su creación, AMC-FCCyT, Febrero, 2005. Disponible en [www.foroconsultivo.org.mx/libros\\_editados/20\\_sni.pdf](http://www.foroconsultivo.org.mx/libros_editados/20_sni.pdf). Fecha de consulta 31/07/2017
- Valenti, G. (2012) El Sistema Nacional de Investigadores en la conformación de la comunidad científica en México. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México

- Valenti, G. (2011) Construyendo puentes entre el Capital Humano y el Sistema de Innovación. FLACSO México, Ciudad de México.
- Valenti, G., Casalet, M.; Gil Antón, M.; González Brambila, C.; Hualde Alfaro, A.; Varela, G. y Villavicencia, D. (2013) Propuestas para contribuir al diseño del PECiTI 2012-2037. FCCyT, Ciudad de México.
- van Dalen, H.P y Henkens, K. (2012) Intended and Unintended Consequences a Publish-or-Perish Culture: A Worldwide Survey CentER, Discussion Paper, No. 2012-003, Enero 2012
- van der Most, F. (2010) Use and non-use of research evaluation: A literature review Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, Working Paper no. 2010/16.
- Vargas, D. (2016) Diseño de Estudios y Análisis de Datos Longitudinales. Curso 27-30 de junio de 2016, UNAM, México.
- Vega y León, S. (coord.) (2012) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México
- Vega y León, S. (2012) El Sistema Nacional de Investigadores y su impacto en el Sistema de Educación Superior. En Vega y León, S. (coord.) *Sistema nacional de investigadores. Retos y perspectivas de la ciencia en México*. UAM-X, México
- Vessuri, H., Guédon, J.C., Cetto, A.M., 2014. Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Curr. Sociol.* 62 (5), 647–665. <http://dx.doi.org/10.1177/0011392113512839>.
- Wang, J., Veugelers, R. y Stephan, P. (2017) Bias against novelty in science: A cautionary tale for users of bibliometric indicators *Research Policy* (2017), <http://dx.doi.org/10.1016/j.respol.2017.06.006>
- Weber, M. (1992) *Economía y Sociedad*. Fondo de Cultura Económica, México.
- Whitley, R. (2007) “Changing Governance of the Public Sciences: The Consequences of Establishing Research Evaluation Systems for Knowledge Production in Different Countries and Scientific Fields”, en Whitley, R. y Gläser, J. (Eds.), *The changing governance of the sciences. The Advent of Research Evaluation Systems*. Springer, Netherlands.
- Whitley, R. 2010. Reconfiguring the Public Sciences: The Impact of Governance Changes on Authority and Innovation in Public Science Systems. In *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*, ed. R. Whitley, J. Gläser and L. Engwall. New York: Oxford University Press Inc.

Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany: State University of New York Press.

Ziman, J. (2000) *Real Science. What it is, and what it means*. Cambridge University Press, Cambridge, United Kingdom

## ANEXO 1: Análisis Factorial Exploratorio. Salida M-PLUS

Mplus VERSION 7.4  
MUTHEN & MUTHEN  
10/03/2016 6:14 PM

### INPUT INSTRUCTIONS

TITLE: Análisis Factorial Exploratorio

DATA: FILE IS enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat;  
FORMAT IS FREE;

VARIABLE: NAMES ARE id\_user edad form niv\_sni cand niv\_I niv\_II niv\_III area\_sni  
f\_m\_ct b\_q m\_cs h\_cc cs b\_ca ing gen\_sni mujer hombre tip\_inst cpi\_e\_n cpi\_cs\_h  
cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ot\_inst ori inst\_t\_d\_k use\_ori  
p\_1\_a\_r p\_1\_b\_r p\_1\_ab\_r p\_2\_r p\_2\_a\_r p\_2\_b\_r p\_2\_c\_r p\_2\_d\_r p\_3\_r p\_3\_r\_d p\_5\_r  
p\_5\_r\_d p\_6\_a\_r p\_6\_b\_r p\_6\_c\_r p\_6\_d\_r p\_6\_e\_r p\_7\_r p\_8\_r p\_9\_r p\_9\_a\_r p\_9\_b\_r  
p\_9\_c\_r p\_9\_d\_r p\_9\_e\_r p\_10\_r p\_11\_r p\_11\_r\_d p\_13\_r p\_14\_a\_r p\_14\_b\_r p\_14\_c\_r  
p\_14\_d\_r p\_14\_e\_r p\_14\_f\_r p\_14\_g\_r p\_14\_h\_r p\_15\_r p\_15\_a\_r p\_15\_b\_r p\_15\_c\_r  
p\_16\_a\_r p\_16\_b\_r p\_16\_c\_r p\_16\_d\_r p\_16\_e\_r p\_16\_f\_r p\_16\_g\_r p\_17\_a\_r p\_17\_b\_r  
p\_17\_c\_r p\_17\_d\_r p\_17\_e\_r p\_17\_f\_r p\_18\_a\_r p\_18\_b\_r p\_18\_c\_r p\_18\_d\_r p\_18\_e\_r  
p\_18\_f\_r p\_18\_g\_r p\_18\_h\_r p\_18\_i\_r p\_18\_j\_r p\_18\_k\_r p\_19\_r p\_20\_a\_r p\_20\_b\_r  
p\_20\_c\_r p\_20\_d\_r p\_20\_e\_r p\_20\_f\_r p\_20\_g\_r p\_20\_h\_r p\_20\_i\_r p\_20\_j\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_m\_r p\_39\_r ;

USEVARIABLES ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r  
p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r  
!p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r  
p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ; !edad !cand !niv\_I !niv\_II !niv\_III !f\_m\_ct  
!b\_q !m\_cs !h\_cc !cs !b\_ca !ing !mujer !hombre !cpi\_e\_n !cpi\_cs\_h !cpi\_dt\_s !cpi\_sec  
!gobierno !inst\_n\_s !inst\_ut !univ ; !ot\_inst !t\_d\_k !use\_ori ;

MISSING ARE p\_1\_a\_r (2) p\_1\_b\_r (2) p\_1\_ab\_r (2) p\_2\_r (5) p\_2\_a\_r (2) p\_2\_b\_r (2)  
p\_2\_c\_r (2) p\_2\_d\_r (2) p\_3\_r\_d (4 5) p\_5\_r\_d (5) p\_6\_a\_r (2) p\_6\_b\_r (2) p\_6\_c\_r  
(2) p\_6\_d\_r (2) p\_6\_e\_r (2) p\_7\_r (2) p\_8\_r (2) p\_9\_r (6) p\_9\_a\_r (2) p\_9\_b\_r (2)  
p\_9\_c\_r (2) p\_9\_d\_r (2) p\_9\_e\_r (2) p\_10\_r (2) p\_11\_r\_d (5) p\_15\_r (4) p\_15\_a\_r (2)  
p\_15\_b\_r (2) p\_15\_c\_r (2) p\_16\_a\_r (2) p\_16\_b\_r (2) p\_16\_c\_r (2) p\_16\_d\_r (2)  
p\_16\_e\_r (2) p\_16\_f\_r (2) p\_16\_g\_r (2) p\_17\_a\_r (2) p\_17\_b\_r (2) p\_17\_c\_r (2)  
p\_17\_d\_r (2) p\_17\_e\_r (2) p\_17\_f\_r (2) p\_19\_r (2) p\_20\_a\_r (2) p\_20\_b\_r (2) p\_20\_c\_r  
(2) p\_20\_d\_r (2) p\_20\_e\_r (2) p\_20\_f\_r (2) p\_20\_g\_r (2) p\_20\_h\_r (2) p\_20\_i\_r (2)  
p\_20\_j\_r (2) p\_20\_k\_r (2) p\_20\_l\_r (2) p\_20\_m\_r (2) . ;

CATEGORICAL ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r  
p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r  
!p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r  
p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ;

AUXILIARY ARE id\_user ;

ANALYSIS:  
TYPE = EFA 1 7;  
OUTPUT: SAMPSTAT ;

\*\*\* WARNING

Data set contains cases with missing on all variables. These cases were not included in the analysis. Number of cases with missing on all variables: 2  
 1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

Análisis Factorial Exploratorio

SUMMARY OF ANALYSIS

Number of groups 1  
 Number of observations 8053

Number of dependent variables 33  
 Number of independent variables 0  
 Number of continuous latent variables 0

Observed dependent variables

Binary and ordered categorical (ordinal)

P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R	P_18_B_R
P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R	P_1_A_R	P_1_B_R
P_2_A_R	P_2_C_R	P_5_R_D	P_7_R	P_8_R	P_9_A_R
P_9_B_R	P_9_D_R	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R
P_17_A_R	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_20_I_R	P_20_E_R	P_20_J_R			

Observed auxiliary variables

ID\_USER

Estimator	WLSMV
Rotation	GEOMIN
Row standardization	CORRELATION
Type of rotation	OBLIQUE
Epsilon value	Varies
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.100D-03
Optimization Specifications for the Exploratory Factor Analysis	
Rotation Algorithm	
Number of random starts	30
Maximum number of iterations	10000
Derivative convergence criterion	0.100D-04

Input data file(s)  
 enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat

Input data format FREE

SUMMARY OF DATA

Number of missing data patterns 1292

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

Covariance Coverage

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	0.972				
P_18_F_R	0.965	0.968			
P_18_G_R	0.969	0.966	0.972		
P_18_I_R	0.969	0.966	0.971	0.973	
P_18_J_R	0.965	0.963	0.967	0.969	0.970
P_18_B_R	0.970	0.967	0.971	0.971	0.968

P_18_D_R	0.968	0.965	0.970	0.970	0.966
P_18_E_R	0.968	0.966	0.970	0.971	0.967
P_18_H_R	0.967	0.965	0.969	0.970	0.966
P_18_K_R	0.968	0.965	0.969	0.971	0.967
P_1_A_R	0.834	0.831	0.834	0.835	0.832
P_1_B_R	0.849	0.846	0.850	0.851	0.847
P_2_A_R	0.948	0.944	0.949	0.950	0.946
P_2_C_R	0.948	0.944	0.949	0.950	0.946
P_5_R_D	0.849	0.846	0.850	0.851	0.847
P_7_R	0.670	0.667	0.670	0.671	0.668
P_8_R	0.614	0.612	0.614	0.615	0.613
P_9_A_R	0.613	0.612	0.614	0.615	0.613
P_9_B_R	0.613	0.612	0.614	0.615	0.613
P_9_D_R	0.613	0.612	0.614	0.615	0.613
P_3_R_D	0.904	0.901	0.905	0.906	0.902
P_10_R	0.834	0.830	0.834	0.834	0.831
P_11_R_D	0.864	0.860	0.865	0.865	0.862
P_9_C_R	0.613	0.612	0.614	0.615	0.613
P_17_A_R	0.910	0.907	0.911	0.912	0.908
P_17_B_R	0.910	0.907	0.911	0.912	0.908
P_17_D_R	0.910	0.907	0.911	0.912	0.908
P_20_G_R	0.760	0.757	0.761	0.762	0.760
P_20_K_R	0.767	0.765	0.768	0.769	0.766
P_20_L_R	0.729	0.726	0.730	0.730	0.728
P_20_I_R	0.778	0.775	0.779	0.780	0.777
P_20_E_R	0.768	0.766	0.770	0.770	0.768
P_20_J_R	0.728	0.725	0.729	0.730	0.728

Covariance Coverage

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_B_R	0.975				
P_18_D_R	0.970	0.972			
P_18_E_R	0.971	0.969	0.973		
P_18_H_R	0.970	0.968	0.969	0.971	
P_18_K_R	0.970	0.969	0.969	0.969	0.972
P_1_A_R	0.836	0.834	0.834	0.833	0.834
P_1_B_R	0.851	0.849	0.850	0.849	0.849
P_2_A_R	0.951	0.948	0.949	0.948	0.948
P_2_C_R	0.951	0.948	0.949	0.948	0.948
P_5_R_D	0.852	0.849	0.850	0.849	0.849
P_7_R	0.672	0.670	0.670	0.669	0.670
P_8_R	0.615	0.614	0.615	0.614	0.614
P_9_A_R	0.615	0.614	0.615	0.614	0.614
P_9_B_R	0.615	0.614	0.615	0.614	0.614
P_9_D_R	0.615	0.614	0.615	0.614	0.614
P_3_R_D	0.907	0.904	0.905	0.904	0.905
P_10_R	0.835	0.833	0.834	0.833	0.833
P_11_R_D	0.866	0.864	0.865	0.864	0.864
P_9_C_R	0.615	0.614	0.615	0.614	0.614
P_17_A_R	0.913	0.910	0.911	0.910	0.910
P_17_B_R	0.913	0.910	0.911	0.910	0.910
P_17_D_R	0.913	0.910	0.911	0.910	0.910
P_20_G_R	0.762	0.761	0.761	0.760	0.761
P_20_K_R	0.769	0.768	0.768	0.768	0.768
P_20_L_R	0.730	0.730	0.730	0.729	0.730
P_20_I_R	0.780	0.779	0.779	0.778	0.779
P_20_E_R	0.770	0.769	0.770	0.769	0.770
P_20_J_R	0.730	0.729	0.729	0.728	0.729

Covariance Coverage

	<u>P_1_A_R</u>	<u>P_1_B_R</u>	<u>P_2_A_R</u>	<u>P_2_C_R</u>	<u>P_5_R_D</u>
P_1_A_R	0.856				
P_1_B_R	0.752	0.873			
P_2_A_R	0.855	0.871	0.975		
P_2_C_R	0.855	0.871	0.975	0.975	
P_5_R_D	0.750	0.775	0.853	0.853	0.871
P_7_R	0.596	0.615	0.675	0.675	0.657
P_8_R	0.544	0.569	0.618	0.618	0.606
P_9_A_R	0.543	0.569	0.618	0.618	0.606
P_9_B_R	0.543	0.569	0.618	0.618	0.606
P_9_D_R	0.543	0.569	0.618	0.618	0.606
P_3_R_D	0.796	0.840	0.911	0.911	0.820
P_10_R	0.733	0.751	0.832	0.832	0.756
P_11_R_D	0.757	0.774	0.860	0.860	0.781
P_9_C_R	0.543	0.569	0.618	0.618	0.606
P_17_A_R	0.790	0.806	0.898	0.898	0.804
P_17_B_R	0.790	0.806	0.898	0.898	0.804
P_17_D_R	0.790	0.806	0.898	0.898	0.804
P_20_G_R	0.666	0.683	0.751	0.751	0.685
P_20_K_R	0.669	0.689	0.758	0.758	0.691
P_20_L_R	0.639	0.657	0.722	0.722	0.658
P_20_I_R	0.678	0.699	0.769	0.769	0.702
P_20_E_R	0.671	0.691	0.760	0.760	0.693
P_20_J_R	0.636	0.658	0.719	0.719	0.668

Covariance Coverage

	<u>P_7_R</u>	<u>P_8_R</u>	<u>P_9_A_R</u>	<u>P_9_B_R</u>	<u>P_9_D_R</u>
P_7_R	0.685				
P_8_R	0.552	0.628			
P_9_A_R	0.552	0.628	0.628		
P_9_B_R	0.552	0.628	0.628	0.628	
P_9_D_R	0.552	0.628	0.628	0.628	0.628
P_3_R_D	0.647	0.598	0.597	0.597	0.597
P_10_R	0.605	0.560	0.560	0.560	0.560
P_11_R_D	0.621	0.570	0.570	0.570	0.570
P_9_C_R	0.552	0.628	0.628	0.628	0.628
P_17_A_R	0.634	0.581	0.581	0.581	0.581
P_17_B_R	0.634	0.581	0.581	0.581	0.581
P_17_D_R	0.634	0.581	0.581	0.581	0.581
P_20_G_R	0.556	0.520	0.520	0.520	0.520
P_20_K_R	0.560	0.527	0.526	0.526	0.526
P_20_L_R	0.539	0.505	0.505	0.505	0.505
P_20_I_R	0.569	0.534	0.534	0.534	0.534
P_20_E_R	0.565	0.529	0.528	0.528	0.528
P_20_J_R	0.548	0.518	0.518	0.518	0.518

Covariance Coverage

	<u>P_3_R_D</u>	<u>P_10_R</u>	<u>P_11_R_D</u>	<u>P_9_C_R</u>	<u>P_17_A_R</u>
P_3_R_D	0.930				
P_10_R	0.793	0.848			
P_11_R_D	0.822	0.830	0.879		
P_9_C_R	0.597	0.560	0.570	0.628	
P_17_A_R	0.856	0.801	0.829	0.581	0.919
P_17_B_R	0.856	0.801	0.829	0.581	0.919
P_17_D_R	0.856	0.801	0.829	0.581	0.919
P_20_G_R	0.722	0.683	0.699	0.520	0.724
P_20_K_R	0.728	0.690	0.707	0.526	0.729
P_20_L_R	0.693	0.655	0.672	0.505	0.693

P_20_I_R	0.738	0.697	0.715	0.534	0.739
P_20_E_R	0.729	0.691	0.707	0.528	0.730
P_20_J_R	0.693	0.657	0.671	0.518	0.693

Covariance Coverage

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_B_R	0.919				
P_17_D_R	0.919	0.919			
P_20_G_R	0.724	0.724	0.764		
P_20_K_R	0.729	0.729	0.737	0.771	
P_20_L_R	0.693	0.693	0.704	0.709	0.733
P_20_I_R	0.739	0.739	0.741	0.750	0.714
P_20_E_R	0.730	0.730	0.738	0.742	0.708
P_20_J_R	0.693	0.693	0.701	0.710	0.677

Covariance Coverage

	P_20_I_R	P_20_E_R	P_20_J_R
P_20_I_R	0.782		
P_20_E_R	0.751	0.773	
P_20_J_R	0.717	0.706	0.732

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

P_18_C_R		
Category 1	0.011	83.000
Category 2	0.041	317.000
Category 3	0.180	1405.000
Category 4	0.467	3658.000
Category 5	0.302	2362.000
P_18_F_R		
Category 1	0.025	193.000
Category 2	0.088	690.000
Category 3	0.242	1884.000
Category 4	0.452	3527.000
Category 5	0.193	1503.000
P_18_G_R		
Category 1	0.134	1047.000
Category 2	0.265	2073.000
Category 3	0.321	2511.000
Category 4	0.226	1768.000
Category 5	0.055	432.000
P_18_I_R		
Category 1	0.048	376.000
Category 2	0.105	825.000
Category 3	0.223	1751.000
Category 4	0.373	2925.000
Category 5	0.250	1962.000
P_18_J_R		
Category 1	0.060	472.000
Category 2	0.175	1365.000
Category 3	0.315	2462.000
Category 4	0.330	2574.000
Category 5	0.120	936.000
P_18_B_R		
Category 1	0.031	240.000
Category 2	0.069	538.000
Category 3	0.198	1550.000
Category 4	0.363	2851.000



Category 5	0.340	2669.000
P_18_D_R		
Category 1	0.020	155.000
Category 2	0.052	410.000
Category 3	0.154	1206.000
Category 4	0.364	2850.000
Category 5	0.410	3206.000
P_18_E_R		
Category 1	0.008	61.000
Category 2	0.027	213.000
Category 3	0.122	956.000
Category 4	0.367	2877.000
Category 5	0.476	3725.000
P_18_H_R		
Category 1	0.016	125.000
Category 2	0.055	427.000
Category 3	0.165	1293.000
Category 4	0.380	2973.000
Category 5	0.384	3005.000
P_18_K_R		
Category 1	0.066	518.000
Category 2	0.146	1141.000
Category 3	0.285	2228.000
Category 4	0.334	2613.000
Category 5	0.170	1327.000
P_1_A_R		
Category 1	0.174	1202.000
Category 2	0.826	5692.000
P_1_B_R		
Category 1	0.237	1669.000
Category 2	0.763	5361.000
P_2_A_R		
Category 1	0.027	213.000
Category 2	0.973	7641.000
P_2_C_R		
Category 1	0.196	1542.000
Category 2	0.804	6312.000
P_5_R_D		
Category 1	0.356	2500.000
Category 2	0.644	4514.000
P_7_R		
Category 1	0.219	1208.000
Category 2	0.781	4308.000
P_8_R		
Category 1	0.381	1926.000
Category 2	0.619	3131.000
P_9_A_R		
Category 1	0.044	224.000
Category 2	0.956	4830.000
P_9_B_R		
Category 1	0.045	225.000
Category 2	0.955	4829.000
P_9_D_R		
Category 1	0.137	693.000
Category 2	0.863	4361.000
P_3_R_D		
Category 1	0.506	3786.000
Category 2	0.494	3703.000
P_10_R		
Category 1	0.451	3083.000
Category 2	0.549	3749.000
P_11_R_D		
Category 1	0.530	3753.000

Category 2	0.470	3326.000
P_9_C_R		
Category 1	0.142	719.000
Category 2	0.858	4335.000
P_17_A_R		
Category 1	0.033	242.000
Category 2	0.967	7158.000
P_17_B_R		
Category 1	0.026	189.000
Category 2	0.974	7211.000
P_17_D_R		
Category 1	0.011	85.000
Category 2	0.989	7315.000
P_20_G_R		
Category 1	0.485	2982.000
Category 2	0.515	3172.000
P_20_K_R		
Category 1	0.330	2047.000
Category 2	0.670	4162.000
P_20_L_R		
Category 1	0.701	4134.000
Category 2	0.299	1767.000
P_20_I_R		
Category 1	0.310	1953.000
Category 2	0.690	4345.000
P_20_E_R		
Category 1	0.428	2661.000
Category 2	0.572	3562.000
P_20_J_R		
Category 1	0.439	2588.000
Category 2	0.561	3306.000

SAMPLE STATISTICS

ESTIMATED SAMPLE STATISTICS

	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_C_R	P_18_C_R	P_18_C_R	P_18_C_R	P_18_F_R
1	<u>-2.304</u>	<u>-1.634</u>	<u>-0.737</u>	<u>0.519</u>	<u>-1.964</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_F_R	P_18_F_R	P_18_F_R	P_18_G_R	P_18_G_R
1	<u>-1.209</u>	<u>-0.372</u>	<u>0.868</u>	<u>-1.109</u>	<u>-0.257</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_G_R	P_18_G_R	P_18_I_R	P_18_I_R	P_18_I_R
1	<u>0.580</u>	<u>1.597</u>	<u>-1.665</u>	<u>-1.023</u>	<u>-0.314</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_I_R	P_18_J_R	P_18_J_R	P_18_J_R	P_18_J_R
1	<u>0.674</u>	<u>-1.551</u>	<u>-0.722</u>	<u>0.127</u>	<u>1.176</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_B_R	P_18_B_R	P_18_B_R	P_18_B_R	P_18_D_R
1	<u>-1.872</u>	<u>-1.287</u>	<u>-0.534</u>	<u>0.412</u>	<u>-2.058</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_D_R	P_18_D_R	P_18_D_R	P_18_E_R	P_18_E_R
	<u></u>	<u></u>	<u></u>	<u></u>	<u></u>

1	-1.460	-0.751	0.229	-2.419	-1.812
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_E_R	P_18_E_R	P_18_H_R	P_18_H_R	P_18_H_R
1	<u>-1.007</u>	<u>0.061</u>	<u>-2.145</u>	<u>-1.472</u>	<u>-0.720</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_18_H_R	P_18_K_R	P_18_K_R	P_18_K_R	P_18_K_R
1	<u>0.295</u>	<u>-1.505</u>	<u>-0.800</u>	<u>-0.008</u>	<u>0.956</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_1_A_R\$	P_1_B_R\$	P_2_A_R\$	P_2_C_R\$	P_5_R_D\$
1	<u>-0.937</u>	<u>-0.715</u>	<u>-1.925</u>	<u>-0.855</u>	<u>-0.368</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_7_R\$1	P_8_R\$1	P_9_A_R\$	P_9_B_R\$	P_9_D_R\$
1	<u>-0.776</u>	<u>-0.303</u>	<u>-1.703</u>	<u>-1.700</u>	<u>-1.093</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_3_R_D\$	P_10_R\$1	P_11_R_D	P_9_C_R\$	P_17_A_R
1	<u>0.014</u>	<u>-0.122</u>	<u>0.076</u>	<u>-1.070</u>	<u>-1.842</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>-1.951</u>	<u>-2.274</u>	<u>-0.039</u>	<u>-0.441</u>	<u>0.526</u>
	MEANS/INTERCEPTS/THRESHOLDS				
	P_20_I_R	P_20_E_R	P_20_J_R		
1	<u>-0.496</u>	<u>-0.182</u>	<u>-0.153</u>		
	CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	<u>0.519</u>				
P_18_F_R	0.310	<u>0.399</u>			
P_18_G_R	0.214	0.272	<u>0.417</u>		
P_18_I_R	0.239	0.231	0.378	<u>0.535</u>	
P_18_J_R	0.347	0.209	0.105	0.056	<u>0.146</u>
P_18_B_R	0.317	0.178	0.055	0.065	0.183
P_18_D_R	0.160	0.052	-0.007	0.038	0.140
P_18_E_R	0.214	0.147	0.060	0.066	0.129
P_18_H_R	0.207	0.188	0.199	0.263	0.372
P_18_K_R	0.098	0.132	0.128	0.285	0.127
P_1_A_R	0.111	0.134	0.150	0.296	0.123
P_1_B_R	0.111	0.057	-0.010	0.022	0.010
P_2_A_R	0.076	0.115	0.145	0.316	0.128
P_2_C_R	0.105	0.160	0.105	0.189	0.135
P_5_R_D	0.114	0.150	0.083	0.230	0.131
P_7_R	0.113	0.202	0.113	0.217	0.128
P_8_R	0.038	0.093	0.066	0.076	0.059
P_9_A_R	0.041	0.076	-0.024	0.054	0.006
P_9_B_R	0.133	0.152	0.082	0.131	0.120
P_9_D_R	0.094	0.083	0.083	0.214	0.085
P_3_R_D	0.140	0.186	0.168	0.307	0.166
P_10_R					

P_11_R_D	0.213	0.251	0.230	0.293	0.237
P_9_C_R	0.008	0.101	0.098	0.173	0.056
P_17_A_R	0.131	0.130	0.062	0.217	0.176
P_17_B_R	0.108	0.172	0.103	0.265	0.212
P_17_D_R	0.030	0.134	0.077	0.217	0.090
P_20_G_R	-0.065	-0.122	-0.152	-0.255	-0.127
P_20_K_R	-0.126	-0.196	-0.213	-0.326	-0.145
P_20_L_R	-0.022	-0.091	-0.146	-0.216	-0.127
P_20_I_R	-0.109	-0.155	-0.169	-0.324	-0.161
P_20_E_R	-0.089	-0.162	-0.156	-0.309	-0.065
P_20_J_R	-0.070	-0.129	-0.095	-0.249	-0.104

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_D_R	0.657				
P_18_E_R	0.418	0.562			
P_18_H_R	0.507	0.537	0.458		
P_18_K_R	0.500	0.492	0.319	0.436	
P_1_A_R	0.058	0.045	0.011	0.042	0.102
P_1_B_R	0.046	0.039	-0.014	0.018	0.109
P_2_A_R	0.084	0.093	0.059	0.107	0.064
P_2_C_R	0.035	0.021	-0.031	0.006	0.100
P_5_R_D	0.136	0.107	0.087	0.089	0.151
P_7_R	0.082	0.082	0.054	0.062	0.111
P_8_R	0.122	0.101	0.061	0.093	0.173
P_9_A_R	0.074	0.043	-0.007	0.000	0.024
P_9_B_R	0.027	-0.018	-0.018	0.031	0.022
P_9_D_R	0.089	0.096	0.066	0.043	0.099
P_3_R_D	0.081	0.075	0.039	0.006	0.092
P_10_R	0.157	0.105	0.059	0.095	0.174
P_11_R_D	0.251	0.195	0.131	0.161	0.242
P_9_C_R	0.056	0.051	0.032	0.078	0.157
P_17_A_R	0.197	0.187	0.146	0.126	0.202
P_17_B_R	0.082	0.104	0.064	0.025	0.127
P_17_D_R	0.102	0.059	0.038	0.085	0.141
P_20_G_R	0.057	0.105	0.191	0.168	0.016
P_20_K_R	-0.036	0.026	0.202	0.071	-0.069
P_20_L_R	0.329	0.294	0.143	0.241	0.319
P_20_I_R	-0.088	-0.039	-0.027	-0.005	-0.116
P_20_E_R	-0.123	-0.071	-0.004	-0.037	-0.131
P_20_J_R	-0.051	-0.025	0.015	0.022	-0.088

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
P_1_B_R	0.887				
P_2_A_R	0.401	0.709			
P_2_C_R	0.911	0.966	-0.618		
P_5_R_D	0.462	0.474	0.264	0.436	
P_7_R	0.592	0.552	0.284	0.537	0.771
P_8_R	0.550	0.529	0.318	0.494	0.919
P_9_A_R	0.255	0.217	0.218	0.185	0.493
P_9_B_R	0.142	0.200	0.137	0.130	0.439
P_9_D_R	0.369	0.352	0.194	0.343	0.710
P_3_R_D	0.375	0.409	0.303	0.362	0.336
P_10_R	0.639	0.656	0.297	0.640	0.659
P_11_R_D	0.516	0.531	0.275	0.502	0.599
P_9_C_R	0.324	0.310	0.085	0.320	0.674
P_17_A_R	0.338	0.366	0.215	0.284	0.430
P_17_B_R	0.351	0.360	0.122	0.324	0.439
P_17_D_R	0.276	0.346	0.106	0.263	0.478
P_20_G_R	-0.306	-0.331	-0.150	-0.330	-0.306

P_20_K_R	-0.450	-0.492	-0.188	-0.513	-0.496
P_20_L_R	-0.222	-0.227	-0.010	-0.238	-0.235
P_20_I_R	-0.574	-0.613	-0.325	-0.605	-0.570
P_20_E_R	-0.352	-0.428	-0.168	-0.416	-0.399
P_20_J_R	-0.419	-0.448	-0.246	-0.435	-0.664

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
P_8_R	0.844				
P_9_A_R	0.365	0.873			
P_9_B_R	0.305	0.874	-0.427		
P_9_D_R	0.705	0.958	-0.594	-0.595	
P_3_R_D	0.327	0.359	0.158	0.159	0.252
P_10_R	0.658	0.699	0.224	0.295	0.539
P_11_R_D	0.581	0.620	0.194	0.233	0.501
P_9_C_R	0.474	0.961	-0.600	-0.600	-0.747
P_17_A_R	0.433	0.521	0.114	0.299	0.284
P_17_B_R	0.413	0.520	0.126	0.113	0.303
P_17_D_R	0.426	0.605	0.098	0.146	0.401
P_20_G_R	-0.256	-0.351	-0.195	-0.169	-0.209
P_20_K_R	-0.429	-0.563	-0.272	-0.179	-0.416
P_20_L_R	-0.202	-0.256	-0.034	-0.140	-0.190
P_20_I_R	-0.539	-0.615	-0.255	-0.255	-0.494
P_20_E_R	-0.359	-0.482	-0.203	-0.157	-0.285
P_20_J_R	-0.550	-0.728	-0.391	-0.255	-0.484

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
P_10_R	0.426				
P_11_R_D	0.417	0.820			
P_9_C_R	0.224	0.504	0.438		
P_17_A_R	0.381	0.521	0.525	0.239	
P_17_B_R	0.270	0.451	0.503	0.306	0.761
P_17_D_R	0.271	0.479	0.422	0.199	0.585
P_20_G_R	-0.362	-0.379	-0.344	-0.242	-0.197
P_20_K_R	-0.449	-0.560	-0.519	-0.429	-0.325
P_20_L_R	-0.192	-0.338	-0.292	-0.149	-0.096
P_20_I_R	-0.497	-0.683	-0.605	-0.431	-0.451
P_20_E_R	-0.407	-0.490	-0.451	-0.376	-0.378
P_20_J_R	-0.412	-0.603	-0.512	-0.599	-0.359

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_D_R	0.401				
P_20_G_R	-0.361	-0.122			
P_20_K_R	-0.329	-0.261	0.889		
P_20_L_R	-0.117	-0.032	0.606	0.718	
P_20_I_R	-0.449	-0.385	0.804	0.934	0.697
P_20_E_R	-0.230	-0.270	0.660	0.835	0.477
P_20_J_R	-0.315	-0.378	0.686	0.859	0.533

CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

	P_20_I_R	P_20_E_R	P_20_J_R
P_20_E_R	0.852		
P_20_J_R	0.887	0.726	

WARNING: THE BIVARIATE TABLE OF P\_2\_C\_R AND P\_2\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_A\_R AND P\_8\_R HAS AN EMPTY CELL.

WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_D\_R HAS AN EMPTY CELL.

SUMMARY OF MODEL FIT INFORMATION

Model	Number of Parameters	Chi-Square	Degrees of Freedom	P-Value
1-factor	33	48802.109	495	0.0000
2-factor	65	22445.137	463	0.0000
3-factor	96	12699.639	432	0.0000
4-factor	126	7009.516	402	0.0000
5-factor	155	4087.423	373	0.0000
6-factor	183	2250.612	345	0.0000
7-factor	210	1429.704	318	0.0000

Models Compared	Chi-Square	Degrees of Freedom	P-Value
1-factor against 2-factor	13168.170	32	0.0000
2-factor against 3-factor	6097.012	31	0.0000
3-factor against 4-factor	3702.120	30	0.0000
4-factor against 5-factor	2228.488	29	0.0000
5-factor against 6-factor	1815.356	28	0.0000
6-factor against 7-factor	817.862	27	0.0000

RESULTS FOR EXPLORATORY FACTOR ANALYSIS

EIGENVALUES FOR SAMPLE CORRELATION MATRIX					
	1	2	3	4	5
1	10.750	3.744	2.385	1.877	1.807
EIGENVALUES FOR SAMPLE CORRELATION MATRIX					
	6	7	8	9	10
1	1.663	1.602	1.467	1.426	1.243
EIGENVALUES FOR SAMPLE CORRELATION MATRIX					
	11	12	13	14	15
1	1.035	0.829	0.700	0.668	0.596
EIGENVALUES FOR SAMPLE CORRELATION MATRIX					
	16	17	18	19	20
1	0.525	0.498	0.463	0.425	0.389
EIGENVALUES FOR SAMPLE CORRELATION MATRIX					
	21	22	23	24	25
1	0.324	0.306	0.229	0.222	0.166
EIGENVALUES FOR SAMPLE CORRELATION MATRIX					
	26	27	28	29	30
1	0.139	0.131	0.101	0.073	-0.010

EIGENVALUES FOR SAMPLE CORRELATION MATRIX

	31	32	33
1	-0.034	-0.619	-2.122

EXPLORATORY FACTOR ANALYSIS WITH 1 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters 33

Chi-Square Test of Model Fit

Value	48802.109*
Degrees of Freedom	495
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.110
90 Percent C.I.	0.109 0.111
Probability RMSEA <= .05	0.000

CFI/TLI

CFI	0.790
TLI	0.776

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.183
-------	-------

MINIMUM ROTATION FUNCTION VALUE 12.18221

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1
P_18_C_R	0.285*
P_18_F_R	0.316*
P_18_G_R	0.292*
P_18_I_R	0.434*
P_18_J_R	0.333*
P_18_B_R	0.308*
P_18_D_R	0.295*
P_18_E_R	0.148*
P_18_H_R	0.200*
P_18_K_R	0.307*
P_1_A_R	0.816*
P_1_B_R	0.945*
P_2_A_R	0.387*

P_2_C_R	0.956*
P_5_R_D	0.820*
P_7_R	0.742*
P_8_R	0.893*
P_9_A_R	0.310*
P_9_B_R	0.266*
P_9_D_R	0.567*
P_3_R_D	0.458*
P_10_R	0.777*
P_11_R_D	0.730*
P_9_C_R	0.512*
P_17_A_R	0.534*
P_17_B_R	0.517*
P_17_D_R	0.456*
P_20_G_R	-0.720*
P_20_K_R	-0.944*
P_20_L_R	-0.460*
P_20_I_R	-0.955*
P_20_E_R	-0.747*
P_20_J_R	-0.829*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

1	<u>1.000</u>				
	ESTIMATED RESIDUAL VARIANCES				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	<u>0.919</u>	<u>0.900</u>	<u>0.915</u>	<u>0.812</u>	<u>0.889</u>
	ESTIMATED RESIDUAL VARIANCES				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	<u>0.905</u>	<u>0.913</u>	<u>0.978</u>	<u>0.960</u>	<u>0.906</u>
	ESTIMATED RESIDUAL VARIANCES				
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	<u>0.335</u>	<u>0.108</u>	<u>0.850</u>	<u>0.086</u>	<u>0.328</u>
	ESTIMATED RESIDUAL VARIANCES				
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	<u>0.450</u>	<u>0.202</u>	<u>0.904</u>	<u>0.929</u>	<u>0.679</u>
	ESTIMATED RESIDUAL VARIANCES				
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	<u>0.790</u>	<u>0.397</u>	<u>0.467</u>	<u>0.738</u>	<u>0.715</u>
	ESTIMATED RESIDUAL VARIANCES				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>0.733</u>	<u>0.792</u>	<u>0.481</u>	<u>0.109</u>	<u>0.789</u>
	ESTIMATED RESIDUAL VARIANCES				
	P_20_I_R	P_20_E_R	P_20_J_R		
1	<u>0.089</u>	<u>0.442</u>	<u>0.312</u>		

S.E. GEOMIN ROTATED LOADINGS  
1



P_18_C_R	0.011
P_18_F_R	0.011
P_18_G_R	0.011
P_18_I_R	0.010
P_18_J_R	0.011
P_18_B_R	0.011
P_18_D_R	0.011
P_18_E_R	0.012
P_18_H_R	0.012
P_18_K_R	0.011
P_1_A_R	0.006
P_1_B_R	0.004
P_2_A_R	0.022
P_2_C_R	0.004
P_5_R_D	0.007
P_7_R	0.009
P_8_R	0.007
P_9_A_R	0.024
P_9_B_R	0.026
P_9_D_R	0.014
P_3_R_D	0.012
P_10_R	0.007
P_11_R_D	0.009
P_9_C_R	0.016
P_17_A_R	0.018
P_17_B_R	0.020
P_17_D_R	0.029
P_20_G_R	0.008
P_20_K_R	0.005
P_20_L_R	0.011
P_20_I_R	0.005
P_20_E_R	0.008
P_20_J_R	0.006

S.E. GEOMIN FACTOR CORRELATIONS

1	0.000
---	-------

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.006	0.007	0.006	0.008	0.007

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.007	0.007	0.004	0.005	0.007

S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.010	0.008	0.017	0.007	0.011

S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.014	0.012	0.015	0.014	0.016

S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R

1            0.011            0.012            0.013            0.016            0.019

S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>0.021</u>	<u>0.027</u>	<u>0.011</u>	<u>0.010</u>	<u>0.011</u>

S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	<u>0.009</u>	<u>0.012</u>	<u>0.010</u>

Est./S.E. GEOMIN ROTATED LOADINGS

1

P_18_C_R	<u>25.301</u>
P_18_F_R	28.738
P_18_G_R	26.847
P_18_I_R	44.648
P_18_J_R	31.600
P_18_B_R	27.977
P_18_D_R	26.734
P_18_E_R	12.189
P_18_H_R	17.348
P_18_K_R	28.948
P_1_A_R	129.357
P_1_B_R	226.268
P_2_A_R	17.536
P_2_C_R	252.578
P_5_R_D	123.365
P_7_R	78.762
P_8_R	133.667
P_9_A_R	13.093
P_9_B_R	10.292
P_9_D_R	39.966
P_3_R_D	37.250
P_10_R	104.041
P_11_R_D	84.721
P_9_C_R	32.841
P_17_A_R	29.529
P_17_B_R	25.942
P_17_D_R	15.558
P_20_G_R	-92.054
P_20_K_R	-172.977
P_20_L_R	-40.221
P_20_I_R	-193.777
P_20_E_R	-93.721
P_20_J_R	-134.515

Est./S.E. GEOMIN FACTOR CORRELATIONS

1

1	<u>0.000</u>
---	--------------

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	<u>143.381</u>	<u>129.386</u>	<u>143.867</u>	<u>96.220</u>	<u>126.779</u>

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	<u>133.300</u>	<u>139.788</u>	<u>271.364</u>	<u>207.355</u>	<u>139.231</u>

	Est./S.E.	ESTIMATED RESIDUAL VARIANCES			
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	32.552	13.670	49.825	11.942	30.151

	Est./S.E.	ESTIMATED RESIDUAL VARIANCES			
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	32.234	16.897	61.436	67.800	42.191

	Est./S.E.	ESTIMATED RESIDUAL VARIANCES			
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	70.243	34.197	37.148	46.206	36.987

	Est./S.E.	ESTIMATED RESIDUAL VARIANCES			
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	35.640	29.582	42.671	10.613	75.084

	Est./S.E.	ESTIMATED RESIDUAL VARIANCES			
	P_20_I_R	P_20_E_R	P_20_J_R		
1	9.408	37.120	30.563		

EXPLORATORY FACTOR ANALYSIS WITH 2 FACTOR(S) :

MODEL FIT INFORMATION

Number of Free Parameters 65

Chi-Square Test of Model Fit

Value	22445.137*
Degrees of Freedom	463
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.077	
90 Percent C.I.	0.076	0.078
Probability RMSEA <= .05	0.000	

CFI/TLI

CFI	0.905
TLI	0.891

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value

0.151

MINIMUM ROTATION FUNCTION VALUE

3.07866

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1	2
P_18_C_R	0.453*	0.136*
P_18_F_R	0.369*	0.209*
P_18_G_R	0.271*	0.224*
P_18_I_R	0.294*	0.372*
P_18_J_R	0.396*	0.218*
P_18_B_R	0.736*	0.003
P_18_D_R	0.790*	-0.043*
P_18_E_R	0.568*	-0.077*
P_18_H_R	0.629*	-0.052*
P_18_K_R	0.630*	0.083*
P_1_A_R	0.029*	0.829*
P_1_B_R	-0.010*	0.950*
P_2_A_R	0.074*	0.389*
P_2_C_R	-0.031*	0.964*
P_5_R_D	0.263*	0.784*
P_7_R	0.214*	0.721*
P_8_R	0.288*	0.855*
P_9_A_R	0.054	0.317*
P_9_B_R	0.029	0.278*
P_9_D_R	0.168*	0.558*
P_3_R_D	0.062*	0.468*
P_10_R	0.234*	0.759*
P_11_R_D	0.325*	0.685*
P_9_C_R	0.107*	0.517*
P_17_A_R	0.276*	0.486*
P_17_B_R	0.199*	0.493*
P_17_D_R	0.179*	0.435*
P_20_G_R	0.288*	-0.752*
P_20_K_R	0.244*	-0.950*
P_20_L_R	0.387*	-0.559*
P_20_I_R	0.109*	-0.960*
P_20_E_R	0.064*	-0.766*
P_20_J_R	0.081*	-0.852*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	1	2
1	1.000	
2	0.040*	1.000

ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.772	0.814	0.872	0.767	0.789

ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.459	0.377	0.675	0.605	0.593

ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.310	0.098	0.841	0.072	0.300

## ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.422	0.167	0.895	0.921	0.653

## ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.775	0.356	0.408	0.717	0.677

## ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.709	0.772	0.370	0.057	0.555

## ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	0.075	0.413	0.273

## S.E. GEOMIN ROTATED LOADINGS

	1	2
P_18_C_R	0.010	0.013
P_18_F_R	0.011	0.013
P_18_G_R	0.012	0.012
P_18_I_R	0.012	0.012
P_18_J_R	0.011	0.013
P_18_B_R	0.006	0.003
P_18_D_R	0.006	0.012
P_18_E_R	0.009	0.013
P_18_H_R	0.008	0.013
P_18_K_R	0.008	0.012
P_1_A_R	0.014	0.006
P_1_B_R	0.004	0.004
P_2_A_R	0.030	0.022
P_2_C_R	0.011	0.004
P_5_R_D	0.018	0.009
P_7_R	0.020	0.011
P_8_R	0.020	0.010
P_9_A_R	0.031	0.025
P_9_B_R	0.032	0.028
P_9_D_R	0.023	0.015
P_3_R_D	0.016	0.013
P_10_R	0.016	0.009
P_11_R_D	0.017	0.011
P_9_C_R	0.023	0.017
P_17_A_R	0.028	0.021
P_17_B_R	0.028	0.022
P_17_D_R	0.042	0.032
P_20_G_R	0.019	0.009
P_20_K_R	0.021	0.007
P_20_L_R	0.018	0.012
P_20_I_R	0.020	0.005
P_20_E_R	0.019	0.008
P_20_J_R	0.020	0.006

## S.E. GEOMIN FACTOR CORRELATIONS

	1	2
1	0.000	
2	0.016	0.000

	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	<u>0.009</u>	<u>0.009</u>	<u>0.008</u>	<u>0.009</u>	<u>0.009</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	<u>0.009</u>	<u>0.010</u>	<u>0.010</u>	<u>0.010</u>	<u>0.010</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	<u>0.010</u>	<u>0.008</u>	<u>0.018</u>	<u>0.007</u>	<u>0.011</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	<u>0.014</u>	<u>0.012</u>	<u>0.016</u>	<u>0.015</u>	<u>0.017</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	<u>0.012</u>	<u>0.012</u>	<u>0.013</u>	<u>0.017</u>	<u>0.021</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>0.022</u>	<u>0.029</u>	<u>0.011</u>	<u>0.009</u>	<u>0.015</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_20_I_R	P_20_E_R	P_20_J_R		
1	<u>0.009</u>	<u>0.012</u>	<u>0.010</u>		
	Est./S.E. GEOMIN ROTATED LOADINGS				
	1	2			
P_18_C_R	<u>44.221</u>	<u>10.588</u>			
P_18_F_R	34.296	16.394			
P_18_G_R	23.018	18.228			
P_18_I_R	25.523	31.419			
P_18_J_R	36.724	16.961			
P_18_B_R	115.369	1.005			
P_18_D_R	124.606	-3.602			
P_18_E_R	62.088	-5.772			
P_18_H_R	76.990	-4.114			
P_18_K_R	80.243	6.695			
P_1_A_R	2.001	131.056			
P_1_B_R	-2.284	222.712			
P_2_A_R	2.463	17.781			
P_2_C_R	-2.891	253.857			
P_5_R_D	14.785	86.191			
P_7_R	10.836	65.193			
P_8_R	14.139	89.229			
P_9_A_R	1.726	12.951			
P_9_B_R	0.904	9.856			
P_9_D_R	7.414	36.469			
P_3_R_D	3.924	36.608			
P_10_R	14.275	84.131			
P_11_R_D	19.598	63.677			
P_9_C_R	4.676	30.712			

P_17_A_R	9.951	23.318
P_17_B_R	6.983	22.947
P_17_D_R	4.286	13.739
P_20_G_R	15.151	-81.528
P_20_K_R	11.698	-136.179
P_20_L_R	21.950	-46.366
P_20_I_R	5.385	-180.144
P_20_E_R	3.390	-93.341
P_20_J_R	4.043	-137.341

Est./S.E. GEOMIN FACTOR CORRELATIONS

	1	2
1	0.000	
2	2.479	0.000

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	83.173	93.817	113.195	84.999	89.048

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	48.938	37.892	64.573	58.990	61.766

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	29.802	12.211	47.082	9.877	27.022

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	29.386	13.741	56.667	60.208	37.967

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	64.706	29.804	31.521	41.976	32.123

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	31.933	26.893	33.483	6.120	37.334

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	8.073	33.738	26.903

FACTOR STRUCTURE

	1	2
P_18_C_R	0.458	0.154
P_18_F_R	0.378	0.224
P_18_G_R	0.280	0.235
P_18_I_R	0.309	0.383
P_18_J_R	0.405	0.234
P_18_B_R	0.736	0.032
P_18_D_R	0.788	-0.011
P_18_E_R	0.565	-0.055
P_18_H_R	0.627	-0.027

P_18_K_R	0.633	0.108
P_1_A_R	0.062	0.830
P_1_B_R	0.028	0.949
P_2_A_R	0.089	0.392
P_2_C_R	0.008	0.963
P_5_R_D	0.294	0.794
P_7_R	0.243	0.729
P_8_R	0.322	0.867
P_9_A_R	0.067	0.320
P_9_B_R	0.040	0.279
P_9_D_R	0.190	0.565
P_3_R_D	0.081	0.470
P_10_R	0.264	0.768
P_11_R_D	0.353	0.698
P_9_C_R	0.128	0.521
P_17_A_R	0.295	0.497
P_17_B_R	0.218	0.501
P_17_D_R	0.196	0.442
P_20_G_R	0.258	-0.740
P_20_K_R	0.206	-0.940
P_20_L_R	0.365	-0.544
P_20_I_R	0.071	-0.955
P_20_E_R	0.033	-0.764
P_20_J_R	0.047	-0.849

EXPLORATORY FACTOR ANALYSIS WITH 3 FACTOR(S) :

MODEL FIT INFORMATION

Number of Free Parameters 96

Chi-Square Test of Model Fit

Value	12699.639*
Degrees of Freedom	432
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.059
90 Percent C.I.	0.058 0.060
Probability RMSEA <= .05	0.000

CFI/TLI

CFI	0.947
TLI	0.935

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.135
-------	-------



MINIMUM ROTATION FUNCTION VALUE 1.27926

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1	2	3
P_18_C_R	-0.073*	0.603*	0.098*
P_18_F_R	-0.035	0.535*	-0.040
P_18_G_R	-0.068*	0.471*	-0.166*
P_18_I_R	0.074	0.477*	-0.222*
P_18_J_R	-0.075*	0.593*	-0.068
P_18_B_R	0.020*	0.706*	0.592*
P_18_D_R	0.002	0.741*	0.680*
P_18_E_R	-0.002	0.501*	0.547*
P_18_H_R	-0.004	0.583*	0.561*
P_18_K_R	0.012	0.669*	0.395*
P_1_A_R	0.931*	-0.217*	0.009
P_1_B_R	1.037*	-0.283*	0.012
P_2_A_R	0.533*	-0.110*	0.113*
P_2_C_R	1.052*	-0.313*	0.002
P_5_R_D	0.878*	-0.021	0.009
P_7_R	0.815*	-0.040	0.006
P_8_R	0.955*	-0.032	-0.001
P_9_A_R	0.345*	-0.018	-0.054
P_9_B_R	0.321*	-0.051	-0.048
P_9_D_R	0.623*	0.010	-0.021
P_3_R_D	0.425*	0.041*	-0.156*
P_10_R	0.763*	0.107*	-0.092*
P_11_R_D	0.651*	0.249*	-0.028
P_9_C_R	0.545*	0.002	-0.089*
P_17_A_R	0.487*	0.203*	0.034
P_17_B_R	0.467*	0.148*	-0.063
P_17_D_R	0.465*	0.080	0.003
P_20_G_R	-0.374*	0.035*	0.737*
P_20_K_R	-0.543*	-0.022*	0.756*
P_20_L_R	-0.287*	0.215*	0.736*
P_20_I_R	-0.681*	-0.042*	0.589*
P_20_E_R	-0.486*	-0.107*	0.510*
P_20_J_R	-0.664*	0.013	0.505*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	1	2	3
1	1.000		
2	0.380*	1.000	
3	-0.117*	-0.354*	1.000

ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.695	0.710	0.718	0.612	0.644

ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.438	0.344	0.643	0.578	0.578

ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.240	0.068	0.740	0.045	0.244

ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R

1	<u>0.360</u>	<u>0.109</u>	<u>0.879</u>	<u>0.902</u>	<u>0.603</u>
---	--------------	--------------	--------------	--------------	--------------

ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
--	---------	--------	----------	---------	----------

1	<u>0.759</u>	<u>0.313</u>	<u>0.381</u>	<u>0.682</u>	<u>0.654</u>
---	--------------	--------------	--------------	--------------	--------------

ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
--	----------	----------	----------	----------	----------

1	<u>0.690</u>	<u>0.749</u>	<u>0.279</u>	<u>0.015</u>	<u>0.439</u>
---	--------------	--------------	--------------	--------------	--------------

ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
--	----------	----------	----------

1	<u>0.054</u>	<u>0.357</u>	<u>0.237</u>
---	--------------	--------------	--------------

S.E. GEOMIN ROTATED LOADINGS

	<u>1</u>	<u>2</u>	<u>3</u>
P_18_C_R	0.028	0.019	0.037
P_18_F_R	0.030	0.018	0.033
P_18_G_R	0.034	0.019	0.031
P_18_I_R	0.039	0.019	0.033
P_18_J_R	0.034	0.020	0.037
P_18_B_R	0.010	0.028	0.039
P_18_D_R	0.007	0.030	0.041
P_18_E_R	0.010	0.025	0.029
P_18_H_R	0.010	0.025	0.033
P_18_K_R	0.009	0.022	0.038
P_1_A_R	0.010	0.029	0.012
P_1_B_R	0.012	0.031	0.009
P_2_A_R	0.024	0.041	0.034
P_2_C_R	0.012	0.029	0.006
P_5_R_D	0.010	0.020	0.015
P_7_R	0.012	0.024	0.018
P_8_R	0.009	0.018	0.016
P_9_A_R	0.031	0.037	0.034
P_9_B_R	0.032	0.039	0.038
P_9_D_R	0.019	0.025	0.023
P_3_R_D	0.021	0.019	0.018
P_10_R	0.020	0.024	0.020
P_11_R_D	0.022	0.025	0.023
P_9_C_R	0.020	0.021	0.025
P_17_A_R	0.029	0.035	0.030
P_17_B_R	0.032	0.035	0.036
P_17_D_R	0.038	0.049	0.042
P_20_G_R	0.047	0.016	0.016
P_20_K_R	0.052	0.009	0.019
P_20_L_R	0.041	0.023	0.019
P_20_I_R	0.045	0.015	0.020
P_20_E_R	0.043	0.018	0.020
P_20_J_R	0.035	0.009	0.018

S.E. GEOMIN FACTOR CORRELATIONS

	<u>1</u>	<u>2</u>	<u>3</u>
1	0.000		
2	0.031	0.000	
3	0.055	0.051	0.000

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	<u>0.010</u>	<u>0.010</u>	<u>0.010</u>	<u>0.011</u>	<u>0.010</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	<u>0.010</u>	<u>0.010</u>	<u>0.011</u>	<u>0.011</u>	<u>0.010</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	<u>0.010</u>	<u>0.009</u>	<u>0.022</u>	<u>0.008</u>	<u>0.012</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	<u>0.015</u>	<u>0.012</u>	<u>0.018</u>	<u>0.017</u>	<u>0.019</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	<u>0.013</u>	<u>0.012</u>	<u>0.013</u>	<u>0.019</u>	<u>0.022</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>0.023</u>	<u>0.031</u>	<u>0.011</u>	<u>0.008</u>	<u>0.016</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_20_I_R	P_20_E_R	P_20_J_R		
1	<u>0.009</u>	<u>0.012</u>	<u>0.010</u>		
	Est./S.E. GEOMIN ROTATED LOADINGS				
	1	2	3		
P_18_C_R	<u>-2.646</u>	<u>32.416</u>	<u>2.677</u>		
P_18_F_R	-1.161	29.631	-1.193		
P_18_G_R	-2.015	25.195	-5.347		
P_18_I_R	1.880	25.202	-6.766		
P_18_J_R	-2.191	30.037	-1.852		
P_18_B_R	2.078	25.453	15.002		
P_18_D_R	0.290	24.638	16.572		
P_18_E_R	-0.162	19.942	19.035		
P_18_H_R	-0.375	22.894	17.000		
P_18_K_R	1.351	30.940	10.284		
P_1_A_R	90.574	-7.427	0.774		
P_1_B_R	90.037	-9.238	1.310		
P_2_A_R	22.315	-2.705	3.345		
P_2_C_R	87.346	-10.680	0.263		
P_5_R_D	89.338	-1.019	0.582		
P_7_R	69.856	-1.700	0.364		
P_8_R	105.305	-1.799	-0.067		
P_9_A_R	11.024	-0.476	-1.573		
P_9_B_R	10.154	-1.321	-1.288		
P_9_D_R	33.523	0.385	-0.910		
P_3_R_D	20.147	2.129	-8.657		
P_10_R	38.434	4.472	-4.615		
P_11_R_D	29.070	9.939	-1.249		
P_9_C_R	27.152	0.108	-3.555		
P_17_A_R	16.658	5.839	1.116		
P_17_B_R	14.412	4.269	-1.743		

P_17_D_R	12.376	1.621	0.079
P_20_G_R	-7.906	2.185	45.337
P_20_K_R	-10.440	-2.367	39.794
P_20_L_R	-6.951	9.235	39.064
P_20_I_R	-15.169	-2.858	29.781
P_20_E_R	-11.329	-5.926	24.971
P_20_J_R	-18.813	1.529	28.658

Est./S.E. GEOMIN FACTOR CORRELATIONS

	1	2	3
1	0.000		
2	12.230	0.000	
3	-2.146	-6.945	0.000

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	69.514	69.853	68.924	57.562	61.915

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	45.264	33.463	57.568	53.184	59.611

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	24.010	7.892	33.919	5.740	21.161

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	24.749	8.925	49.206	52.068	31.933

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	59.546	25.433	28.563	36.561	29.186

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	29.675	24.471	24.978	1.808	26.662

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	5.780	30.002	23.096

FACTOR STRUCTURE

	1	2	3
P_18_C_R	0.145	0.541	-0.107
P_18_F_R	0.173	0.536	-0.225
P_18_G_R	0.130	0.503	-0.325
P_18_I_R	0.281	0.583	-0.400
P_18_J_R	0.158	0.589	-0.269
P_18_B_R	0.219	0.505	0.340
P_18_D_R	0.204	0.501	0.418
P_18_E_R	0.125	0.307	0.370
P_18_H_R	0.152	0.383	0.355

P_18_K_R	0.220	0.534	0.157
P_1_A_R	0.847	0.134	-0.023
P_1_B_R	0.928	0.107	-0.010
P_2_A_R	0.478	0.053	0.089
P_2_C_R	0.933	0.086	-0.011
P_5_R_D	0.869	0.310	-0.087
P_7_R	0.799	0.267	-0.075
P_8_R	0.943	0.331	-0.102
P_9_A_R	0.345	0.132	-0.088
P_9_B_R	0.308	0.088	-0.068
P_9_D_R	0.629	0.254	-0.097
P_3_R_D	0.459	0.258	-0.221
P_10_R	0.814	0.429	-0.219
P_11_R_D	0.749	0.506	-0.193
P_9_C_R	0.556	0.241	-0.154
P_17_A_R	0.560	0.376	-0.095
P_17_B_R	0.530	0.348	-0.170
P_17_D_R	0.495	0.255	-0.080
P_20_G_R	-0.447	-0.368	0.768
P_20_K_R	-0.641	-0.496	0.827
P_20_L_R	-0.292	-0.154	0.694
P_20_I_R	-0.766	-0.509	0.684
P_20_E_R	-0.586	-0.472	0.604
P_20_J_R	-0.719	-0.417	0.578

EXPLORATORY FACTOR ANALYSIS WITH 4 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters 126

Chi-Square Test of Model Fit

Value	7009.516*
Degrees of Freedom	402
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.045	
90 Percent C.I.	0.044	0.046
Probability RMSEA <= .05	1.000	

CFI/TLI

CFI	0.971
TLI	0.962

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.125
-------	-------

MINIMUM ROTATION FUNCTION VALUE

1.77341

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1	2	3	4
P_18_C_R	0.398*	0.435*	0.016	0.063*
P_18_F_R	0.274*	0.466*	0.053*	0.115*
P_18_G_R	0.122*	0.559*	0.113*	0.059*
P_18_I_R	0.067	0.595*	0.266*	0.111*
P_18_J_R	0.234*	0.576*	0.166*	-0.001
P_18_B_R	0.785*	0.093*	-0.047	0.035
P_18_D_R	0.848*	0.066	-0.051	-0.010
P_18_E_R	0.617*	-0.015	-0.017	-0.066*
P_18_H_R	0.665*	0.054	0.004	-0.066*
P_18_K_R	0.599*	0.259*	0.097*	-0.026*
P_1_A_R	-0.291*	0.036*	0.950*	0.013
P_1_B_R	-0.341*	0.064*	1.023*	0.002
P_2_A_R	-0.112*	-0.081*	0.608*	-0.063
P_2_C_R	-0.382*	0.085*	1.048*	-0.028*
P_5_R_D	0.088*	-0.370*	0.420*	0.559*
P_7_R	-0.007	-0.283*	0.554*	0.395*
P_8_R	0.074*	-0.391*	0.473*	0.593*
P_9_A_R	0.007	-0.135*	0.162*	0.273*
P_9_B_R	-0.009	-0.169*	0.143*	0.252*
P_9_D_R	0.050*	-0.241*	0.348*	0.412*
P_3_R_D	-0.031	0.010	0.245*	0.349*
P_10_R	-0.002	-0.039*	0.583*	0.424*
P_11_R_D	0.131*	0.036	0.490*	0.399*
P_9_C_R	0.042	-0.208*	0.229*	0.460*
P_17_A_R	0.169*	-0.053	0.316*	0.332*
P_17_B_R	0.067	-0.006	0.327*	0.324*
P_17_D_R	0.069	-0.112*	0.312*	0.282*
P_20_G_R	0.220*	-0.150*	0.138*	-0.859*
P_20_K_R	0.177*	-0.145*	0.024*	-0.968*
P_20_L_R	0.415*	-0.123*	0.068*	-0.658*
P_20_I_R	0.114*	-0.055	-0.155*	-0.897*
P_20_E_R	0.011	-0.078*	0.026	-0.826*
P_20_J_R	0.064*	0.114*	-0.064	-0.894*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	1	2	3	4
1	1.000			
2	-0.128*	1.000		
3	0.329*	0.021	1.000	
4	0.085*	0.145*	0.433*	1.000

ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.675	0.687	0.646	0.504	0.590

ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.410	0.318	0.624	0.566	0.571

ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.180	0.053	0.687	0.023	0.188

## ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.310	0.061	0.853	0.871	0.539

## ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.749	0.271	0.359	0.618	0.626

## ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.674	0.721	0.242	0.010	0.406

## ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	0.047	0.312	0.175

## S.E. GEOMIN ROTATED LOADINGS

	1	2	3	4
P_18_C_R	0.031	0.019	0.019	0.021
P_18_F_R	0.034	0.015	0.025	0.024
P_18_G_R	0.039	0.011	0.042	0.028
P_18_I_R	0.043	0.012	0.045	0.029
P_18_J_R	0.042	0.015	0.041	0.021
P_18_B_R	0.009	0.037	0.025	0.018
P_18_D_R	0.009	0.040	0.027	0.019
P_18_E_R	0.012	0.031	0.026	0.022
P_18_H_R	0.012	0.034	0.022	0.020
P_18_K_R	0.023	0.033	0.018	0.013
P_1_A_R	0.050	0.017	0.013	0.014
P_1_B_R	0.054	0.018	0.011	0.011
P_2_A_R	0.040	0.035	0.031	0.039
P_2_C_R	0.057	0.018	0.012	0.010
P_5_R_D	0.016	0.024	0.071	0.049
P_7_R	0.013	0.026	0.057	0.044
P_8_R	0.013	0.027	0.077	0.054
P_9_A_R	0.035	0.035	0.055	0.043
P_9_B_R	0.037	0.039	0.053	0.045
P_9_D_R	0.022	0.027	0.058	0.042
P_3_R_D	0.017	0.017	0.025	0.021
P_10_R	0.020	0.017	0.035	0.028
P_11_R_D	0.025	0.019	0.033	0.027
P_9_C_R	0.024	0.026	0.053	0.038
P_17_A_R	0.034	0.029	0.054	0.043
P_17_B_R	0.035	0.030	0.053	0.041
P_17_D_R	0.048	0.047	0.065	0.055
P_20_G_R	0.041	0.044	0.023	0.018
P_20_K_R	0.038	0.041	0.012	0.012
P_20_L_R	0.033	0.045	0.024	0.022
P_20_I_R	0.028	0.028	0.024	0.013
P_20_E_R	0.026	0.029	0.022	0.014
P_20_J_R	0.026	0.029	0.038	0.020

## S.E. GEOMIN FACTOR CORRELATIONS

	1	2	3	4
1	0.000			
2	0.027	0.000		
3	0.024	0.046	0.000	

4	0.030	0.020	0.023	0.000	
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	<u>0.010</u>	<u>0.010</u>	<u>0.011</u>	<u>0.012</u>	<u>0.011</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	<u>0.010</u>	<u>0.010</u>	<u>0.011</u>	<u>0.011</u>	<u>0.010</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	<u>0.010</u>	<u>0.009</u>	<u>0.025</u>	<u>0.008</u>	<u>0.012</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	<u>0.015</u>	<u>0.012</u>	<u>0.019</u>	<u>0.020</u>	<u>0.020</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	<u>0.013</u>	<u>0.013</u>	<u>0.014</u>	<u>0.020</u>	<u>0.024</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>0.024</u>	<u>0.033</u>	<u>0.012</u>	<u>0.009</u>	<u>0.017</u>
	S.E. ESTIMATED RESIDUAL VARIANCES				
	P_20_I_R	P_20_E_R	P_20_J_R		
1	<u>0.010</u>	<u>0.012</u>	<u>0.010</u>		
	Est./S.E. GEOMIN ROTATED LOADINGS				
	1	2	3	4	
P_18_C_R	<u>13.050</u>	<u>22.949</u>	<u>0.872</u>	<u>3.061</u>	
P_18_F_R	8.124	31.163	2.129	4.869	
P_18_G_R	3.117	48.993	2.703	2.108	
P_18_I_R	1.561	48.822	5.980	3.775	
P_18_J_R	5.519	38.257	4.026	-0.046	
P_18_B_R	83.551	2.495	-1.917	1.919	
P_18_D_R	97.546	1.669	-1.909	-0.507	
P_18_E_R	52.939	-0.477	-0.653	-3.003	
P_18_H_R	57.509	1.602	0.171	-3.368	
P_18_K_R	26.597	7.934	5.482	-2.030	
P_1_A_R	-5.778	2.129	72.325	0.942	
P_1_B_R	-6.332	3.616	90.565	0.148	
P_2_A_R	-2.787	-2.299	19.714	-1.601	
P_2_C_R	-6.694	4.736	86.592	-2.778	
P_5_R_D	5.646	-15.505	5.902	11.340	
P_7_R	-0.491	-11.063	9.682	9.039	
P_8_R	5.854	-14.419	6.165	11.063	
P_9_A_R	0.199	-3.914	2.955	6.324	
P_9_B_R	-0.244	-4.384	2.691	5.636	
P_9_D_R	2.273	-8.950	6.047	9.761	
P_3_R_D	-1.771	0.609	9.632	16.613	
P_10_R	-0.106	-2.291	16.784	15.148	
P_11_R_D	5.337	1.928	14.692	14.555	



P_9_C_R	1.766	-7.851	4.286	12.112
P_17_A_R	4.978	-1.804	5.832	7.713
P_17_B_R	1.918	-0.193	6.168	7.831
P_17_D_R	1.452	-2.394	4.820	5.151
P_20_G_R	5.423	-3.393	6.112	-46.559
P_20_K_R	4.660	-3.556	2.021	-79.309
P_20_L_R	12.518	-2.699	2.805	-29.787
P_20_I_R	4.037	-1.936	-6.424	-69.076
P_20_E_R	0.444	-2.663	1.192	-59.943
P_20_J_R	2.447	3.909	-1.703	-45.333

Est./S.E. GEOMIN FACTOR CORRELATIONS

	1	2	3	4
1	0.000			
2	-4.666	0.000		
3	13.543	0.459	0.000	
4	2.827	7.300	19.246	0.000

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	65.559	65.832	58.413	43.579	53.701

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	41.502	30.579	54.340	50.908	58.209

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	17.534	5.778	27.610	2.852	15.548

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	20.513	5.064	44.132	42.567	27.075

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	56.401	21.136	26.367	30.646	25.937

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	27.636	22.104	20.452	1.121	23.220

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	4.906	26.102	17.802

FACTOR STRUCTURE

	1	2	3	4
P_18_C_R	0.353	0.394	0.184	0.167
P_18_F_R	0.241	0.449	0.203	0.229
P_18_G_R	0.093	0.554	0.191	0.200
P_18_I_R	0.088	0.608	0.349	0.319
P_18_J_R	0.215	0.550	0.255	0.175
P_18_B_R	0.761	-0.003	0.229	0.095

P_18_D_R	0.822	-0.044	0.226	0.050
P_18_E_R	0.608	-0.104	0.158	-0.023
P_18_H_R	0.654	-0.041	0.196	0.000
P_18_K_R	0.596	0.180	0.289	0.104
P_1_A_R	0.018	0.095	0.860	0.405
P_1_B_R	-0.012	0.129	0.913	0.425
P_2_A_R	0.093	-0.063	0.542	0.179
P_2_C_R	-0.050	0.152	0.912	0.406
P_5_R_D	0.321	-0.291	0.684	0.695
P_7_R	0.245	-0.213	0.717	0.593
P_8_R	0.330	-0.304	0.746	0.747
P_9_A_R	0.101	-0.093	0.279	0.324
P_9_B_R	0.081	-0.128	0.245	0.289
P_9_D_R	0.230	-0.180	0.538	0.532
P_3_R_D	0.078	0.070	0.386	0.454
P_10_R	0.231	0.035	0.765	0.671
P_11_R_D	0.322	0.088	0.706	0.627
P_9_C_R	0.183	-0.142	0.438	0.532
P_17_A_R	0.308	-0.019	0.515	0.476
P_17_B_R	0.203	0.040	0.489	0.470
P_17_D_R	0.210	-0.074	0.454	0.406
P_20_G_R	0.212	-0.300	-0.165	-0.803
P_20_K_R	0.121	-0.307	-0.340	-0.963
P_20_L_R	0.397	-0.270	-0.083	-0.611
P_20_I_R	-0.006	-0.203	-0.508	-0.963
P_20_E_R	-0.040	-0.199	-0.330	-0.825
P_20_J_R	-0.047	-0.025	-0.428	-0.899

EXPLORATORY FACTOR ANALYSIS WITH 5 FACTOR(S) :

MODEL FIT INFORMATION

Number of Free Parameters 155

Chi-Square Test of Model Fit

Value	4087.423*
Degrees of Freedom	373
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.035
90 Percent C.I.	0.034 0.036
Probability RMSEA <= .05	1.000

CFI/TLI

CFI	0.984
TLI	0.977

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value 0.123

MINIMUM ROTATION FUNCTION VALUE 1.04493

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1	2	3	4	5
P_18_C_R	0.493*	0.212*	0.150*	-0.154*	0.015
P_18_F_R	0.565*	0.053*	0.222*	-0.161*	0.010
P_18_G_R	0.629*	-0.075*	0.038*	-0.004	-0.012
P_18_I_R	0.639*	-0.067*	-0.050*	0.220*	-0.095*
P_18_J_R	0.625*	0.067*	-0.048*	0.112*	0.022
P_18_B_R	0.061*	0.760*	0.022	-0.057*	-0.051*
P_18_D_R	0.013	0.849*	-0.029*	-0.031*	-0.040*
P_18_E_R	-0.047*	0.630*	0.001	-0.001	0.050*
P_18_H_R	0.029*	0.657*	0.007	0.001	0.052*
P_18_K_R	0.258*	0.546*	-0.013	0.082*	0.026
P_1_A_R	0.009	-0.026*	0.121*	0.861*	0.024
P_1_B_R	-0.003	-0.018	0.046*	0.940*	-0.027*
P_2_A_R	-0.141*	0.122*	0.004	0.611*	0.031
P_2_C_R	0.003	-0.042*	-0.017	0.991*	-0.031
P_5_R_D	-0.032*	-0.006	0.920*	-0.005	-0.021
P_7_R	0.001	-0.055*	0.788*	0.213*	0.089*
P_8_R	-0.027*	-0.026*	0.968*	0.021	-0.038*
P_9_A_R	-0.025	-0.026	0.434*	-0.029	-0.042
P_9_B_R	-0.053	-0.050	0.458*	-0.055	-0.008
P_9_D_R	0.004	-0.036	0.796*	-0.016	0.038
P_3_R_D	-0.007	0.063*	0.109*	0.227*	-0.297*
P_10_R	0.068*	0.051*	0.503*	0.377*	-0.139*
P_11_R_D	0.146*	0.137*	0.475*	0.277*	-0.130*
P_9_C_R	-0.049*	0.002	0.620*	-0.034	-0.126*
P_17_A_R	0.071*	0.138*	0.483*	0.126*	-0.032
P_17_B_R	0.144*	0.006	0.487*	0.124*	-0.009
P_17_D_R	0.055	0.016	0.530*	0.091	0.054
P_20_G_R	-0.048*	0.134*	0.112*	0.010	0.914*
P_20_K_R	-0.075*	0.093*	-0.050*	-0.032*	0.931*
P_20_L_R	-0.100*	0.399*	0.004	0.037	0.666*
P_20_I_R	0.047*	-0.055*	-0.107*	-0.209*	0.831*
P_20_E_R	0.036*	-0.125*	0.016	-0.072*	0.832*
P_20_J_R	0.100*	-0.006	-0.372*	0.032*	0.716*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	1	2	3	4	5
1	1.000				
2	0.217*	1.000			
3	0.158*	0.176*	1.000		
4	0.181*	0.076*	0.472*	1.000	
5	-0.230*	0.019	-0.490*	-0.409*	1.000

ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.644	0.618	0.609	0.476	0.567

ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.396	0.286	0.609	0.554	0.561

ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.166	0.055	0.630	0.011	0.150
ESTIMATED RESIDUAL VARIANCES					
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.265	0.021	0.810	0.816	0.413
ESTIMATED RESIDUAL VARIANCES					
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.732	0.251	0.345	0.554	0.604
ESTIMATED RESIDUAL VARIANCES					
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.635	0.677	0.213	0.011	0.383
ESTIMATED RESIDUAL VARIANCES					
	P_20_I_R	P_20_E_R	P_20_J_R		
1	0.020	0.270	0.150		
S.E. GEOMIN ROTATED LOADINGS					
	1	2	3	4	5
P_18_C_R	0.013	0.015	0.023	0.022	0.015
P_18_F_R	0.012	0.013	0.023	0.023	0.015
P_18_G_R	0.011	0.013	0.017	0.013	0.016
P_18_I_R	0.013	0.012	0.013	0.024	0.018
P_18_J_R	0.012	0.013	0.014	0.023	0.017
P_18_B_R	0.010	0.008	0.014	0.014	0.014
P_18_D_R	0.008	0.008	0.014	0.012	0.013
P_18_E_R	0.012	0.011	0.018	0.016	0.018
P_18_H_R	0.010	0.010	0.016	0.014	0.016
P_18_K_R	0.012	0.011	0.014	0.016	0.014
P_1_A_R	0.013	0.013	0.025	0.013	0.016
P_1_B_R	0.010	0.010	0.017	0.011	0.013
P_2_A_R	0.038	0.034	0.038	0.028	0.042
P_2_C_R	0.009	0.010	0.023	0.012	0.017
P_5_R_D	0.014	0.014	0.016	0.016	0.019
P_7_R	0.015	0.018	0.020	0.023	0.023
P_8_R	0.011	0.012	0.013	0.011	0.017
P_9_A_R	0.036	0.035	0.046	0.049	0.046
P_9_B_R	0.040	0.038	0.045	0.047	0.044
P_9_D_R	0.022	0.023	0.025	0.029	0.026
P_3_R_D	0.017	0.017	0.025	0.023	0.023
P_10_R	0.016	0.015	0.022	0.022	0.022
P_11_R_D	0.018	0.017	0.022	0.024	0.023
P_9_C_R	0.025	0.024	0.030	0.031	0.031
P_17_A_R	0.034	0.035	0.040	0.042	0.038
P_17_B_R	0.036	0.035	0.041	0.046	0.040
P_17_D_R	0.050	0.049	0.052	0.058	0.052
P_20_G_R	0.013	0.021	0.026	0.015	0.016
P_20_K_R	0.012	0.016	0.020	0.016	0.012
P_20_L_R	0.019	0.020	0.018	0.020	0.019
P_20_I_R	0.011	0.013	0.022	0.020	0.014
P_20_E_R	0.012	0.018	0.016	0.021	0.014
P_20_J_R	0.015	0.008	0.021	0.016	0.015

S.E. GEOMIN FACTOR CORRELATIONS

1 2 3 4 5

1	0.000				
2	0.016	0.000			
3	0.024	0.017	0.000		
4	0.025	0.016	0.016	0.000	
5	0.021	0.019	0.017	0.016	0.000
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.011	0.012	0.012	0.012	0.011
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.010	0.011	0.012	0.011	0.010
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.010	0.009	0.024	0.008	0.014
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.016	0.013	0.024	0.027	0.025
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.014	0.013	0.014	0.023	0.026
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.028	0.035	0.013	0.009	0.019
S.E. ESTIMATED RESIDUAL VARIANCES					
	P_20_I_R	P_20_E_R	P_20_J_R		
1	0.010	0.012	0.011		
Est./S.E. GEOMIN ROTATED LOADINGS					
	1	2	3	4	5
P_18_C_R	38.994	14.395	6.597	-6.966	1.008
P_18_F_R	45.918	4.111	9.507	-7.074	0.678
P_18_G_R	55.407	-5.792	2.215	-0.353	-0.739
P_18_I_R	50.613	-5.612	-3.918	9.145	-5.144
P_18_J_R	51.691	5.300	-3.320	4.948	1.302
P_18_B_R	5.855	92.719	1.587	-4.184	-3.602
P_18_D_R	1.765	111.696	-2.155	-2.640	-3.181
P_18_E_R	-3.991	58.850	0.068	-0.058	2.830
P_18_H_R	2.887	68.246	0.459	0.097	3.215
P_18_K_R	21.464	51.526	-0.954	5.088	1.787
P_1_A_R	0.732	-2.064	4.861	65.052	1.450
P_1_B_R	-0.311	-1.794	2.674	86.200	-2.107
P_2_A_R	-3.744	3.573	0.093	21.938	0.727
P_2_C_R	0.381	-4.019	-0.731	83.734	-1.752
P_5_R_D	-2.329	-0.430	58.750	-0.287	-1.084
P_7_R	0.047	-3.063	39.670	9.253	3.774
P_8_R	-2.509	-2.170	74.006	1.938	-2.235

P_9_A_R	-0.699	-0.740	9.489	-0.588	-0.902
P_9_B_R	-1.329	-1.299	10.066	-1.171	-0.178
P_9_D_R	0.191	-1.573	31.516	-0.565	1.433
P_3_R_D	-0.434	3.679	4.361	9.991	-12.837
P_10_R	4.144	3.315	23.299	17.457	-6.354
P_11_R_D	8.101	8.005	21.358	11.737	-5.598
P_9_C_R	-1.973	0.103	20.985	-1.108	-4.078
P_17_A_R	2.094	3.948	12.010	2.995	-0.847
P_17_B_R	3.941	0.168	11.752	2.688	-0.218
P_17_D_R	1.103	0.323	10.144	1.565	1.027
P_20_G_R	-3.743	6.500	4.251	0.667	56.773
P_20_K_R	-6.162	5.842	-2.519	-1.995	75.493
P_20_L_R	-5.241	20.379	0.252	1.856	35.680
P_20_I_R	4.245	-4.358	-4.834	-10.484	59.110
P_20_E_R	2.884	-6.893	1.000	-3.417	58.728
P_20_J_R	6.510	-0.729	-17.468	1.990	46.439

Est./S.E. GEOMIN FACTOR CORRELATIONS

	1	2	3	4	5
1	0.000				
2	13.801	0.000			
3	6.670	10.639	0.000		
4	7.319	4.803	29.163	0.000	
5	-10.863	1.018	-28.340	-25.072	0.000

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	58.553	51.065	52.504	38.976	49.705

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	39.367	26.357	51.691	49.233	56.176

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	15.891	5.830	25.798	1.372	10.995

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	16.476	1.695	33.409	30.233	16.729

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	52.002	19.109	25.084	24.044	23.113

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	22.412	19.097	16.821	1.266	20.707

Est./S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	2.050	22.575	14.198

FACTOR STRUCTURE

	1	2	3	4	5
--	---	---	---	---	---

P_18_C_R	0.531	0.334	0.185	0.016	-0.105
P_18_F_R	0.580	0.202	0.239	0.046	-0.162
P_18_G_R	0.620	0.067	0.127	0.126	-0.174
P_18_I_R	0.678	0.078	0.190	0.346	-0.309
P_18_J_R	0.648	0.204	0.105	0.199	-0.143
P_18_B_R	0.230	0.772	0.163	0.043	-0.038
P_18_D_R	0.197	0.844	0.127	0.038	0.000
P_18_E_R	0.079	0.621	0.080	0.019	0.072
P_18_H_R	0.161	0.666	0.102	0.039	0.054
P_18_K_R	0.383	0.606	0.150	0.154	-0.050
P_1_A_R	0.173	0.063	0.512	0.908	-0.390
P_1_B_R	0.176	0.060	0.499	0.971	-0.433
P_2_A_R	-0.010	0.139	0.276	0.584	-0.186
P_2_C_R	0.178	0.031	0.459	0.993	-0.429
P_5_R_D	0.115	0.148	0.922	0.432	-0.462
P_7_R	0.131	0.101	0.836	0.545	-0.386
P_8_R	0.132	0.139	0.988	0.487	-0.515
P_9_A_R	0.042	0.042	0.433	0.187	-0.237
P_9_B_R	0.000	0.015	0.419	0.151	-0.198
P_9_D_R	0.110	0.104	0.764	0.342	-0.347
P_3_R_D	0.133	0.092	0.372	0.404	-0.440
P_10_R	0.259	0.180	0.768	0.687	-0.554
P_11_R_D	0.331	0.271	0.717	0.591	-0.507
P_9_C_R	0.072	0.096	0.659	0.302	-0.404
P_17_A_R	0.208	0.247	0.594	0.391	-0.334
P_17_B_R	0.246	0.132	0.573	0.384	-0.331
P_17_D_R	0.146	0.129	0.559	0.330	-0.256
P_20_G_R	-0.210	0.161	-0.315	-0.309	0.869
P_20_K_R	-0.283	0.083	-0.517	-0.443	0.988
P_20_L_R	-0.159	0.394	-0.250	-0.222	0.680
P_20_I_R	-0.211	-0.063	-0.615	-0.596	0.957
P_20_E_R	-0.193	-0.104	-0.441	-0.407	0.843
P_20_J_R	-0.119	-0.033	-0.693	-0.419	0.862

EXPLORATORY FACTOR ANALYSIS WITH 6 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters 183

Chi-Square Test of Model Fit

Value	2250.612*
Degrees of Freedom	345
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.026
90 Percent C.I.	0.025 0.027
Probability RMSEA <= .05	1.000

CFI/TLI

CFI	0.992
TLI	0.987

Chi-Square Test of Model Fit for the Baseline Model

Value 230934.785  
 Degrees of Freedom 528  
 P-Value 0.0000

SRMR (Standardized Root Mean Square Residual)

Value 0.121

MINIMUM ROTATION FUNCTION VALUE 0.89970

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1	2	3	4	5
P_18_C_R	0.621*	0.273*	0.017	0.003	0.035*
P_18_F_R	0.699*	0.095*	0.073*	0.097*	0.007
P_18_G_R	0.384*	-0.047*	0.414*	-0.019	-0.001
P_18_I_R	0.100*	-0.071*	0.684*	-0.021	0.069*
P_18_J_R	0.068*	0.056*	0.743*	0.007	-0.086*
P_18_B_R	0.140*	0.773*	-0.037*	-0.013	0.001
P_18_D_R	0.060*	0.845*	-0.010	-0.038*	-0.008
P_18_E_R	-0.072*	0.618*	0.039*	0.025	-0.034
P_18_H_R	0.023	0.651*	0.040*	0.012	-0.004
P_18_K_R	-0.013	0.545*	0.351*	0.024	-0.023
P_1_A_R	-0.005	-0.035*	0.012	0.109*	0.870*
P_1_B_R	0.008	-0.026*	-0.018	0.032*	0.952*
P_2_A_R	0.027	0.119*	-0.208*	-0.033	0.696*
P_2_C_R	-0.020	-0.052*	0.013	-0.029	0.996*
P_5_R_D	0.018	-0.002	-0.016	0.928*	-0.016
P_7_R	0.029	-0.054*	0.001	0.788*	0.214*
P_8_R	0.039*	-0.020	-0.031*	0.973*	0.018
P_9_A_R	0.063	-0.022	-0.072	0.442*	-0.017
P_9_B_R	0.021	-0.046	-0.069	0.466*	-0.048
P_9_D_R	0.085*	-0.029	-0.043	0.803*	-0.009
P_3_R_D	-0.034	0.061*	0.026	0.110*	0.222*
P_10_R	-0.002	0.049*	0.102*	0.498*	0.368*
P_11_R_D	0.060*	0.141*	0.144*	0.463*	0.270*
P_9_C_R	-0.047	-0.002	0.014	0.654*	-0.075*
P_17_A_R	-0.199*	0.139*	0.268*	0.477*	0.047
P_17_B_R	-0.165*	0.002	0.336*	0.476*	0.035
P_17_D_R	-0.103	0.013	0.161*	0.536*	0.041
P_20_G_R	-0.037*	0.124*	-0.020	0.109*	0.015
P_20_K_R	-0.107*	0.076*	0.006	-0.046*	-0.043*
P_20_L_R	-0.073*	0.387*	-0.046*	0.009	0.038
P_20_I_R	0.073*	-0.052*	-0.021	-0.121*	-0.188*
P_20_E_R	0.021	-0.127*	0.017	0.006	-0.066*
P_20_J_R	0.053*	-0.007	0.051*	-0.395*	0.054*

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	6
P_18_C_R	-0.008
P_18_F_R	0.000
P_18_G_R	-0.017
P_18_I_R	-0.086*
P_18_J_R	0.049
P_18_B_R	-0.052*
P_18_D_R	-0.033*
P_18_E_R	0.060*
P_18_H_R	0.060*
P_18_K_R	0.039*
P_1_A_R	0.028



P\_1\_B\_R -0.028\*  
 P\_2\_A\_R 0.029  
 P\_2\_C\_R -0.029  
 P\_5\_R\_D -0.012  
 P\_7\_R 0.097\*  
 P\_8\_R -0.030  
 P\_9\_A\_R -0.037  
 P\_9\_B\_R -0.004  
 P\_9\_D\_R 0.045  
 P\_3\_R\_D -0.293\*  
 P\_10\_R -0.127\*  
 P\_11\_R\_D -0.119\*  
 P\_9\_C\_R -0.114\*  
 P\_17\_A\_R -0.039  
 P\_17\_B\_R -0.014  
 P\_17\_D\_R 0.055  
 P\_20\_G\_R 0.914\*  
 P\_20\_K\_R 0.935\*  
 P\_20\_L\_R 0.670\*  
 P\_20\_I\_R 0.826\*  
 P\_20\_E\_R 0.828\*  
 P\_20\_J\_R 0.711\*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	1	2	3	4	5
1	1.000				
2	0.067	1.000			
3	0.244*	0.180*	1.000		
4	0.067*	0.164*	0.229*	1.000	
5	0.079*	0.080*	0.296*	0.478*	1.000
6	-0.125*	0.027	-0.281*	-0.487*	-0.419*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	6
6	1.000

ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.501	0.434	0.609	0.430	0.446

ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.383	0.290	0.602	0.556	0.516

ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.162	0.052	0.569	0.011	0.147

ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.263	0.018	0.801	0.811	0.401

ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.730	0.246	0.343	0.533	0.535

ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.564	0.654	0.215	0.002	0.383

ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	0.016	0.272	0.146

S.E. GEOMIN ROTATED LOADINGS

	1	2	3	4	5
P_18_C_R	0.016	0.034	0.010	0.010	0.016
P_18_F_R	0.018	0.036	0.019	0.020	0.014
P_18_G_R	0.020	0.012	0.017	0.015	0.014
P_18_I_R	0.026	0.017	0.017	0.013	0.023
P_18_J_R	0.025	0.020	0.017	0.013	0.026
P_18_B_R	0.016	0.012	0.010	0.013	0.012
P_18_D_R	0.015	0.008	0.009	0.014	0.012
P_18_E_R	0.015	0.011	0.013	0.019	0.018
P_18_H_R	0.012	0.010	0.012	0.016	0.016
P_18_K_R	0.010	0.014	0.013	0.014	0.016
P_1_A_R	0.015	0.012	0.014	0.024	0.014
P_1_B_R	0.012	0.010	0.012	0.016	0.011
P_2_A_R	0.030	0.034	0.042	0.038	0.028
P_2_C_R	0.015	0.011	0.009	0.022	0.013
P_5_R_D	0.014	0.013	0.015	0.016	0.020
P_7_R	0.017	0.018	0.016	0.020	0.025
P_8_R	0.013	0.010	0.012	0.013	0.016
P_9_A_R	0.041	0.034	0.040	0.046	0.050
P_9_B_R	0.039	0.037	0.041	0.047	0.051
P_9_D_R	0.028	0.022	0.026	0.026	0.029
P_3_R_D	0.018	0.017	0.019	0.025	0.024
P_10_R	0.015	0.015	0.019	0.021	0.023
P_11_R_D	0.017	0.017	0.020	0.023	0.024
P_9_C_R	0.026	0.024	0.025	0.030	0.035
P_17_A_R	0.037	0.036	0.036	0.042	0.034
P_17_B_R	0.043	0.024	0.039	0.043	0.036
P_17_D_R	0.054	0.045	0.052	0.051	0.057
P_20_G_R	0.014	0.020	0.015	0.026	0.016
P_20_K_R	0.015	0.017	0.011	0.019	0.017
P_20_L_R	0.020	0.020	0.019	0.019	0.021
P_20_I_R	0.014	0.012	0.012	0.022	0.022
P_20_E_R	0.013	0.018	0.015	0.016	0.022
P_20_J_R	0.015	0.008	0.015	0.021	0.020

S.E. GEOMIN ROTATED LOADINGS

	6
P_18_C_R	0.015
P_18_F_R	0.014
P_18_G_R	0.015
P_18_I_R	0.023
P_18_J_R	0.025
P_18_B_R	0.014
P_18_D_R	0.012
P_18_E_R	0.018
P_18_H_R	0.017
P_18_K_R	0.018
P_1_A_R	0.018

P\_1\_B\_R 0.013  
 P\_2\_A\_R 0.039  
 P\_2\_C\_R 0.019  
 P\_5\_R\_D 0.019  
 P\_7\_R 0.024  
 P\_8\_R 0.016  
 P\_9\_A\_R 0.046  
 P\_9\_B\_R 0.043  
 P\_9\_D\_R 0.026  
 P\_3\_R\_D 0.023  
 P\_10\_R 0.023  
 P\_11\_R\_D 0.024  
 P\_9\_C\_R 0.031  
 P\_17\_A\_R 0.032  
 P\_17\_B\_R 0.032  
 P\_17\_D\_R 0.052  
 P\_20\_G\_R 0.016  
 P\_20\_K\_R 0.012  
 P\_20\_L\_R 0.019  
 P\_20\_I\_R 0.015  
 P\_20\_E\_R 0.014  
 P\_20\_J\_R 0.016

S.E. GEOMIN FACTOR CORRELATIONS

	1	2	3	4	5
1	0.000				
2	0.045	0.000			
3	0.029	0.029	0.000		
4	0.022	0.017	0.021	0.000	
5	0.024	0.016	0.025	0.017	0.000
6	0.023	0.019	0.025	0.017	0.017

S.E. GEOMIN FACTOR CORRELATIONS

	6
6	0.000

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.017	0.020	0.012	0.014	0.015

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.010	0.011	0.012	0.011	0.011

S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.011	0.010	0.028	0.008	0.014

S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.016	0.013	0.025	0.028	0.025

S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.014	0.013	0.014	0.024	0.030

S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.031	0.036	0.013	0.009	0.018

S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
1	0.010	0.012	0.011

Est./S.E. GEOMIN ROTATED LOADINGS

	1	2	3	4	5
P_18_C_R	37.747	8.095	1.820	0.289	2.217
P_18_F_R	37.961	2.657	3.850	4.800	0.530
P_18_G_R	19.362	-3.787	25.077	-1.314	-0.036
P_18_I_R	3.798	-4.176	40.012	-1.666	3.011
P_18_J_R	2.719	2.832	42.883	0.550	-3.358
P_18_B_R	8.620	66.375	-3.730	-1.036	0.089
P_18_D_R	4.038	99.640	-1.132	-2.688	-0.678
P_18_E_R	-4.820	53.988	2.955	1.341	-1.906
P_18_H_R	1.932	67.907	3.272	0.720	-0.280
P_18_K_R	-1.329	38.407	26.592	1.752	-1.421
P_1_A_R	-0.324	-2.808	0.890	4.473	63.559
P_1_B_R	0.708	-2.628	-1.601	1.987	85.713
P_2_A_R	0.878	3.473	-4.984	-0.876	24.608
P_2_C_R	-1.320	-4.956	1.369	-1.301	78.213
P_5_R_D	1.227	-0.156	-1.083	57.599	-0.810
P_7_R	1.638	-3.060	0.064	39.446	8.578
P_8_R	3.004	-1.953	-2.653	72.292	1.148
P_9_A_R	1.514	-0.645	-1.791	9.596	-0.336
P_9_B_R	0.546	-1.236	-1.691	9.966	-0.947
P_9_D_R	3.006	-1.347	-1.629	31.350	-0.312
P_3_R_D	-1.839	3.650	1.347	4.352	9.220
P_10_R	-0.154	3.200	5.325	23.288	16.260
P_11_R_D	3.586	8.349	7.079	20.540	11.096
P_9_C_R	-1.805	-0.071	0.570	21.563	-2.144
P_17_A_R	-5.312	3.908	7.356	11.490	1.376
P_17_B_R	-3.827	0.092	8.639	11.203	0.992
P_17_D_R	-1.916	0.294	3.109	10.418	0.711
P_20_G_R	-2.653	6.088	-1.387	4.130	0.942
P_20_K_R	-7.135	4.490	0.542	-2.355	-2.559
P_20_L_R	-3.672	19.434	-2.445	0.485	1.825
P_20_I_R	5.119	-4.410	-1.708	-5.429	-8.654
P_20_E_R	1.603	-7.248	1.133	0.375	-3.008
P_20_J_R	3.633	-0.791	3.296	-18.672	2.710

Est./S.E. GEOMIN ROTATED LOADINGS

	6
P_18_C_R	-0.552
P_18_F_R	0.027
P_18_G_R	-1.104
P_18_I_R	-3.660
P_18_J_R	1.953
P_18_B_R	-3.662
P_18_D_R	-2.642
P_18_E_R	3.298
P_18_H_R	3.554
P_18_K_R	2.213
P_1_A_R	1.565

P\_1\_B\_R -2.069  
 P\_2\_A\_R 0.739  
 P\_2\_C\_R -1.519  
 P\_5\_R\_D -0.656  
 P\_7\_R 4.081  
 P\_8\_R -1.908  
 P\_9\_A\_R -0.815  
 P\_9\_B\_R -0.097  
 P\_9\_D\_R 1.725  
 P\_3\_R\_D -12.542  
 P\_10\_R -5.587  
 P\_11\_R\_D -4.997  
 P\_9\_C\_R -3.647  
 P\_17\_A\_R -1.233  
 P\_17\_B\_R -0.427  
 P\_17\_D\_R 1.056  
 P\_20\_G\_R 57.671  
 P\_20\_K\_R 75.924  
 P\_20\_L\_R 35.410  
 P\_20\_I\_R 56.073  
 P\_20\_E\_R 57.945  
 P\_20\_J\_R 45.305

	Est./S.E. GEOMIN FACTOR CORRELATIONS				
	1	2	3	4	5
1	0.000				
2	1.505	0.000			
3	8.403	6.287	0.000		
4	3.071	9.770	10.978	0.000	
5	3.305	4.913	11.701	28.416	0.000
6	-5.455	1.435	-11.048	-28.260	-24.586

Est./S.E. GEOMIN FACTOR CORRELATIONS  
6

6	0.000
---	-------

	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	29.275	21.522	52.215	31.193	29.240

	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	38.082	27.268	50.578	49.608	48.471

	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	15.441	5.392	20.601	1.308	10.629

	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	16.314	1.434	31.623	29.258	15.856

	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	51.596	18.719	25.023	21.927	18.012

	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	<u>18.000</u>	<u>18.214</u>	<u>16.932</u>	<u>0.241</u>	<u>20.754</u>

	Est./S.E. ESTIMATED RESIDUAL VARIANCES		
	P_20_I_R	P_20_E_R	P_20_J_R
1	<u>1.608</u>	<u>22.673</u>	<u>13.486</u>

	FACTOR STRUCTURE				
	1	2	3	4	5
P_18_C_R	<u>0.647</u>	<u>0.321</u>	<u>0.231</u>	<u>0.114</u>	<u>0.116</u>
P_18_F_R	0.730	0.172	0.284	0.179	0.138
P_18_G_R	0.483	0.050	0.500	0.101	0.146
P_18_I_R	0.277	0.059	0.736	0.205	0.300
P_18_J_R	0.241	0.190	0.732	0.126	0.127
P_18_B_R	0.188	0.772	0.148	0.140	0.078
P_18_D_R	0.116	0.839	0.155	0.114	0.056
P_18_E_R	-0.030	0.623	0.111	0.084	0.007
P_18_H_R	0.070	0.663	0.148	0.098	0.042
P_18_K_R	0.104	0.611	0.434	0.163	0.118
P_1_A_R	0.068	0.055	0.279	0.508	0.911
P_1_B_R	0.083	0.051	0.276	0.493	0.972
P_2_A_R	0.033	0.135	0.010	0.259	0.618
P_2_C_R	0.059	0.022	0.295	0.455	0.992
P_5_R_D	0.076	0.146	0.199	0.923	0.429
P_7_R	0.083	0.097	0.215	0.837	0.549
P_8_R	0.101	0.137	0.212	0.989	0.489
P_9_A_R	0.077	0.040	0.046	0.436	0.192
P_9_B_R	0.029	0.016	0.022	0.423	0.154
P_9_D_R	0.120	0.101	0.141	0.767	0.348
P_3_R_D	0.038	0.091	0.202	0.372	0.407
P_10_R	0.104	0.175	0.369	0.767	0.694
P_11_R_D	0.172	0.265	0.403	0.710	0.600
P_9_C_R	0.008	0.096	0.162	0.674	0.286
P_17_A_R	-0.084	0.255	0.379	0.589	0.366
P_17_B_R	-0.046	0.132	0.420	0.566	0.356
P_17_D_R	-0.031	0.128	0.257	0.561	0.315
P_20_G_R	-0.140	0.162	-0.234	-0.317	-0.315
P_20_K_R	-0.224	0.085	-0.292	-0.516	-0.457
P_20_L_R	-0.138	0.396	-0.169	-0.251	-0.226
P_20_I_R	-0.062	-0.063	-0.327	-0.621	-0.596
P_20_E_R	-0.092	-0.104	-0.251	-0.444	-0.413
P_20_J_R	-0.046	-0.035	-0.211	-0.702	-0.415

FACTOR STRUCTURE  
6

P_18_C_R	<u>-0.100</u>
P_18_F_R	-0.155
P_18_G_R	-0.173
P_18_I_R	-0.311
P_18_J_R	-0.134
P_18_B_R	-0.032
P_18_D_R	0.008
P_18_E_R	0.078
P_18_H_R	0.060
P_18_K_R	-0.045
P_1_A_R	-0.393

P_1_B_R	-0.439
P_2_A_R	-0.188
P_2_C_R	-0.435
P_5_R_D	-0.456
P_7_R	-0.382
P_8_R	-0.509
P_9_A_R	-0.234
P_9_B_R	-0.196
P_9_D_R	-0.341
P_3_R_D	-0.441
P_10_R	-0.551
P_11_R_D	-0.502
P_9_C_R	-0.399
P_17_A_R	-0.338
P_17_B_R	-0.334
P_17_D_R	-0.255
P_20_G_R	0.869
P_20_K_R	0.990
P_20_L_R	0.682
P_20_I_R	0.959
P_20_E_R	0.842
P_20_J_R	0.860

EXPLORATORY FACTOR ANALYSIS WITH 7 FACTOR(S) :

MODEL FIT INFORMATION

Number of Free Parameters 210

Chi-Square Test of Model Fit

Value	1429.704*
Degrees of Freedom	318
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.021	
90 Percent C.I.	0.020	0.022
Probability RMSEA <= .05	1.000	

CFI/TLI

CFI	0.995
TLI	0.992

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.122
-------	-------

MINIMUM ROTATION FUNCTION VALUE 0.79257

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	1	2	3	4	5
P_18_C_R	0.659*	0.185*	0.006	0.009	-0.024
P_18_F_R	0.714*	0.014	0.070*	-0.011	0.088*
P_18_G_R	0.358*	-0.050*	0.412*	0.012	-0.005
P_18_I_R	0.048*	-0.015	0.671*	0.103*	-0.028
P_18_J_R	0.016	0.121*	0.730*	-0.049*	-0.002
P_18_B_R	0.190*	0.741*	-0.031*	-0.005	-0.006
P_18_D_R	0.112*	0.817*	-0.010	-0.020	-0.056*
P_18_E_R	-0.036*	0.607*	0.020	-0.052*	-0.047*
P_18_H_R	0.056*	0.644*	0.045*	-0.005	0.006
P_18_K_R	-0.023*	0.597*	0.377*	0.013	0.077*
P_1_A_R	-0.012	-0.015	0.033*	0.864*	0.071*
P_1_B_R	0.007	-0.014	0.002	0.941*	-0.002
P_2_A_R	0.043	0.115*	-0.176*	0.698*	-0.023
P_2_C_R	-0.038*	-0.021*	0.053*	0.993*	-0.022
P_5_R_D	-0.001	0.024	0.003	0.023	0.843*
P_7_R	0.015	-0.036*	0.007	0.231*	0.690*
P_8_R	0.017	0.010	0.003	0.054*	0.947*
P_9_A_R	0.025	0.016	0.008	0.013	0.636*
P_9_B_R	0.025	-0.053	-0.079	-0.053	0.419*
P_9_D_R	0.069*	-0.022	-0.024	-0.005	0.790*
P_3_R_D	-0.015	0.043*	-0.009	0.201*	0.012
P_10_R	0.044*	-0.017	-0.022	0.303*	0.193*
P_11_R_D	0.130*	0.065*	0.009	0.166*	0.127*
P_9_C_R	-0.078*	0.036	0.057*	-0.051	0.725*
P_17_A_R	-0.044	0.070*	0.003	-0.011	-0.016
P_17_B_R	-0.014	-0.065	0.085*	-0.004	0.024
P_17_D_R	-0.038	-0.013	0.040	0.019	0.272*
P_20_G_R	-0.023	0.131*	-0.032*	0.013	0.044*
P_20_K_R	-0.093*	0.086*	-0.021	-0.054*	-0.117*
P_20_L_R	-0.061*	0.426*	-0.006	0.085*	0.091*
P_20_I_R	0.071*	-0.040*	0.000	-0.161*	-0.060*
P_20_E_R	0.020	-0.117*	0.007	-0.060*	-0.021
P_20_J_R	0.079*	-0.024*	0.017	0.033	-0.427*

GEOMIN ROTATED LOADINGS (\* significant at 5% level)

	6	7
P_18_C_R	0.005	0.028
P_18_F_R	0.012	-0.009
P_18_G_R	-0.019	-0.023
P_18_I_R	-0.096*	0.041*
P_18_J_R	0.030*	0.040*
P_18_B_R	-0.056*	0.003
P_18_D_R	-0.036*	0.053*
P_18_E_R	0.057*	0.143*
P_18_H_R	0.050*	0.022
P_18_K_R	0.014	-0.073*
P_1_A_R	0.023	0.050*
P_1_B_R	-0.029*	0.048*
P_2_A_R	0.026	-0.052
P_2_C_R	-0.037	-0.022
P_5_R_D	-0.032	0.077*
P_7_R	0.083*	0.120*
P_8_R	-0.042*	0.023
P_9_A_R	-0.046	-0.268*
P_9_B_R	-0.002	0.086
P_9_D_R	0.040	0.042
P_3_R_D	-0.277*	0.199*
P_10_R	-0.096*	0.563*



P_11_R_D	-0.090*	0.628*
P_9_C_R	-0.117*	-0.065
P_17_A_R	-0.010	0.833*
P_17_B_R	0.014	0.748*
P_17_D_R	0.069	0.448*
P_20_G_R	0.885*	0.026
P_20_K_R	0.911*	0.052*
P_20_L_R	0.643*	-0.207*
P_20_I_R	0.806*	-0.191*
P_20_E_R	0.804*	-0.037
P_20_J_R	0.701*	0.006

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	1	2	3	4	5
1	1.000				
2	0.136*	1.000			
3	0.329*	0.049*	1.000		
4	0.130*	0.042*	0.202*	1.000	
5	0.118*	0.113*	0.149*	0.457*	1.000
6	-0.169*	0.058*	-0.229*	-0.404*	-0.422*
7	0.146*	0.174*	0.257*	0.390*	0.530*

GEOMIN FACTOR CORRELATIONS (\* significant at 5% level)

	6	7
6	1.000	
7	-0.360*	1.000

ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.492	0.435	0.608	0.428	0.440

ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.382	0.291	0.593	0.556	0.492

ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.161	0.053	0.554	0.005	0.162

ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.266	-0.017	0.677	0.806	0.362

ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.722	0.165	0.223	0.464	0.304

ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.408	0.616	0.218	-0.006	0.345

ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R

1            0.004            0.274            0.136

S.E. GEOMIN ROTATED LOADINGS

	1	2	3	4	5
P_18_C_R	0.016	0.017	0.010	0.012	0.013
P_18_F_R	0.018	0.009	0.018	0.012	0.022
P_18_G_R	0.014	0.011	0.014	0.013	0.016
P_18_I_R	0.014	0.009	0.014	0.019	0.015
P_18_J_R	0.011	0.018	0.013	0.016	0.014
P_18_B_R	0.016	0.009	0.010	0.012	0.013
P_18_D_R	0.017	0.009	0.010	0.012	0.016
P_18_E_R	0.014	0.012	0.013	0.019	0.020
P_18_H_R	0.014	0.010	0.013	0.015	0.017
P_18_K_R	0.009	0.012	0.013	0.011	0.018
P_1_A_R	0.015	0.013	0.015	0.014	0.024
P_1_B_R	0.011	0.010	0.011	0.012	0.015
P_2_A_R	0.035	0.035	0.041	0.029	0.041
P_2_C_R	0.016	0.010	0.011	0.012	0.022
P_5_R_D	0.016	0.014	0.015	0.020	0.023
P_7_R	0.019	0.017	0.018	0.027	0.025
P_8_R	0.013	0.010	0.012	0.016	0.017
P_9_A_R	0.038	0.032	0.034	0.045	0.060
P_9_B_R	0.040	0.038	0.041	0.052	0.059
P_9_D_R	0.029	0.023	0.026	0.032	0.031
P_3_R_D	0.019	0.017	0.019	0.026	0.027
P_10_R	0.021	0.015	0.016	0.034	0.039
P_11_R_D	0.025	0.019	0.016	0.038	0.040
P_9_C_R	0.029	0.024	0.027	0.036	0.034
P_17_A_R	0.028	0.033	0.025	0.030	0.027
P_17_B_R	0.030	0.036	0.040	0.033	0.032
P_17_D_R	0.052	0.043	0.049	0.053	0.060
P_20_G_R	0.014	0.020	0.016	0.016	0.021
P_20_K_R	0.016	0.017	0.012	0.018	0.023
P_20_L_R	0.018	0.020	0.015	0.027	0.030
P_20_I_R	0.015	0.011	0.012	0.024	0.021
P_20_E_R	0.014	0.018	0.015	0.022	0.018
P_20_J_R	0.017	0.010	0.014	0.018	0.022

S.E. GEOMIN ROTATED LOADINGS

	6	7
P_18_C_R	0.012	0.017
P_18_F_R	0.012	0.015
P_18_G_R	0.014	0.018
P_18_I_R	0.019	0.017
P_18_J_R	0.015	0.018
P_18_B_R	0.014	0.013
P_18_D_R	0.012	0.018
P_18_E_R	0.018	0.025
P_18_H_R	0.016	0.018
P_18_K_R	0.011	0.019
P_1_A_R	0.017	0.024
P_1_B_R	0.013	0.017
P_2_A_R	0.037	0.053
P_2_C_R	0.019	0.021
P_5_R_D	0.017	0.027
P_7_R	0.022	0.033
P_8_R	0.012	0.016
P_9_A_R	0.041	0.072
P_9_B_R	0.042	0.069
P_9_D_R	0.026	0.035
P_3_R_D	0.023	0.030

P_10_R	0.029	0.031
P_11_R_D	0.032	0.035
P_9_C_R	0.031	0.039
P_17_A_R	0.024	0.043
P_17_B_R	0.026	0.047
P_17_D_R	0.051	0.059
P_20_G_R	0.014	0.021
P_20_K_R	0.013	0.021
P_20_L_R	0.022	0.034
P_20_I_R	0.015	0.025
P_20_E_R	0.015	0.023
P_20_J_R	0.015	0.018

S.E. GEOMIN FACTOR CORRELATIONS

	1	2	3	4	5
1	0.000				
2	0.022	0.000			
3	0.024	0.021	0.000		
4	0.019	0.018	0.021	0.000	
5	0.023	0.018	0.020	0.017	0.000
6	0.019	0.020	0.021	0.018	0.018
7	0.026	0.025	0.022	0.030	0.028

S.E. GEOMIN FACTOR CORRELATIONS

	6	7
6	0.000	
7	0.027	0.000

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	0.017	0.020	0.012	0.014	0.015

S.E. ESTIMATED RESIDUAL VARIANCES

	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	0.010	0.011	0.012	0.011	0.012

S.E. ESTIMATED RESIDUAL VARIANCES

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	0.011	0.010	0.029	0.009	0.015

S.E. ESTIMATED RESIDUAL VARIANCES

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	0.016	0.015	0.046	0.029	0.027

S.E. ESTIMATED RESIDUAL VARIANCES

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	0.015	0.017	0.022	0.028	0.049

S.E. ESTIMATED RESIDUAL VARIANCES

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	0.046	0.038	0.013	0.010	0.021

S.E. ESTIMATED RESIDUAL VARIANCES

	P_20_I_R	P_20_E_R	P_20_J_R
--	----------	----------	----------

1	0.010	0.012	0.011		
Est./S.E. GEOMIN ROTATED LOADINGS					
	1	2	3	4	5
P_18_C_R	40.146	11.188	0.571	0.692	-1.866
P_18_F_R	40.258	1.499	3.947	-0.865	4.068
P_18_G_R	24.822	-4.703	29.819	0.875	-0.291
P_18_I_R	3.462	-1.626	49.347	5.364	-1.907
P_18_J_R	1.466	6.879	54.818	-3.165	-0.109
P_18_B_R	11.751	80.635	-2.938	-0.446	-0.442
P_18_D_R	6.773	95.048	-1.020	-1.598	-3.437
P_18_E_R	-2.561	51.135	1.499	-2.734	-2.304
P_18_H_R	3.948	65.597	3.458	-0.354	0.330
P_18_K_R	-2.575	48.333	28.035	1.185	4.200
P_1_A_R	-0.813	-1.228	2.238	63.122	2.917
P_1_B_R	0.616	-1.325	0.184	78.865	-0.156
P_2_A_R	1.224	3.302	-4.259	24.134	-0.566
P_2_C_R	-2.388	-2.221	4.762	79.649	-1.010
P_5_R_D	-0.069	1.728	0.209	1.146	37.064
P_7_R	0.803	-2.074	0.422	8.663	27.235
P_8_R	1.341	0.996	0.252	3.423	56.273
P_9_A_R	0.666	0.496	0.220	0.281	10.640
P_9_B_R	0.641	-1.395	-1.917	-1.017	7.122
P_9_D_R	2.352	-0.956	-0.921	-0.164	25.568
P_3_R_D	-0.768	2.563	-0.491	7.748	0.443
P_10_R	2.080	-1.108	-1.404	8.856	4.963
P_11_R_D	5.148	3.355	0.569	4.360	3.155
P_9_C_R	-2.665	1.503	2.090	-1.419	21.496
P_17_A_R	-1.547	2.103	0.137	-0.381	-0.600
P_17_B_R	-0.479	-1.799	2.160	-0.114	0.738
P_17_D_R	-0.729	-0.298	0.814	0.352	4.532
P_20_G_R	-1.651	6.446	-1.983	0.836	2.092
P_20_K_R	-5.819	5.077	-1.761	-3.104	-5.161
P_20_L_R	-3.394	21.425	-0.396	3.177	3.017
P_20_I_R	4.649	-3.718	-0.014	-6.592	-2.853
P_20_E_R	1.414	-6.529	0.498	-2.668	-1.193
P_20_J_R	4.534	-2.320	1.235	1.880	-19.141

Est./S.E. GEOMIN ROTATED LOADINGS

	6	7
P_18_C_R	0.407	1.629
P_18_F_R	0.974	-0.591
P_18_G_R	-1.376	-1.260
P_18_I_R	-5.134	2.417
P_18_J_R	2.034	2.291
P_18_B_R	-3.905	0.201
P_18_D_R	-2.944	2.954
P_18_E_R	3.190	5.745
P_18_H_R	3.172	1.200
P_18_K_R	1.204	-3.862
P_1_A_R	1.315	2.107
P_1_B_R	-2.174	2.832
P_2_A_R	0.703	-0.984
P_2_C_R	-1.933	-1.068
P_5_R_D	-1.885	2.872
P_7_R	3.801	3.611
P_8_R	-3.462	1.444
P_9_A_R	-1.117	-3.741
P_9_B_R	-0.056	1.252
P_9_D_R	1.565	1.206

P_3_R_D	-11.807	6.647
P_10_R	-3.326	18.075
P_11_R_D	-2.807	17.898
P_9_C_R	-3.732	-1.655
P_17_A_R	-0.425	19.221
P_17_B_R	0.535	15.894
P_17_D_R	1.353	7.547
P_20_G_R	63.978	1.213
P_20_K_R	70.085	2.462
P_20_L_R	28.647	-6.106
P_20_I_R	52.047	-7.524
P_20_E_R	55.060	-1.613
P_20_J_R	46.698	0.322

	Est./S.E. GEOMIN FACTOR CORRELATIONS				
	1	2	3	4	5
1	0.000				
2	6.270	0.000			
3	13.708	2.307	0.000		
4	6.735	2.347	9.516	0.000	
5	5.159	6.218	7.317	26.304	0.000
6	-8.752	2.858	-11.079	-22.662	-23.184
7	5.540	7.083	11.669	13.096	19.154
	Est./S.E. GEOMIN FACTOR CORRELATIONS				
	6	7			
6	0.000				
7	-13.549	0.000			
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
1	28.521	21.928	51.893	31.021	28.989
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
1	37.800	27.205	47.919	49.716	42.763
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
1	15.265	5.470	19.337	0.589	11.126
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
1	16.531	-1.097	14.659	28.228	13.293
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
1	49.515	9.564	10.269	16.644	6.157
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
1	8.941	16.020	17.115	-0.640	16.760
	Est./S.E. ESTIMATED RESIDUAL VARIANCES				
	P_20_I_R	P_20_E_R	P_20_J_R		

1	0.364	22.884	11.972		
FACTOR STRUCTURE					
	1	2	3	4	5
P_18_C_R	0.687	0.278	0.236	0.101	0.091
P_18_F_R	0.744	0.123	0.311	0.129	0.169
P_18_G_R	0.487	0.014	0.527	0.136	0.095
P_18_I_R	0.299	0.028	0.736	0.286	0.185
P_18_J_R	0.267	0.165	0.735	0.108	0.109
P_18_B_R	0.290	0.761	0.080	0.065	0.118
P_18_D_R	0.225	0.832	0.077	0.037	0.082
P_18_E_R	0.053	0.624	0.045	-0.016	0.048
P_18_H_R	0.153	0.661	0.089	0.029	0.079
P_18_K_R	0.180	0.609	0.391	0.113	0.160
P_1_A_R	0.120	0.039	0.221	0.911	0.485
P_1_B_R	0.140	0.033	0.212	0.971	0.465
P_2_A_R	0.077	0.131	-0.038	0.632	0.249
P_2_C_R	0.107	0.010	0.240	0.995	0.437
P_5_R_D	0.122	0.131	0.162	0.453	0.911
P_7_R	0.127	0.080	0.172	0.561	0.823
P_8_R	0.148	0.122	0.177	0.516	1.005
P_9_A_R	0.075	0.043	0.056	0.223	0.525
P_9_B_R	0.048	0.006	0.001	0.158	0.427
P_9_D_R	0.150	0.085	0.117	0.359	0.795
P_3_R_D	0.091	0.068	0.145	0.394	0.328
P_10_R	0.195	0.115	0.249	0.650	0.670
P_11_R_D	0.286	0.209	0.290	0.527	0.598
P_9_C_R	0.034	0.090	0.141	0.306	0.720
P_17_A_R	0.087	0.206	0.204	0.308	0.428
P_17_B_R	0.114	0.071	0.269	0.306	0.416
P_17_D_R	0.061	0.097	0.171	0.292	0.488
P_20_G_R	-0.154	0.188	-0.219	-0.318	-0.302
P_20_K_R	-0.255	0.119	-0.271	-0.469	-0.503
P_20_L_R	-0.121	0.433	-0.174	-0.205	-0.212
P_20_I_R	-0.126	-0.030	-0.254	-0.582	-0.571
P_20_E_R	-0.145	-0.079	-0.201	-0.410	-0.417
P_20_J_R	-0.082	-0.017	-0.174	-0.430	-0.695

FACTOR STRUCTURE					
	6	7			
P_18_C_R	-0.100	0.146			
P_18_F_R	-0.153	0.154			
P_18_G_R	-0.171	0.136			
P_18_I_R	-0.303	0.278			
P_18_J_R	-0.127	0.221			
P_18_B_R	-0.034	0.166			
P_18_D_R	0.007	0.185			
P_18_E_R	0.083	0.183			
P_18_H_R	0.060	0.137			
P_18_K_R	-0.045	0.165			
P_1_A_R	-0.381	0.420			
P_1_B_R	-0.428	0.423			
P_2_A_R	-0.188	0.179			
P_2_C_R	-0.428	0.372			
P_5_R_D	-0.424	0.549			
P_7_R	-0.351	0.544			
P_8_R	-0.475	0.566			
P_9_A_R	-0.228	0.099			
P_9_B_R	-0.178	0.262			
P_9_D_R	-0.314	0.444			
P_3_R_D	-0.428	0.386			

P_10_R	-0.506	0.816
P_11_R_D	-0.457	0.825
P_9_C_R	-0.377	0.351
P_17_A_R	-0.288	0.830
P_17_B_R	-0.284	0.762
P_17_D_R	-0.217	0.577
P_20_G_R	0.871	-0.253
P_20_K_R	0.989	-0.363
P_20_L_R	0.681	-0.293
P_20_I_R	0.951	-0.573
P_20_E_R	0.838	-0.377
P_20_J_R	0.847	-0.448

DIAGRAM INFORMATION

Mplus diagrams are currently not available for EFA.  
No diagram output was produced.

Beginning Time: 18:14:36  
Ending Time: 18:14:54  
Elapsed Time: 00:00:18

MUTHEN & MUTHEN  
3463 Stoner Ave.  
Los Angeles, CA 90066

Tel: (310) 391-9971  
Fax: (310) 391-8971  
Web: [www.StatModel.com](http://www.StatModel.com)  
Support: [Support@StatModel.com](mailto:Support@StatModel.com)

Copyright (c) 1998-2015 Muthen & Muthen

## ANEXO 2: Análisis Factorial Confirmatorio. Salida M-PLUS

Mplus VERSION 7.4

MUTHEN & MUTHEN

11/28/2016 12:47 PM

### INPUT INSTRUCTIONS

TITLE: Análisis Factorial Confirmatorio del Modelo de Medida para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

DATA: FILE IS enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat ;  
FORMAT IS FREE ;

VARIABLE: NAMES ARE id\_user edad form niv\_sni cand niv\_I niv\_II niv\_III area\_sni f\_m\_ct b\_q m\_cs h\_cc cs b\_ca ing gen\_sni mujer hombre tip\_inst cpi\_e\_n cpi\_cs\_h cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ot\_inst ori\_inst t\_d\_k use\_ori p\_1\_a\_r p\_1\_b\_r p\_1\_ab\_r p\_2\_r p\_2\_a\_r p\_2\_b\_r p\_2\_c\_r p\_2\_d\_r p\_3\_r p\_3\_r\_d p\_5\_r p\_5\_r\_d p\_6\_a\_r p\_6\_b\_r p\_6\_c\_r p\_6\_d\_r p\_6\_e\_r p\_7\_r p\_8\_r p\_9\_r p\_9\_a\_r p\_9\_b\_r p\_9\_c\_r p\_9\_d\_r p\_9\_e\_r p\_10\_r p\_11\_r p\_11\_r\_d p\_13\_r p\_14\_a\_r p\_14\_b\_r p\_14\_c\_r p\_14\_d\_r p\_14\_e\_r p\_14\_f\_r p\_14\_g\_r p\_14\_h\_r p\_15\_r p\_15\_a\_r p\_15\_b\_r p\_15\_c\_r p\_16\_a\_r p\_16\_b\_r p\_16\_c\_r p\_16\_d\_r p\_16\_e\_r p\_16\_f\_r p\_16\_g\_r p\_17\_a\_r p\_17\_b\_r p\_17\_c\_r p\_17\_d\_r p\_17\_e\_r p\_17\_f\_r p\_18\_a\_r p\_18\_b\_r p\_18\_c\_r p\_18\_d\_r p\_18\_e\_r p\_18\_f\_r p\_18\_g\_r p\_18\_h\_r p\_18\_i\_r p\_18\_j\_r p\_18\_k\_r p\_19\_r p\_20\_a\_r p\_20\_b\_r p\_20\_c\_r p\_20\_d\_r p\_20\_e\_r p\_20\_f\_r p\_20\_g\_r p\_20\_h\_r p\_20\_i\_r p\_20\_j\_r p\_20\_k\_r p\_20\_l\_r p\_20\_m\_r p\_39\_r ;

USEVARIABLES ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ; !edad !cand !niv\_I !niv\_II !niv\_III !f\_m\_ct !b\_q !m\_cs !h\_cc !cs !b\_ca !ing !mujer !hombre !cpi\_e\_n !cpi\_cs\_h !cpi\_dt\_s !cpi\_sec !gobierno !inst\_n\_s !inst\_ut !univ ; !ot\_inst !t\_d\_k !use\_ori ;

MISSING ARE p\_1\_a\_r (2) p\_1\_b\_r (2) p\_1\_ab\_r (2) p\_2\_r (5) p\_2\_a\_r (2) p\_2\_b\_r (2) p\_2\_c\_r (2) p\_2\_d\_r (2) p\_3\_r\_d (4 5) p\_5\_r\_d (5) p\_6\_a\_r (2) p\_6\_b\_r (2) p\_6\_c\_r (2) p\_6\_d\_r (2) p\_6\_e\_r (2) p\_7\_r (2) p\_8\_r (2) p\_9\_r (6) p\_9\_a\_r (2) p\_9\_b\_r (2) p\_9\_c\_r (2) p\_9\_d\_r (2) p\_9\_e\_r (2) p\_10\_r (2) p\_11\_r\_d (5) p\_15\_r (4) p\_15\_a\_r (2) p\_15\_b\_r (2) p\_15\_c\_r (2) p\_16\_a\_r (2) p\_16\_b\_r (2) p\_16\_c\_r (2) p\_16\_d\_r (2) p\_16\_e\_r (2) p\_16\_f\_r (2) p\_16\_g\_r (2) p\_17\_a\_r (2) p\_17\_b\_r (2) p\_17\_c\_r (2) p\_17\_d\_r (2) p\_17\_e\_r (2) p\_17\_f\_r (2) p\_19\_r (2) p\_20\_a\_r (2) p\_20\_b\_r (2) p\_20\_c\_r (2) p\_20\_d\_r (2) p\_20\_e\_r (2) p\_20\_f\_r (2) p\_20\_g\_r (2) p\_20\_h\_r (2) p\_20\_i\_r (2) p\_20\_j\_r (2) p\_20\_k\_r (2) p\_20\_l\_r (2) p\_20\_m\_r (2) . ;

CATEGORICAL ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ;

AUXILIARY ARE id\_user ;

MODEL: PROP\_M1 BY p\_18\_c\_r\* p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r@1 p\_18\_j\_r ;  
PROP\_M2 BY p\_18\_b\_r\* p\_18\_d\_r@1 p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r ;



```

SATPROEV BY p_1_a_r* p_1_b_r p_2_a_r p_2_c_r@1 p_5_r_d p_7_r p_8_r p_9_a_r
p_9_b_r p_9_d_r ;
FLXEPSNI BY p_3_r_d* p_10_r@1 p_11_r_d p_9_c_r p_17_a_r p_17_b_r p_17_d_r;
DIRREF BY p_20_g_r* p_20_k_r p_20_l_r p_20_i_r@1 p_20_j_r p_20_e_r;

```

```

OUTPUT: stdyx;
SAVEDATA: DIFFTEST IS deriv_final.dat;

```

\*\*\* WARNING

Data set contains cases with missing on all variables. These cases were not included in the analysis. Number of cases with missing on all variables: 2  
 1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

Análisis Factorial Confirmatorio del Modelo de Medida para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	8053
Number of dependent variables	33
Number of independent variables	0
Number of continuous latent variables	5

Observed dependent variables

Binary and ordered categorical (ordinal)

P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R	P_18_B_R
P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R	P_1_A_R	P_1_B_R
P_2_A_R	P_2_C_R	P_5_R_D	P_7_R	P_8_R	P_9_A_R
P_9_B_R	P_9_D_R	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R
P_17_A_R	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_20_I_R	P_20_E_R	P_20_J_R			

Observed auxiliary variables

ID\_USER

Continuous latent variables

PROP_M1	PROP_M2	SATPROEV	FLXEPSNI	DIRREF
---------	---------	----------	----------	--------

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.100D-03
Parameterization	DELTA

Input data file(s)

enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat

Input data format FREE

SUMMARY OF DATA

Number of missing data patterns	1292
---------------------------------	------

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

	Covariance Coverage				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	0.972				
P_18_F_R	0.965	0.968			
P_18_G_R	0.969	0.966	0.972		
P_18_I_R	0.969	0.966	0.971	0.973	
P_18_J_R	0.965	0.963	0.967	0.969	0.970
P_18_B_R	0.970	0.967	0.971	0.971	0.968
P_18_D_R	0.968	0.965	0.970	0.970	0.966
P_18_E_R	0.968	0.966	0.970	0.971	0.967
P_18_H_R	0.967	0.965	0.969	0.970	0.966
P_18_K_R	0.968	0.965	0.969	0.971	0.967
P_1_A_R	0.834	0.831	0.834	0.835	0.832
P_1_B_R	0.849	0.846	0.850	0.851	0.847
P_2_A_R	0.948	0.944	0.949	0.950	0.946
P_2_C_R	0.948	0.944	0.949	0.950	0.946
P_5_R_D	0.849	0.846	0.850	0.851	0.847
P_7_R	0.670	0.667	0.670	0.671	0.668
P_8_R	0.614	0.612	0.614	0.615	0.613
P_9_A_R	0.613	0.612	0.614	0.615	0.613
P_9_B_R	0.613	0.612	0.614	0.615	0.613
P_9_D_R	0.613	0.612	0.614	0.615	0.613
P_3_R_D	0.904	0.901	0.905	0.906	0.902
P_10_R	0.834	0.830	0.834	0.834	0.831
P_11_R_D	0.864	0.860	0.865	0.865	0.862
P_9_C_R	0.613	0.612	0.614	0.615	0.613
P_17_A_R	0.910	0.907	0.911	0.912	0.908
P_17_B_R	0.910	0.907	0.911	0.912	0.908
P_17_D_R	0.910	0.907	0.911	0.912	0.908
P_20_G_R	0.760	0.757	0.761	0.762	0.760
P_20_K_R	0.767	0.765	0.768	0.769	0.766
P_20_L_R	0.729	0.726	0.730	0.730	0.728
P_20_I_R	0.778	0.775	0.779	0.780	0.777
P_20_E_R	0.768	0.766	0.770	0.770	0.768
P_20_J_R	0.728	0.725	0.729	0.730	0.728

	Covariance Coverage				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_B_R	0.975				
P_18_D_R	0.970	0.972			
P_18_E_R	0.971	0.969	0.973		
P_18_H_R	0.970	0.968	0.969	0.971	
P_18_K_R	0.970	0.969	0.969	0.969	0.972
P_1_A_R	0.836	0.834	0.834	0.833	0.834
P_1_B_R	0.851	0.849	0.850	0.849	0.849
P_2_A_R	0.951	0.948	0.949	0.948	0.948
P_2_C_R	0.951	0.948	0.949	0.948	0.948
P_5_R_D	0.852	0.849	0.850	0.849	0.849
P_7_R	0.672	0.670	0.670	0.669	0.670
P_8_R	0.615	0.614	0.615	0.614	0.614
P_9_A_R	0.615	0.614	0.615	0.614	0.614
P_9_B_R	0.615	0.614	0.615	0.614	0.614
P_9_D_R	0.615	0.614	0.615	0.614	0.614
P_3_R_D	0.907	0.904	0.905	0.904	0.905
P_10_R	0.835	0.833	0.834	0.833	0.833

P_11_R_D	0.866	0.864	0.865	0.864	0.864
P_9_C_R	0.615	0.614	0.615	0.614	0.614
P_17_A_R	0.913	0.910	0.911	0.910	0.910
P_17_B_R	0.913	0.910	0.911	0.910	0.910
P_17_D_R	0.913	0.910	0.911	0.910	0.910
P_20_G_R	0.762	0.761	0.761	0.760	0.761
P_20_K_R	0.769	0.768	0.768	0.768	0.768
P_20_L_R	0.730	0.730	0.730	0.729	0.730
P_20_I_R	0.780	0.779	0.779	0.778	0.779
P_20_E_R	0.770	0.769	0.770	0.769	0.770
P_20_J_R	0.730	0.729	0.729	0.728	0.729

Covariance Coverage

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
P_1_A_R	0.856				
P_1_B_R	0.752	0.873			
P_2_A_R	0.855	0.871	0.975		
P_2_C_R	0.855	0.871	0.975	0.975	
P_5_R_D	0.750	0.775	0.853	0.853	0.871
P_7_R	0.596	0.615	0.675	0.675	0.657
P_8_R	0.544	0.569	0.618	0.618	0.606
P_9_A_R	0.543	0.569	0.618	0.618	0.606
P_9_B_R	0.543	0.569	0.618	0.618	0.606
P_9_D_R	0.543	0.569	0.618	0.618	0.606
P_3_R_D	0.796	0.840	0.911	0.911	0.820
P_10_R	0.733	0.751	0.832	0.832	0.756
P_11_R_D	0.757	0.774	0.860	0.860	0.781
P_9_C_R	0.543	0.569	0.618	0.618	0.606
P_17_A_R	0.790	0.806	0.898	0.898	0.804
P_17_B_R	0.790	0.806	0.898	0.898	0.804
P_17_D_R	0.790	0.806	0.898	0.898	0.804
P_20_G_R	0.666	0.683	0.751	0.751	0.685
P_20_K_R	0.669	0.689	0.758	0.758	0.691
P_20_L_R	0.639	0.657	0.722	0.722	0.658
P_20_I_R	0.678	0.699	0.769	0.769	0.702
P_20_E_R	0.671	0.691	0.760	0.760	0.693
P_20_J_R	0.636	0.658	0.719	0.719	0.668

Covariance Coverage

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
P_7_R	0.685				
P_8_R	0.552	0.628			
P_9_A_R	0.552	0.628	0.628		
P_9_B_R	0.552	0.628	0.628	0.628	
P_9_D_R	0.552	0.628	0.628	0.628	0.628
P_3_R_D	0.647	0.598	0.597	0.597	0.597
P_10_R	0.605	0.560	0.560	0.560	0.560
P_11_R_D	0.621	0.570	0.570	0.570	0.570
P_9_C_R	0.552	0.628	0.628	0.628	0.628
P_17_A_R	0.634	0.581	0.581	0.581	0.581
P_17_B_R	0.634	0.581	0.581	0.581	0.581
P_17_D_R	0.634	0.581	0.581	0.581	0.581
P_20_G_R	0.556	0.520	0.520	0.520	0.520
P_20_K_R	0.560	0.527	0.526	0.526	0.526
P_20_L_R	0.539	0.505	0.505	0.505	0.505
P_20_I_R	0.569	0.534	0.534	0.534	0.534
P_20_E_R	0.565	0.529	0.528	0.528	0.528
P_20_J_R	0.548	0.518	0.518	0.518	0.518

Covariance Coverage

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
--	---------	--------	----------	---------	----------

P_3_R_D	0.930				
P_10_R	0.793	0.848			
P_11_R_D	0.822	0.830	0.879		
P_9_C_R	0.597	0.560	0.570	0.628	
P_17_A_R	0.856	0.801	0.829	0.581	0.919
P_17_B_R	0.856	0.801	0.829	0.581	0.919
P_17_D_R	0.856	0.801	0.829	0.581	0.919
P_20_G_R	0.722	0.683	0.699	0.520	0.724
P_20_K_R	0.728	0.690	0.707	0.526	0.729
P_20_L_R	0.693	0.655	0.672	0.505	0.693
P_20_I_R	0.738	0.697	0.715	0.534	0.739
P_20_E_R	0.729	0.691	0.707	0.528	0.730
P_20_J_R	0.693	0.657	0.671	0.518	0.693

Covariance Coverage

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_B_R	0.919				
P_17_D_R	0.919	0.919			
P_20_G_R	0.724	0.724	0.764		
P_20_K_R	0.729	0.729	0.737	0.771	
P_20_L_R	0.693	0.693	0.704	0.709	0.733
P_20_I_R	0.739	0.739	0.741	0.750	0.714
P_20_E_R	0.730	0.730	0.738	0.742	0.708
P_20_J_R	0.693	0.693	0.701	0.710	0.677

Covariance Coverage

	P_20_I_R	P_20_E_R	P_20_J_R
P_20_I_R	0.782		
P_20_E_R	0.751	0.773	
P_20_J_R	0.717	0.706	0.732

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

P_18_C_R		
Category 1	0.011	83.000
Category 2	0.041	317.000
Category 3	0.180	1405.000
Category 4	0.467	3658.000
Category 5	0.302	2362.000
P_18_F_R		
Category 1	0.025	193.000
Category 2	0.088	690.000
Category 3	0.242	1884.000
Category 4	0.452	3527.000
Category 5	0.193	1503.000
P_18_G_R		
Category 1	0.134	1047.000
Category 2	0.265	2073.000
Category 3	0.321	2511.000
Category 4	0.226	1768.000
Category 5	0.055	432.000
P_18_I_R		
Category 1	0.048	376.000
Category 2	0.105	825.000
Category 3	0.223	1751.000
Category 4	0.373	2925.000
Category 5	0.250	1962.000
P_18_J_R		
Category 1	0.060	472.000
Category 2	0.175	1365.000

Category 3	0.315	2462.000
Category 4	0.330	2574.000
Category 5	0.120	936.000
P_18_B_R		
Category 1	0.031	240.000
Category 2	0.069	538.000
Category 3	0.198	1550.000
Category 4	0.363	2851.000
Category 5	0.340	2669.000
P_18_D_R		
Category 1	0.020	155.000
Category 2	0.052	410.000
Category 3	0.154	1206.000
Category 4	0.364	2850.000
Category 5	0.410	3206.000
P_18_E_R		
Category 1	0.008	61.000
Category 2	0.027	213.000
Category 3	0.122	956.000
Category 4	0.367	2877.000
Category 5	0.476	3725.000
P_18_H_R		
Category 1	0.016	125.000
Category 2	0.055	427.000
Category 3	0.165	1293.000
Category 4	0.380	2973.000
Category 5	0.384	3005.000
P_18_K_R		
Category 1	0.066	518.000
Category 2	0.146	1141.000
Category 3	0.285	2228.000
Category 4	0.334	2613.000
Category 5	0.170	1327.000
P_1_A_R		
Category 1	0.174	1202.000
Category 2	0.826	5692.000
P_1_B_R		
Category 1	0.237	1669.000
Category 2	0.763	5361.000
P_2_A_R		
Category 1	0.027	213.000
Category 2	0.973	7641.000
P_2_C_R		
Category 1	0.196	1542.000
Category 2	0.804	6312.000
P_5_R_D		
Category 1	0.356	2500.000
Category 2	0.644	4514.000
P_7_R		
Category 1	0.219	1208.000
Category 2	0.781	4308.000
P_8_R		
Category 1	0.381	1926.000
Category 2	0.619	3131.000
P_9_A_R		
Category 1	0.044	224.000
Category 2	0.956	4830.000
P_9_B_R		
Category 1	0.045	225.000
Category 2	0.955	4829.000
P_9_D_R		
Category 1	0.137	693.000
Category 2	0.863	4361.000

P_3_R_D			
Category 1	0.506		3786.000
Category 2	0.494		3703.000
P_10_R			
Category 1	0.451		3083.000
Category 2	0.549		3749.000
P_11_R_D			
Category 1	0.530		3753.000
Category 2	0.470		3326.000
P_9_C_R			
Category 1	0.142		719.000
Category 2	0.858		4335.000
P_17_A_R			
Category 1	0.033		242.000
Category 2	0.967		7158.000
P_17_B_R			
Category 1	0.026		189.000
Category 2	0.974		7211.000
P_17_D_R			
Category 1	0.011		85.000
Category 2	0.989		7315.000
P_20_G_R			
Category 1	0.485		2982.000
Category 2	0.515		3172.000
P_20_K_R			
Category 1	0.330		2047.000
Category 2	0.670		4162.000
P_20_L_R			
Category 1	0.701		4134.000
Category 2	0.299		1767.000
P_20_I_R			
Category 1	0.310		1953.000
Category 2	0.690		4345.000
P_20_E_R			
Category 1	0.428		2661.000
Category 2	0.572		3562.000
P_20_J_R			
Category 1	0.439		2588.000
Category 2	0.561		3306.000

WARNING: THE BIVARIATE TABLE OF P\_2\_C\_R AND P\_2\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_A\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_D\_R HAS AN EMPTY CELL.

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 106

Chi-Square Test of Model Fit

Value	10890.792*
Degrees of Freedom	485
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.052	
90 Percent C.I.	0.051	0.052
Probability RMSEA <= .05	0.001	

CFI/TLI

CFI	0.955
TLI	0.951

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

WRMR (Weighted Root Mean Square Residual)

Value	4.775
-------	-------

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.823	0.019	42.219	0.000
P_18_F_R	0.844	0.019	43.401	0.000
P_18_G_R	0.783	0.019	42.331	0.000
P_18_I_R	1.000	0.000	999.000	999.000
P_18_J_R	0.882	0.020	44.695	0.000
PROP_M2 BY				
P_18_B_R	0.937	0.011	83.477	0.000
P_18_D_R	1.000	0.000	999.000	999.000
P_18_E_R	0.716	0.013	56.636	0.000
P_18_H_R	0.795	0.011	69.377	0.000
P_18_K_R	0.827	0.011	73.047	0.000
SATPROEV BY				
P_1_A_R	0.897	0.007	123.275	0.000
P_1_B_R	0.994	0.008	127.770	0.000
P_2_A_R	0.481	0.025	19.038	0.000
P_2_C_R	1.000	0.000	999.000	999.000
P_5_R_D	0.914	0.008	111.814	0.000
P_7_R	0.844	0.010	82.058	0.000
P_8_R	0.996	0.008	121.154	0.000
P_9_A_R	0.386	0.028	13.738	0.000
P_9_B_R	0.339	0.031	11.018	0.000
P_9_D_R	0.686	0.016	41.716	0.000
FLXEPSNI BY				
P_3_R_D	0.588	0.016	36.599	0.000

P_10_R	1.000	0.000	999.000	999.000
P_11_R_D	0.915	0.013	70.812	0.000
P_9_C_R	0.656	0.020	32.110	0.000
P_17_A_R	0.695	0.023	30.519	0.000
P_17_B_R	0.671	0.025	27.198	0.000
P_17_D_R	0.589	0.036	16.309	0.000
DIRREF BY				
P_20_G_R	0.831	0.008	107.720	0.000
P_20_K_R	0.986	0.008	120.106	0.000
P_20_L_R	0.637	0.011	56.008	0.000
P_20_I_R	1.000	0.000	999.000	999.000
P_20_J_R	0.921	0.008	122.170	0.000
P_20_E_R	0.840	0.009	96.921	0.000
PROP_M2 WITH				
PROP_M1	0.225	0.008	28.700	0.000
SATPROEV WITH				
PROP_M1	0.191	0.010	19.238	0.000
PROP_M2	0.092	0.011	8.148	0.000
FLXEPSNI WITH				
PROP_M1	0.253	0.010	24.664	0.000
PROP_M2	0.183	0.012	15.631	0.000
SATPROEV	0.680	0.010	69.984	0.000
DIRREF WITH				
PROP_M1	-0.216	0.011	-19.226	0.000
PROP_M2	0.037	0.013	2.875	0.004
SATPROEV	-0.561	0.011	-52.404	0.000
FLXEPSNI	-0.620	0.012	-52.064	0.000
Thresholds				
P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000
P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000
P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000
P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000
P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000



P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000
P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000

Variiances

PROP_M1	0.503	0.013	37.966	0.000
PROP_M2	0.677	0.011	63.784	0.000
SATPROEV	0.946	0.008	123.406	0.000
FLXEPSNI	0.861	0.014	60.991	0.000
DIRREF	0.978	0.010	101.423	0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.584	0.011	53.301	0.000
P_18_F_R	0.599	0.011	55.848	0.000
P_18_G_R	0.556	0.011	51.084	0.000
P_18_I_R	0.710	0.009	75.931	0.000
P_18_J_R	0.626	0.010	62.046	0.000
PROP_M2 BY				
P_18_B_R	0.771	0.007	114.790	0.000
P_18_D_R	0.823	0.006	127.568	0.000
P_18_E_R	0.589	0.010	61.525	0.000
P_18_H_R	0.654	0.008	77.680	0.000
P_18_K_R	0.680	0.008	84.493	0.000

SATPROEV BY				
P_1_A_R	0.872	0.006	143.130	0.000
P_1_B_R	0.967	0.004	222.082	0.000
P_2_A_R	0.468	0.024	19.821	0.000
P_2_C_R	0.972	0.004	246.812	0.000
P_5_R_D	0.889	0.007	127.724	0.000
P_7_R	0.821	0.009	87.335	0.000
P_8_R	0.968	0.007	141.638	0.000
P_9_A_R	0.376	0.027	13.769	0.000
P_9_B_R	0.329	0.030	11.043	0.000
P_9_D_R	0.667	0.016	42.503	0.000
FLXEPSNI BY				
P_3_R_D	0.546	0.014	38.337	0.000
P_10_R	0.928	0.008	121.983	0.000
P_11_R_D	0.849	0.009	98.690	0.000
P_9_C_R	0.609	0.018	33.088	0.000
P_17_A_R	0.645	0.020	31.618	0.000
P_17_B_R	0.623	0.022	27.783	0.000
P_17_D_R	0.547	0.033	16.414	0.000
DIRREF BY				
P_20_G_R	0.822	0.007	115.740	0.000
P_20_K_R	0.976	0.005	195.047	0.000
P_20_L_R	0.630	0.011	56.745	0.000
P_20_I_R	0.989	0.005	202.846	0.000
P_20_J_R	0.911	0.006	157.781	0.000
P_20_E_R	0.831	0.008	108.179	0.000
PROP_M2 WITH				
PROP_M1	0.385	0.012	31.809	0.000
SATPROEV WITH				
PROP_M1	0.276	0.014	19.981	0.000
PROP_M2	0.115	0.014	8.161	0.000
FLXEPSNI WITH				
PROP_M1	0.385	0.014	26.811	0.000
PROP_M2	0.240	0.015	15.943	0.000
SATPROEV	0.754	0.009	84.451	0.000
DIRREF WITH				
PROP_M1	-0.308	0.015	-20.058	0.000
PROP_M2	0.046	0.016	2.877	0.004
SATPROEV	-0.583	0.010	-56.271	0.000
FLXEPSNI	-0.675	0.011	-58.995	0.000
Thresholds				
P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000
P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000

P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000
P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000
P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000
P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000
P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000

Variiances

PROP_M1	1.000	0.000	999.000	999.000
PROP_M2	1.000	0.000	999.000	999.000
SATPROEV	1.000	0.000	999.000	999.000
FLXEPSNI	1.000	0.000	999.000	999.000
DIRREF	1.000	0.000	999.000	999.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Residual Variance
P_18_C_R	0.341	0.013	26.651	0.000	0.659
P_18_F_R	0.358	0.013	27.924	0.000	0.642

P_18_G_R	0.309	0.012	25.542	0.000	0.691
P_18_I_R	0.503	0.013	37.966	0.000	0.497
P_18_J_R	0.391	0.013	31.023	0.000	0.609
P_18_B_R	0.594	0.010	57.395	0.000	0.406
P_18_D_R	0.677	0.011	63.784	0.000	0.323
P_18_E_R	0.346	0.011	30.763	0.000	0.654
P_18_H_R	0.427	0.011	38.840	0.000	0.573
P_18_K_R	0.463	0.011	42.247	0.000	0.537
P_1_A_R	0.761	0.011	71.565	0.000	0.239
P_1_B_R	0.935	0.008	111.041	0.000	0.065
P_2_A_R	0.219	0.022	9.911	0.000	0.781
P_2_C_R	0.946	0.008	123.406	0.000	0.054
P_5_R_D	0.791	0.012	63.862	0.000	0.209
P_7_R	0.673	0.015	43.667	0.000	0.327
P_8_R	0.937	0.013	70.819	0.000	0.063
P_9_A_R	0.141	0.020	6.884	0.000	0.859
P_9_B_R	0.109	0.020	5.522	0.000	0.891
P_9_D_R	0.446	0.021	21.252	0.000	0.554
P_3_R_D	0.298	0.016	19.169	0.000	0.702
P_10_R	0.861	0.014	60.991	0.000	0.139
P_11_R_D	0.721	0.015	49.345	0.000	0.279
P_9_C_R	0.371	0.022	16.544	0.000	0.629
P_17_A_R	0.417	0.026	15.809	0.000	0.583
P_17_B_R	0.388	0.028	13.891	0.000	0.612
P_17_D_R	0.299	0.036	8.207	0.000	0.701
P_20_G_R	0.675	0.012	57.870	0.000	0.325
P_20_K_R	0.952	0.010	97.524	0.000	0.048
P_20_L_R	0.396	0.014	28.372	0.000	0.604
P_20_I_R	0.978	0.010	101.423	0.000	0.022
P_20_E_R	0.691	0.013	54.089	0.000	0.309
P_20_J_R	0.830	0.011	78.891	0.000	0.170

#### QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.226E-02  
 (ratio of smallest to largest eigenvalue)

#### SAVEDATA INFORMATION

Difference testing

Save file  
 deriv\_final.dat  
 Save format Free

#### DIAGRAM INFORMATION

Use View Diagram under the Diagram menu in the Mplus Editor to view the diagram.  
 If running Mplus from the Mplus Diagrammer, the diagram opens automatically.

Diagram output  
 c:\users\repodrida\documents\nata\documents\doctorado en ciencias  
 sociales\_uam\_x\2016\_capítulos\

Beginning Time: 12:47:05  
 Ending Time: 12:47:14  
 Elapsed Time: 00:00:09

MUTHEN & MUTHEN  
 3463 Stoner Ave.  
 Los Angeles, CA 90066

Tel: (310) 391-9971

Fax: (310) 391-8971  
Web: www.StatModel.com  
Support: Support@StatModel.com

Copyright (c) 1998-2015 Muthen & Muthen

## ANEXO 3: Prueba de la diferencia de la Ji-cuadrado entre el modelo multifactorial de cinco factores *versus* un modelo unidimensional. Salida M-PLUS

Mplus VERSION 7.4  
MUTHEN & MUTHEN  
11/28/2016 12:47 PM

### INPUT INSTRUCTIONS

TITLE: Análisis Factorial Confirmatorio del Modelo de Medida para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento. Prueba de la diferencia de la Chi-cuadrado entre el modelo multifactorial de cinco factores vs un modelo unidimensional.

DATA: FILE IS enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat ;  
FORMAT IS FREE ;

VARIABLE: NAMES ARE id\_user edad form niv\_sni cand niv\_I niv\_II niv\_III area\_sni f\_m\_ct b\_q m\_cs h\_cc cs b\_ca ing gen\_sni mujer hombre tip\_inst cpi\_e\_n cpi\_cs\_h cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ot\_inst ori\_inst t\_d\_k use\_ori p\_1\_a\_r p\_1\_b\_r p\_1\_ab\_r p\_2\_r p\_2\_a\_r p\_2\_b\_r p\_2\_c\_r p\_2\_d\_r p\_3\_r p\_3\_r\_d p\_5\_r p\_5\_r\_d p\_6\_a\_r p\_6\_b\_r p\_6\_c\_r p\_6\_d\_r p\_6\_e\_r p\_7\_r p\_8\_r p\_9\_r p\_9\_a\_r p\_9\_b\_r p\_9\_c\_r p\_9\_d\_r p\_9\_e\_r p\_10\_r p\_11\_r p\_11\_r\_d p\_13\_r p\_14\_a\_r p\_14\_b\_r p\_14\_c\_r p\_14\_d\_r p\_14\_e\_r p\_14\_f\_r p\_14\_g\_r p\_14\_h\_r p\_15\_r p\_15\_a\_r p\_15\_b\_r p\_15\_c\_r p\_16\_a\_r p\_16\_b\_r p\_16\_c\_r p\_16\_d\_r p\_16\_e\_r p\_16\_f\_r p\_16\_g\_r p\_17\_a\_r p\_17\_b\_r p\_17\_c\_r p\_17\_d\_r p\_17\_e\_r p\_17\_f\_r p\_18\_a\_r p\_18\_b\_r p\_18\_c\_r p\_18\_d\_r p\_18\_e\_r p\_18\_f\_r p\_18\_g\_r p\_18\_h\_r p\_18\_i\_r p\_18\_j\_r p\_18\_k\_r p\_19\_r p\_20\_a\_r p\_20\_b\_r p\_20\_c\_r p\_20\_d\_r p\_20\_e\_r p\_20\_f\_r p\_20\_g\_r p\_20\_h\_r p\_20\_i\_r p\_20\_j\_r p\_20\_k\_r p\_20\_l\_r p\_20\_m\_r p\_39\_r ;

USEVARIABLES ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ; !edad !cand !niv\_I !niv\_II !niv\_III !f\_m\_ct !b\_q !m\_cs !h\_cc !cs !b\_ca !ing !mujer !hombre !cpi\_e\_n !cpi\_cs\_h !cpi\_dt\_s !cpi\_sec !gobierno !inst\_n\_s !inst\_ut !univ ; !ot\_inst !t\_d\_k !use\_ori ;

MISSING ARE p\_1\_a\_r (2) p\_1\_b\_r (2) p\_1\_ab\_r (2) p\_2\_r (5) p\_2\_a\_r (2) p\_2\_b\_r (2) p\_2\_c\_r (2) p\_2\_d\_r (2) p\_3\_r\_d (4 5) p\_5\_r\_d (5) p\_6\_a\_r (2) p\_6\_b\_r (2) p\_6\_c\_r (2) p\_6\_d\_r (2) p\_6\_e\_r (2) p\_7\_r (2) p\_8\_r (2) p\_9\_r (6) p\_9\_a\_r (2) p\_9\_b\_r (2) p\_9\_c\_r (2) p\_9\_d\_r (2) p\_9\_e\_r (2) p\_10\_r (2) p\_11\_r\_d (5) p\_15\_r (4) p\_15\_a\_r (2) p\_15\_b\_r (2) p\_15\_c\_r (2) p\_16\_a\_r (2) p\_16\_b\_r (2) p\_16\_c\_r (2) p\_16\_d\_r (2) p\_16\_e\_r (2) p\_16\_f\_r (2) p\_16\_g\_r (2) p\_17\_a\_r (2) p\_17\_b\_r (2) p\_17\_c\_r (2) p\_17\_d\_r (2) p\_17\_e\_r (2) p\_17\_f\_r (2) p\_19\_r (2) p\_20\_a\_r (2) p\_20\_b\_r (2) p\_20\_c\_r (2) p\_20\_d\_r (2) p\_20\_e\_r (2) p\_20\_f\_r (2) p\_20\_g\_r (2) p\_20\_h\_r (2) p\_20\_i\_r (2) p\_20\_j\_r (2) p\_20\_k\_r (2) p\_20\_l\_r (2) p\_20\_m\_r (2) . ;

CATEGORICAL ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r

p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ;

AUXILIARY ARE id\_user ;

ANALYSIS: DIFFTEST IS deriv\_final.dat;

MODEL: f1 BY p\_18\_c\_r\* p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r  
p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r p\_2\_c\_r@1 p\_5\_r\_d p\_7\_r p\_8\_r  
p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r  
p\_17\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_j\_r p\_20\_e\_r;

OUTPUT: stdyx;

\*\*\* WARNING

Data set contains cases with missing on all variables. These cases were not included  
in the analysis. Number of cases with missing on all variables: 2

1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

Análisis Factorial Confirmatorio del Modelo de Medida para la relación entre la  
evaluación del desempeño individual de investigación de los investigadores SNI y  
la producción de conocimiento. Prueba de la diferencia de la Chi-cuadrado entre el  
modelo multifactorial de cinco factores vs un modelo unidimensional.

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	8053
Number of dependent variables	33
Number of independent variables	0
Number of continuous latent variables	1

Observed dependent variables

Binary and ordered categorical (ordinal)

P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R	P_18_B_R
P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R	P_1_A_R	P_1_B_R
P_2_A_R	P_2_C_R	P_5_R_D	P_7_R	P_8_R	P_9_A_R
P_9_B_R	P_9_D_R	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R
P_17_A_R	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_20_I_R	P_20_E_R	P_20_J_R			

Observed auxiliary variables

ID\_USER

Continuous latent variables

F1

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.100D-03
Parameterization	DELTA

Input data file(s)

enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat

Input data format FREE

SUMMARY OF DATA

Number of missing data patterns

1292

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

	Covariance Coverage				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	0.972				
P_18_F_R	0.965	0.968			
P_18_G_R	0.969	0.966	0.972		
P_18_I_R	0.969	0.966	0.971	0.973	
P_18_J_R	0.965	0.963	0.967	0.969	0.970
P_18_B_R	0.970	0.967	0.971	0.971	0.968
P_18_D_R	0.968	0.965	0.970	0.970	0.966
P_18_E_R	0.968	0.966	0.970	0.971	0.967
P_18_H_R	0.967	0.965	0.969	0.970	0.966
P_18_K_R	0.968	0.965	0.969	0.971	0.967
P_1_A_R	0.834	0.831	0.834	0.835	0.832
P_1_B_R	0.849	0.846	0.850	0.851	0.847
P_2_A_R	0.948	0.944	0.949	0.950	0.946
P_2_C_R	0.948	0.944	0.949	0.950	0.946
P_5_R_D	0.849	0.846	0.850	0.851	0.847
P_7_R	0.670	0.667	0.670	0.671	0.668
P_8_R	0.614	0.612	0.614	0.615	0.613
P_9_A_R	0.613	0.612	0.614	0.615	0.613
P_9_B_R	0.613	0.612	0.614	0.615	0.613
P_9_D_R	0.613	0.612	0.614	0.615	0.613
P_3_R_D	0.904	0.901	0.905	0.906	0.902
P_10_R	0.834	0.830	0.834	0.834	0.831
P_11_R_D	0.864	0.860	0.865	0.865	0.862
P_9_C_R	0.613	0.612	0.614	0.615	0.613
P_17_A_R	0.910	0.907	0.911	0.912	0.908
P_17_B_R	0.910	0.907	0.911	0.912	0.908
P_17_D_R	0.910	0.907	0.911	0.912	0.908
P_20_G_R	0.760	0.757	0.761	0.762	0.760
P_20_K_R	0.767	0.765	0.768	0.769	0.766
P_20_L_R	0.729	0.726	0.730	0.730	0.728
P_20_I_R	0.778	0.775	0.779	0.780	0.777
P_20_E_R	0.768	0.766	0.770	0.770	0.768
P_20_J_R	0.728	0.725	0.729	0.730	0.728

	Covariance Coverage				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_B_R	0.975				
P_18_D_R	0.970	0.972			
P_18_E_R	0.971	0.969	0.973		
P_18_H_R	0.970	0.968	0.969	0.971	
P_18_K_R	0.970	0.969	0.969	0.969	0.972
P_1_A_R	0.836	0.834	0.834	0.833	0.834
P_1_B_R	0.851	0.849	0.850	0.849	0.849
P_2_A_R	0.951	0.948	0.949	0.948	0.948
P_2_C_R	0.951	0.948	0.949	0.948	0.948
P_5_R_D	0.852	0.849	0.850	0.849	0.849
P_7_R	0.672	0.670	0.670	0.669	0.670
P_8_R	0.615	0.614	0.615	0.614	0.614
P_9_A_R	0.615	0.614	0.615	0.614	0.614
P_9_B_R	0.615	0.614	0.615	0.614	0.614
P_9_D_R	0.615	0.614	0.615	0.614	0.614

P_3_R_D	0.907	0.904	0.905	0.904	0.905
P_10_R	0.835	0.833	0.834	0.833	0.833
P_11_R_D	0.866	0.864	0.865	0.864	0.864
P_9_C_R	0.615	0.614	0.615	0.614	0.614
P_17_A_R	0.913	0.910	0.911	0.910	0.910
P_17_B_R	0.913	0.910	0.911	0.910	0.910
P_17_D_R	0.913	0.910	0.911	0.910	0.910
P_20_G_R	0.762	0.761	0.761	0.760	0.761
P_20_K_R	0.769	0.768	0.768	0.768	0.768
P_20_L_R	0.730	0.730	0.730	0.729	0.730
P_20_I_R	0.780	0.779	0.779	0.778	0.779
P_20_E_R	0.770	0.769	0.770	0.769	0.770
P_20_J_R	0.730	0.729	0.729	0.728	0.729

Covariance Coverage

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
P_1_A_R	0.856				
P_1_B_R	0.752	0.873			
P_2_A_R	0.855	0.871	0.975		
P_2_C_R	0.855	0.871	0.975	0.975	
P_5_R_D	0.750	0.775	0.853	0.853	0.871
P_7_R	0.596	0.615	0.675	0.675	0.657
P_8_R	0.544	0.569	0.618	0.618	0.606
P_9_A_R	0.543	0.569	0.618	0.618	0.606
P_9_B_R	0.543	0.569	0.618	0.618	0.606
P_9_D_R	0.543	0.569	0.618	0.618	0.606
P_3_R_D	0.796	0.840	0.911	0.911	0.820
P_10_R	0.733	0.751	0.832	0.832	0.756
P_11_R_D	0.757	0.774	0.860	0.860	0.781
P_9_C_R	0.543	0.569	0.618	0.618	0.606
P_17_A_R	0.790	0.806	0.898	0.898	0.804
P_17_B_R	0.790	0.806	0.898	0.898	0.804
P_17_D_R	0.790	0.806	0.898	0.898	0.804
P_20_G_R	0.666	0.683	0.751	0.751	0.685
P_20_K_R	0.669	0.689	0.758	0.758	0.691
P_20_L_R	0.639	0.657	0.722	0.722	0.658
P_20_I_R	0.678	0.699	0.769	0.769	0.702
P_20_E_R	0.671	0.691	0.760	0.760	0.693
P_20_J_R	0.636	0.658	0.719	0.719	0.668

Covariance Coverage

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
P_7_R	0.685				
P_8_R	0.552	0.628			
P_9_A_R	0.552	0.628	0.628		
P_9_B_R	0.552	0.628	0.628	0.628	
P_9_D_R	0.552	0.628	0.628	0.628	0.628
P_3_R_D	0.647	0.598	0.597	0.597	0.597
P_10_R	0.605	0.560	0.560	0.560	0.560
P_11_R_D	0.621	0.570	0.570	0.570	0.570
P_9_C_R	0.552	0.628	0.628	0.628	0.628
P_17_A_R	0.634	0.581	0.581	0.581	0.581
P_17_B_R	0.634	0.581	0.581	0.581	0.581
P_17_D_R	0.634	0.581	0.581	0.581	0.581
P_20_G_R	0.556	0.520	0.520	0.520	0.520
P_20_K_R	0.560	0.527	0.526	0.526	0.526
P_20_L_R	0.539	0.505	0.505	0.505	0.505
P_20_I_R	0.569	0.534	0.534	0.534	0.534
P_20_E_R	0.565	0.529	0.528	0.528	0.528
P_20_J_R	0.548	0.518	0.518	0.518	0.518



Covariance Coverage					
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
P_3_R_D	0.930				
P_10_R	0.793	0.848			
P_11_R_D	0.822	0.830	0.879		
P_9_C_R	0.597	0.560	0.570	0.628	
P_17_A_R	0.856	0.801	0.829	0.581	0.919
P_17_B_R	0.856	0.801	0.829	0.581	0.919
P_17_D_R	0.856	0.801	0.829	0.581	0.919
P_20_G_R	0.722	0.683	0.699	0.520	0.724
P_20_K_R	0.728	0.690	0.707	0.526	0.729
P_20_L_R	0.693	0.655	0.672	0.505	0.693
P_20_I_R	0.738	0.697	0.715	0.534	0.739
P_20_E_R	0.729	0.691	0.707	0.528	0.730
P_20_J_R	0.693	0.657	0.671	0.518	0.693

Covariance Coverage					
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_B_R	0.919				
P_17_D_R	0.919	0.919			
P_20_G_R	0.724	0.724	0.764		
P_20_K_R	0.729	0.729	0.737	0.771	
P_20_L_R	0.693	0.693	0.704	0.709	0.733
P_20_I_R	0.739	0.739	0.741	0.750	0.714
P_20_E_R	0.730	0.730	0.738	0.742	0.708
P_20_J_R	0.693	0.693	0.701	0.710	0.677

Covariance Coverage			
	P_20_I_R	P_20_E_R	P_20_J_R
P_20_I_R	0.782		
P_20_E_R	0.751	0.773	
P_20_J_R	0.717	0.706	0.732

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

P_18_C_R		
Category 1	0.011	83.000
Category 2	0.041	317.000
Category 3	0.180	1405.000
Category 4	0.467	3658.000
Category 5	0.302	2362.000
P_18_F_R		
Category 1	0.025	193.000
Category 2	0.088	690.000
Category 3	0.242	1884.000
Category 4	0.452	3527.000
Category 5	0.193	1503.000
P_18_G_R		
Category 1	0.134	1047.000
Category 2	0.265	2073.000
Category 3	0.321	2511.000
Category 4	0.226	1768.000
Category 5	0.055	432.000
P_18_I_R		
Category 1	0.048	376.000
Category 2	0.105	825.000
Category 3	0.223	1751.000
Category 4	0.373	2925.000
Category 5	0.250	1962.000
P_18_J_R		

Category 1	0.060	472.000
Category 2	0.175	1365.000
Category 3	0.315	2462.000
Category 4	0.330	2574.000
Category 5	0.120	936.000
P_18_B_R		
Category 1	0.031	240.000
Category 2	0.069	538.000
Category 3	0.198	1550.000
Category 4	0.363	2851.000
Category 5	0.340	2669.000
P_18_D_R		
Category 1	0.020	155.000
Category 2	0.052	410.000
Category 3	0.154	1206.000
Category 4	0.364	2850.000
Category 5	0.410	3206.000
P_18_E_R		
Category 1	0.008	61.000
Category 2	0.027	213.000
Category 3	0.122	956.000
Category 4	0.367	2877.000
Category 5	0.476	3725.000
P_18_H_R		
Category 1	0.016	125.000
Category 2	0.055	427.000
Category 3	0.165	1293.000
Category 4	0.380	2973.000
Category 5	0.384	3005.000
P_18_K_R		
Category 1	0.066	518.000
Category 2	0.146	1141.000
Category 3	0.285	2228.000
Category 4	0.334	2613.000
Category 5	0.170	1327.000
P_1_A_R		
Category 1	0.174	1202.000
Category 2	0.826	5692.000
P_1_B_R		
Category 1	0.237	1669.000
Category 2	0.763	5361.000
P_2_A_R		
Category 1	0.027	213.000
Category 2	0.973	7641.000
P_2_C_R		
Category 1	0.196	1542.000
Category 2	0.804	6312.000
P_5_R_D		
Category 1	0.356	2500.000
Category 2	0.644	4514.000
P_7_R		
Category 1	0.219	1208.000
Category 2	0.781	4308.000
P_8_R		
Category 1	0.381	1926.000
Category 2	0.619	3131.000
P_9_A_R		
Category 1	0.044	224.000
Category 2	0.956	4830.000
P_9_B_R		
Category 1	0.045	225.000
Category 2	0.955	4829.000
P_9_D_R		

Category 1	0.137	693.000
Category 2	0.863	4361.000
P_3_R_D		
Category 1	0.506	3786.000
Category 2	0.494	3703.000
P_10_R		
Category 1	0.451	3083.000
Category 2	0.549	3749.000
P_11_R_D		
Category 1	0.530	3753.000
Category 2	0.470	3326.000
P_9_C_R		
Category 1	0.142	719.000
Category 2	0.858	4335.000
P_17_A_R		
Category 1	0.033	242.000
Category 2	0.967	7158.000
P_17_B_R		
Category 1	0.026	189.000
Category 2	0.974	7211.000
P_17_D_R		
Category 1	0.011	85.000
Category 2	0.989	7315.000
P_20_G_R		
Category 1	0.485	2982.000
Category 2	0.515	3172.000
P_20_K_R		
Category 1	0.330	2047.000
Category 2	0.670	4162.000
P_20_L_R		
Category 1	0.701	4134.000
Category 2	0.299	1767.000
P_20_I_R		
Category 1	0.310	1953.000
Category 2	0.690	4345.000
P_20_E_R		
Category 1	0.428	2661.000
Category 2	0.572	3562.000
P_20_J_R		
Category 1	0.439	2588.000
Category 2	0.561	3306.000

WARNING: THE BIVARIATE TABLE OF P\_2\_C\_R AND P\_2\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_A\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_D\_R HAS AN EMPTY CELL.

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 96

Chi-Square Test of Model Fit

Value	48802.108*
Degrees of Freedom	495

P-Value 0.0000

Chi-Square Test for Difference Testing

Value 12119.064
Degrees of Freedom 10
P-Value 0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate 0.110
90 Percent C.I. 0.109 0.111
Probability RMSEA <= .05 0.000

CFI/TLI

CFI 0.790
TLI 0.776

Chi-Square Test of Model Fit for the Baseline Model

Value 230934.785
Degrees of Freedom 528
P-Value 0.0000

WRMR (Weighted Root Mean Square Residual)

Value 10.508

MODEL RESULTS

Table with 5 columns: Parameter Name, Estimate, S.E., Est./S.E., Two-Tailed P-Value. Rows include F1, BY, P\_18\_C\_R, P\_18\_F\_R, P\_18\_G\_R, P\_18\_I\_R, P\_18\_J\_R, P\_18\_B\_R, P\_18\_D\_R, P\_18\_E\_R, P\_18\_H\_R, P\_18\_K\_R, P\_1\_A\_R, P\_1\_B\_R, P\_2\_A\_R, P\_2\_C\_R, P\_5\_R\_D, P\_7\_R, P\_8\_R, P\_9\_A\_R, P\_9\_B\_R, P\_9\_D\_R, P\_3\_R\_D, P\_10\_R.

P_11_R_D	0.764	0.010	80.051	0.000
P_9_C_R	0.536	0.016	32.712	0.000
P_17_A_R	0.559	0.019	29.203	0.000
P_17_B_R	0.540	0.021	25.792	0.000
P_17_D_R	0.477	0.031	15.517	0.000
P_20_G_R	-0.754	0.009	-86.984	0.000
P_20_K_R	-0.987	0.007	-145.027	0.000
P_20_L_R	-0.481	0.012	-39.681	0.000
P_20_I_R	-0.999	0.006	-153.710	0.000
P_20_J_R	-0.868	0.007	-119.491	0.000
P_20_E_R	-0.782	0.009	-88.262	0.000

Thresholds

P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000
P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000
P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000
P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000
P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000
P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000
P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000

P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000
Variances				
F1	0.914	0.007	126.289	0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1				
BY				
P_18_C_R	0.285	0.011	25.301	0.000
P_18_F_R	0.316	0.011	28.738	0.000
P_18_G_R	0.292	0.011	26.847	0.000
P_18_I_R	0.434	0.010	44.649	0.000
P_18_J_R	0.333	0.011	31.600	0.000
P_18_B_R	0.308	0.011	27.977	0.000
P_18_D_R	0.295	0.011	26.734	0.000
P_18_E_R	0.148	0.012	12.189	0.000
P_18_H_R	0.200	0.012	17.347	0.000
P_18_K_R	0.307	0.011	28.948	0.000
P_1_A_R	0.816	0.006	129.357	0.000
P_1_B_R	0.945	0.004	226.268	0.000
P_2_A_R	0.387	0.022	17.536	0.000
P_2_C_R	0.956	0.004	252.578	0.000
P_5_R_D	0.820	0.007	123.365	0.000
P_7_R	0.742	0.009	78.762	0.000
P_8_R	0.893	0.007	133.667	0.000
P_9_A_R	0.310	0.024	13.093	0.000
P_9_B_R	0.266	0.026	10.292	0.000
P_9_D_R	0.567	0.014	39.965	0.000
P_3_R_D	0.458	0.012	37.249	0.000
P_10_R	0.777	0.007	104.041	0.000
P_11_R_D	0.730	0.009	84.721	0.000
P_9_C_R	0.512	0.016	32.841	0.000
P_17_A_R	0.534	0.018	29.529	0.000
P_17_B_R	0.517	0.020	25.942	0.000
P_17_D_R	0.456	0.029	15.558	0.000
P_20_G_R	-0.720	0.008	-92.054	0.000
P_20_K_R	-0.944	0.005	-172.977	0.000
P_20_L_R	-0.460	0.011	-40.221	0.000
P_20_I_R	-0.955	0.005	-193.777	0.000
P_20_J_R	-0.829	0.006	-134.516	0.000
P_20_E_R	-0.747	0.008	-93.721	0.000
Thresholds				
P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000

P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000
P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000
P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000
P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000
P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000
P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000
Variances				
F1	1.000	0.000	999.000	999.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Residual Variance
P_18_C_R	0.081	0.006	12.650	0.000	0.919
P_18_F_R	0.100	0.007	14.369	0.000	0.900
P_18_G_R	0.085	0.006	13.423	0.000	0.915
P_18_I_R	0.188	0.008	22.324	0.000	0.812
P_18_J_R	0.111	0.007	15.800	0.000	0.889
P_18_B_R	0.095	0.007	13.989	0.000	0.905
P_18_D_R	0.087	0.007	13.367	0.000	0.913
P_18_E_R	0.022	0.004	6.094	0.000	0.978
P_18_H_R	0.040	0.005	8.674	0.000	0.960
P_18_K_R	0.094	0.007	14.474	0.000	0.906
P_1_A_R	0.665	0.010	64.678	0.000	0.335
P_1_B_R	0.892	0.008	113.134	0.000	0.108
P_2_A_R	0.150	0.017	8.768	0.000	0.850
P_2_C_R	0.914	0.007	126.289	0.000	0.086
P_5_R_D	0.672	0.011	61.683	0.000	0.328
P_7_R	0.550	0.014	39.381	0.000	0.450
P_8_R	0.798	0.012	66.833	0.000	0.202
P_9_A_R	0.096	0.015	6.546	0.000	0.904
P_9_B_R	0.071	0.014	5.146	0.000	0.929
P_9_D_R	0.321	0.016	19.983	0.000	0.679
P_3_R_D	0.210	0.011	18.624	0.000	0.790
P_10_R	0.603	0.012	52.021	0.000	0.397
P_11_R_D	0.533	0.013	42.360	0.000	0.467
P_9_C_R	0.262	0.016	16.420	0.000	0.738
P_17_A_R	0.285	0.019	14.764	0.000	0.715
P_17_B_R	0.267	0.021	12.971	0.000	0.733
P_17_D_R	0.208	0.027	7.779	0.000	0.792
P_20_G_R	0.519	0.011	46.027	0.000	0.481
P_20_K_R	0.891	0.010	86.489	0.000	0.109
P_20_L_R	0.211	0.011	20.110	0.000	0.789
P_20_I_R	0.911	0.009	96.889	0.000	0.089
P_20_E_R	0.558	0.012	46.860	0.000	0.442
P_20_J_R	0.688	0.010	67.258	0.000	0.312

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.119E-02  
 (ratio of smallest to largest eigenvalue)

DIAGRAM INFORMATION

Use View Diagram under the Diagram menu in the Mplus Editor to view the diagram.  
 If running Mplus from the Mplus Diagrammer, the diagram opens automatically.

Diagram output

c:\users\repodrida\documents\nata\documents\doctorado en ciencias  
 sociales\_uam\_x\2016\_capitulos\

Beginning Time: 18:59:04  
 Ending Time: 18:59:12  
 Elapsed Time: 00:00:08

MUTHEN & MUTHEN  
 3463 Stoner Ave.  
 Los Angeles, CA 90066

Tel: (310) 391-9971  
 Fax: (310) 391-8971



Web: www.StatModel.com  
Support: Support@StatModel.com

Copyright (c) 1998-2015 Muthen & Muthen

## ANEXO 4: Modelo de Ecuaciones Estructurales. Salida M-PLUS

Mplus VERSION 7.4  
MUTHEN & MUTHEN  
02/13/2017 1:14 PM

### INPUT INSTRUCTIONS

TITLE: Modelo de Ecuaciones Estructurales para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

DATA: FILE IS enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat ;  
FORMAT IS FREE ;

VARIABLE: NAMES ARE id\_user edad form niv\_sni cand niv\_I niv\_II niv\_III area\_sni f\_m\_ct b\_q m\_cs h\_cc cs b\_ca ing gen\_sni mujer hombre tip\_inst cpi\_e\_n cpi\_cs\_h cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ot\_inst ori\_inst t\_d\_k use\_ori p\_1\_a\_r p\_1\_b\_r p\_1\_ab\_r p\_2\_r p\_2\_a\_r p\_2\_b\_r p\_2\_c\_r p\_2\_d\_r p\_3\_r p\_3\_r\_d p\_5\_r p\_5\_r\_d p\_6\_a\_r p\_6\_b\_r p\_6\_c\_r p\_6\_d\_r p\_6\_e\_r p\_7\_r p\_8\_r p\_9\_r p\_9\_a\_r p\_9\_b\_r p\_9\_c\_r p\_9\_d\_r p\_9\_e\_r p\_10\_r p\_11\_r p\_11\_r\_d p\_13\_r p\_14\_a\_r p\_14\_b\_r p\_14\_c\_r p\_14\_d\_r p\_14\_e\_r p\_14\_f\_r p\_14\_g\_r p\_14\_h\_r p\_15\_r p\_15\_a\_r p\_15\_b\_r p\_15\_c\_r p\_16\_a\_r p\_16\_b\_r p\_16\_c\_r p\_16\_d\_r p\_16\_e\_r p\_16\_f\_r p\_16\_g\_r p\_17\_a\_r p\_17\_b\_r p\_17\_c\_r p\_17\_d\_r p\_17\_e\_r p\_17\_f\_r p\_18\_a\_r p\_18\_b\_r p\_18\_c\_r p\_18\_d\_r p\_18\_e\_r p\_18\_f\_r p\_18\_g\_r p\_18\_h\_r p\_18\_i\_r p\_18\_j\_r p\_18\_k\_r p\_19\_r p\_20\_a\_r p\_20\_b\_r p\_20\_c\_r p\_20\_d\_r p\_20\_e\_r p\_20\_f\_r p\_20\_g\_r p\_20\_h\_r p\_20\_i\_r p\_20\_j\_r p\_20\_k\_r p\_20\_l\_r p\_20\_m\_r p\_39\_r ;

USEVARIABLES ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ; edad !cand !niv\_I !niv\_II !niv\_III !f\_m\_ct !b\_q !m\_cs !h\_cc !cs !b\_ca !ing !mujer !hombre !cpi\_e\_n !cpi\_cs\_h !cpi\_dt\_s !cpi\_sec !gobierno !inst\_n\_s !inst\_ut !univ ; !ot\_inst !t\_d\_k !use\_ori ;

MISSING ARE p\_1\_a\_r (2) p\_1\_b\_r (2) p\_1\_ab\_r (2) p\_2\_r (5) p\_2\_a\_r (2) p\_2\_b\_r (2) p\_2\_c\_r (2) p\_2\_d\_r (2) p\_3\_r\_d (4 5) p\_5\_r\_d (5) p\_6\_a\_r (2) p\_6\_b\_r (2) p\_6\_c\_r (2) p\_6\_d\_r (2) p\_6\_e\_r (2) p\_7\_r (2) p\_8\_r (2) p\_9\_r (6) p\_9\_a\_r (2) p\_9\_b\_r (2) p\_9\_c\_r (2) p\_9\_d\_r (2) p\_9\_e\_r (2) p\_10\_r (2) p\_11\_r\_d (5) p\_15\_r (4) p\_15\_a\_r (2) p\_15\_b\_r (2) p\_15\_c\_r (2) p\_16\_a\_r (2) p\_16\_b\_r (2) p\_16\_c\_r (2) p\_16\_d\_r (2) p\_16\_e\_r (2) p\_16\_f\_r (2) p\_16\_g\_r (2) p\_17\_a\_r (2) p\_17\_b\_r (2) p\_17\_c\_r (2) p\_17\_d\_r (2) p\_17\_e\_r (2) p\_17\_f\_r (2) p\_19\_r (2) p\_20\_a\_r (2) p\_20\_b\_r (2) p\_20\_c\_r (2) p\_20\_d\_r (2) p\_20\_e\_r (2) p\_20\_f\_r (2) p\_20\_g\_r (2) p\_20\_h\_r (2) p\_20\_i\_r (2) p\_20\_j\_r (2) p\_20\_k\_r (2) p\_20\_l\_r (2) p\_20\_m\_r (2). ;

CATEGORICAL ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ;

AUXILIARY ARE id\_user ;

MODEL: PROP\_M1 BY p\_18\_c\_r\* p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r@1 p\_18\_j\_r ;  
PROP\_M2 BY p\_18\_b\_r\* p\_18\_d\_r@1 p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r ;

```

SATPROEV BY p_1_a_r* p_1_b_r p_2_a_r p_2_c_r@1 p_5_r_d p_7_r p_8_r p_9_a_r
p_9_b_r p_9_d_r ;
FLXEPSNI BY p_3_r_d* p_10_r@1 p_11_r_d p_9_c_r p_17_a_r p_17_b_r p_17_d_r;
DIRREF BY p_20_g_r* p_20_k_r p_20_l_r p_20_i_r@1 p_20_j_r p_20_e_r;

```

```

FLXEPSNI ON PROP_M1 PROP_M2 ;
SATPROEV ON PROP_M1 PROP_M2 FLXEPSNI ;
DIRREF ON SATPROEV ;
OUTPUT: stdyx;

```

\*\*\* WARNING

Data set contains cases with missing on all variables. These cases were not included in the analysis. Number of cases with missing on all variables: 2  
 1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

Modelo de Ecuaciones Estructurales para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	8053
Number of dependent variables	33
Number of independent variables	0
Number of continuous latent variables	5

Observed dependent variables

Binary and ordered categorical (ordinal)

P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R	P_18_B_R
P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R	P_1_A_R	P_1_B_R
P_2_A_R	P_2_C_R	P_5_R_D	P_7_R	P_8_R	P_9_A_R
P_9_B_R	P_9_D_R	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R
P_17_A_R	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_20_I_R	P_20_E_R	P_20_J_R			

Observed auxiliary variables

ID\_USER

Continuous latent variables

PROP_M1	PROP_M2	SATPROEV	FLXEPSNI	DIRREF
---------	---------	----------	----------	--------

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.100D-03
Parameterization	DELTA

Input data file(s)

enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat

Input data format FREE

SUMMARY OF DATA

Number of missing data patterns	1292
---------------------------------	------

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

	Covariance Coverage				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	0.972				
P_18_F_R	0.965	0.968			
P_18_G_R	0.969	0.966	0.972		
P_18_I_R	0.969	0.966	0.971	0.973	
P_18_J_R	0.965	0.963	0.967	0.969	0.970
P_18_B_R	0.970	0.967	0.971	0.971	0.968
P_18_D_R	0.968	0.965	0.970	0.970	0.966
P_18_E_R	0.968	0.966	0.970	0.971	0.967
P_18_H_R	0.967	0.965	0.969	0.970	0.966
P_18_K_R	0.968	0.965	0.969	0.971	0.967
P_1_A_R	0.834	0.831	0.834	0.835	0.832
P_1_B_R	0.849	0.846	0.850	0.851	0.847
P_2_A_R	0.948	0.944	0.949	0.950	0.946
P_2_C_R	0.948	0.944	0.949	0.950	0.946
P_5_R_D	0.849	0.846	0.850	0.851	0.847
P_7_R	0.670	0.667	0.670	0.671	0.668
P_8_R	0.614	0.612	0.614	0.615	0.613
P_9_A_R	0.613	0.612	0.614	0.615	0.613
P_9_B_R	0.613	0.612	0.614	0.615	0.613
P_9_D_R	0.613	0.612	0.614	0.615	0.613
P_3_R_D	0.904	0.901	0.905	0.906	0.902
P_10_R	0.834	0.830	0.834	0.834	0.831
P_11_R_D	0.864	0.860	0.865	0.865	0.862
P_9_C_R	0.613	0.612	0.614	0.615	0.613
P_17_A_R	0.910	0.907	0.911	0.912	0.908
P_17_B_R	0.910	0.907	0.911	0.912	0.908
P_17_D_R	0.910	0.907	0.911	0.912	0.908
P_20_G_R	0.760	0.757	0.761	0.762	0.760
P_20_K_R	0.767	0.765	0.768	0.769	0.766
P_20_L_R	0.729	0.726	0.730	0.730	0.728
P_20_I_R	0.778	0.775	0.779	0.780	0.777
P_20_E_R	0.768	0.766	0.770	0.770	0.768
P_20_J_R	0.728	0.725	0.729	0.730	0.728

	Covariance Coverage				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_B_R	0.975				
P_18_D_R	0.970	0.972			
P_18_E_R	0.971	0.969	0.973		
P_18_H_R	0.970	0.968	0.969	0.971	
P_18_K_R	0.970	0.969	0.969	0.969	0.972
P_1_A_R	0.836	0.834	0.834	0.833	0.834
P_1_B_R	0.851	0.849	0.850	0.849	0.849
P_2_A_R	0.951	0.948	0.949	0.948	0.948
P_2_C_R	0.951	0.948	0.949	0.948	0.948
P_5_R_D	0.852	0.849	0.850	0.849	0.849
P_7_R	0.672	0.670	0.670	0.669	0.670
P_8_R	0.615	0.614	0.615	0.614	0.614
P_9_A_R	0.615	0.614	0.615	0.614	0.614
P_9_B_R	0.615	0.614	0.615	0.614	0.614
P_9_D_R	0.615	0.614	0.615	0.614	0.614
P_3_R_D	0.907	0.904	0.905	0.904	0.905
P_10_R	0.835	0.833	0.834	0.833	0.833
P_11_R_D	0.866	0.864	0.865	0.864	0.864
P_9_C_R	0.615	0.614	0.615	0.614	0.614

P_17_A_R	0.913	0.910	0.911	0.910	0.910
P_17_B_R	0.913	0.910	0.911	0.910	0.910
P_17_D_R	0.913	0.910	0.911	0.910	0.910
P_20_G_R	0.762	0.761	0.761	0.760	0.761
P_20_K_R	0.769	0.768	0.768	0.768	0.768
P_20_L_R	0.730	0.730	0.730	0.729	0.730
P_20_I_R	0.780	0.779	0.779	0.778	0.779
P_20_E_R	0.770	0.769	0.770	0.769	0.770
P_20_J_R	0.730	0.729	0.729	0.728	0.729

Covariance Coverage

	<u>P_1_A_R</u>	<u>P_1_B_R</u>	<u>P_2_A_R</u>	<u>P_2_C_R</u>	<u>P_5_R_D</u>
P_1_A_R	0.856				
P_1_B_R	0.752	0.873			
P_2_A_R	0.855	0.871	0.975		
P_2_C_R	0.855	0.871	0.975	0.975	
P_5_R_D	0.750	0.775	0.853	0.853	0.871
P_7_R	0.596	0.615	0.675	0.675	0.657
P_8_R	0.544	0.569	0.618	0.618	0.606
P_9_A_R	0.543	0.569	0.618	0.618	0.606
P_9_B_R	0.543	0.569	0.618	0.618	0.606
P_9_D_R	0.543	0.569	0.618	0.618	0.606
P_3_R_D	0.796	0.840	0.911	0.911	0.820
P_10_R	0.733	0.751	0.832	0.832	0.756
P_11_R_D	0.757	0.774	0.860	0.860	0.781
P_9_C_R	0.543	0.569	0.618	0.618	0.606
P_17_A_R	0.790	0.806	0.898	0.898	0.804
P_17_B_R	0.790	0.806	0.898	0.898	0.804
P_17_D_R	0.790	0.806	0.898	0.898	0.804
P_20_G_R	0.666	0.683	0.751	0.751	0.685
P_20_K_R	0.669	0.689	0.758	0.758	0.691
P_20_L_R	0.639	0.657	0.722	0.722	0.658
P_20_I_R	0.678	0.699	0.769	0.769	0.702
P_20_E_R	0.671	0.691	0.760	0.760	0.693
P_20_J_R	0.636	0.658	0.719	0.719	0.668

Covariance Coverage

	<u>P_7_R</u>	<u>P_8_R</u>	<u>P_9_A_R</u>	<u>P_9_B_R</u>	<u>P_9_D_R</u>
P_7_R	0.685				
P_8_R	0.552	0.628			
P_9_A_R	0.552	0.628	0.628		
P_9_B_R	0.552	0.628	0.628	0.628	
P_9_D_R	0.552	0.628	0.628	0.628	0.628
P_3_R_D	0.647	0.598	0.597	0.597	0.597
P_10_R	0.605	0.560	0.560	0.560	0.560
P_11_R_D	0.621	0.570	0.570	0.570	0.570
P_9_C_R	0.552	0.628	0.628	0.628	0.628
P_17_A_R	0.634	0.581	0.581	0.581	0.581
P_17_B_R	0.634	0.581	0.581	0.581	0.581
P_17_D_R	0.634	0.581	0.581	0.581	0.581
P_20_G_R	0.556	0.520	0.520	0.520	0.520
P_20_K_R	0.560	0.527	0.526	0.526	0.526
P_20_L_R	0.539	0.505	0.505	0.505	0.505
P_20_I_R	0.569	0.534	0.534	0.534	0.534
P_20_E_R	0.565	0.529	0.528	0.528	0.528
P_20_J_R	0.548	0.518	0.518	0.518	0.518

Covariance Coverage

	<u>P_3_R_D</u>	<u>P_10_R</u>	<u>P_11_R_D</u>	<u>P_9_C_R</u>	<u>P_17_A_R</u>
P_3_R_D	0.930				

P_10_R	0.793	0.848			
P_11_R_D	0.822	0.830	0.879		
P_9_C_R	0.597	0.560	0.570	0.628	
P_17_A_R	0.856	0.801	0.829	0.581	0.919
P_17_B_R	0.856	0.801	0.829	0.581	0.919
P_17_D_R	0.856	0.801	0.829	0.581	0.919
P_20_G_R	0.722	0.683	0.699	0.520	0.724
P_20_K_R	0.728	0.690	0.707	0.526	0.729
P_20_L_R	0.693	0.655	0.672	0.505	0.693
P_20_I_R	0.738	0.697	0.715	0.534	0.739
P_20_E_R	0.729	0.691	0.707	0.528	0.730
P_20_J_R	0.693	0.657	0.671	0.518	0.693

Covariance Coverage

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_B_R	0.919				
P_17_D_R	0.919	0.919			
P_20_G_R	0.724	0.724	0.764		
P_20_K_R	0.729	0.729	0.737	0.771	
P_20_L_R	0.693	0.693	0.704	0.709	0.733
P_20_I_R	0.739	0.739	0.741	0.750	0.714
P_20_E_R	0.730	0.730	0.738	0.742	0.708
P_20_J_R	0.693	0.693	0.701	0.710	0.677

Covariance Coverage

	P_20_I_R	P_20_E_R	P_20_J_R
P_20_I_R	0.782		
P_20_E_R	0.751	0.773	
P_20_J_R	0.717	0.706	0.732

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

P_18_C_R		
Category 1	0.011	83.000
Category 2	0.041	317.000
Category 3	0.180	1405.000
Category 4	0.467	3658.000
Category 5	0.302	2362.000
P_18_F_R		
Category 1	0.025	193.000
Category 2	0.088	690.000
Category 3	0.242	1884.000
Category 4	0.452	3527.000
Category 5	0.193	1503.000
P_18_G_R		
Category 1	0.134	1047.000
Category 2	0.265	2073.000
Category 3	0.321	2511.000
Category 4	0.226	1768.000
Category 5	0.055	432.000
P_18_I_R		
Category 1	0.048	376.000
Category 2	0.105	825.000
Category 3	0.223	1751.000
Category 4	0.373	2925.000
Category 5	0.250	1962.000
P_18_J_R		
Category 1	0.060	472.000
Category 2	0.175	1365.000
Category 3	0.315	2462.000
Category 4	0.330	2574.000

Category 5	0.120	936.000
P_18_B_R		
Category 1	0.031	240.000
Category 2	0.069	538.000
Category 3	0.198	1550.000
Category 4	0.363	2851.000
Category 5	0.340	2669.000
P_18_D_R		
Category 1	0.020	155.000
Category 2	0.052	410.000
Category 3	0.154	1206.000
Category 4	0.364	2850.000
Category 5	0.410	3206.000
P_18_E_R		
Category 1	0.008	61.000
Category 2	0.027	213.000
Category 3	0.122	956.000
Category 4	0.367	2877.000
Category 5	0.476	3725.000
P_18_H_R		
Category 1	0.016	125.000
Category 2	0.055	427.000
Category 3	0.165	1293.000
Category 4	0.380	2973.000
Category 5	0.384	3005.000
P_18_K_R		
Category 1	0.066	518.000
Category 2	0.146	1141.000
Category 3	0.285	2228.000
Category 4	0.334	2613.000
Category 5	0.170	1327.000
P_1_A_R		
Category 1	0.174	1202.000
Category 2	0.826	5692.000
P_1_B_R		
Category 1	0.237	1669.000
Category 2	0.763	5361.000
P_2_A_R		
Category 1	0.027	213.000
Category 2	0.973	7641.000
P_2_C_R		
Category 1	0.196	1542.000
Category 2	0.804	6312.000
P_5_R_D		
Category 1	0.356	2500.000
Category 2	0.644	4514.000
P_7_R		
Category 1	0.219	1208.000
Category 2	0.781	4308.000
P_8_R		
Category 1	0.381	1926.000
Category 2	0.619	3131.000
P_9_A_R		
Category 1	0.044	224.000
Category 2	0.956	4830.000
P_9_B_R		
Category 1	0.045	225.000
Category 2	0.955	4829.000
P_9_D_R		
Category 1	0.137	693.000
Category 2	0.863	4361.000
P_3_R_D		
Category 1	0.506	3786.000

Category 2	0.494	3703.000
P_10_R		
Category 1	0.451	3083.000
Category 2	0.549	3749.000
P_11_R_D		
Category 1	0.530	3753.000
Category 2	0.470	3326.000
P_9_C_R		
Category 1	0.142	719.000
Category 2	0.858	4335.000
P_17_A_R		
Category 1	0.033	242.000
Category 2	0.967	7158.000
P_17_B_R		
Category 1	0.026	189.000
Category 2	0.974	7211.000
P_17_D_R		
Category 1	0.011	85.000
Category 2	0.989	7315.000
P_20_G_R		
Category 1	0.485	2982.000
Category 2	0.515	3172.000
P_20_K_R		
Category 1	0.330	2047.000
Category 2	0.670	4162.000
P_20_L_R		
Category 1	0.701	4134.000
Category 2	0.299	1767.000
P_20_I_R		
Category 1	0.310	1953.000
Category 2	0.690	4345.000
P_20_E_R		
Category 1	0.428	2661.000
Category 2	0.572	3562.000
P_20_J_R		
Category 1	0.439	2588.000
Category 2	0.561	3306.000

WARNING: THE BIVARIATE TABLE OF P\_2\_C\_R AND P\_2\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_A\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_8\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
 WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_D\_R HAS AN EMPTY CELL.

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 103

Chi-Square Test of Model Fit

Value	11855.593*
Degrees of Freedom	488
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used

for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULMSV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.054	
90 Percent C.I.	0.053	0.055
Probability RMSEA <= .05	0.000	

CFI/TLI

CFI	0.951
TLI	0.947

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

WRMR (Weighted Root Mean Square Residual)

Value	5.089
-------	-------

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.834	0.020	42.495	0.000
P_18_F_R	0.851	0.020	43.612	0.000
P_18_G_R	0.788	0.019	42.341	0.000
P_18_I_R	1.000	0.000	999.000	999.000
P_18_J_R	0.891	0.020	44.623	0.000
PROP_M2 BY				
P_18_B_R	0.939	0.011	82.816	0.000
P_18_D_R	1.000	0.000	999.000	999.000
P_18_E_R	0.712	0.013	55.943	0.000
P_18_H_R	0.792	0.012	68.625	0.000
P_18_K_R	0.829	0.011	72.914	0.000
SATPROEV BY				
P_1_A_R	0.887	0.007	121.916	0.000
P_1_B_R	0.992	0.008	129.536	0.000
P_2_A_R	0.461	0.025	18.484	0.000
P_2_C_R	1.000	0.000	999.000	999.000
P_5_R_D	0.901	0.008	111.919	0.000
P_7_R	0.829	0.010	80.429	0.000
P_8_R	0.983	0.008	120.514	0.000
P_9_A_R	0.371	0.027	13.603	0.000
P_9_B_R	0.326	0.030	10.846	0.000
P_9_D_R	0.665	0.016	41.118	0.000
FLXEPSNI BY				
P_3_R_D	0.585	0.016	36.195	0.000
P_10_R	1.000	0.000	999.000	999.000
P_11_R_D	0.919	0.013	71.470	0.000
P_9_C_R	0.660	0.021	32.097	0.000
P_17_A_R	0.694	0.023	30.349	0.000
P_17_B_R	0.670	0.025	26.883	0.000
P_17_D_R	0.590	0.036	16.446	0.000



DIRREF	BY				
	P_20_G_R	0.828	0.008	106.721	0.000
	P_20_K_R	0.984	0.008	118.392	0.000
	P_20_L_R	0.623	0.011	54.412	0.000
	P_20_I_R	1.000	0.000	999.000	999.000
	P_20_J_R	0.921	0.008	121.908	0.000
	P_20_E_R	0.841	0.009	96.818	0.000
FLXEPSNI	ON				
	PROP_M1	0.451	0.023	19.575	0.000
	PROP_M2	0.121	0.019	6.235	0.000
SATPROEV	ON				
	PROP_M1	0.096	0.021	4.571	0.000
	PROP_M2	-0.194	0.017	-11.582	0.000
	FLXEPSNI	0.884	0.014	61.563	0.000
DIRREF	ON				
	SATPROEV	-0.675	0.011	-61.580	0.000
PROP_M2	WITH				
	PROP_M1	0.223	0.008	28.738	0.000
Thresholds					
	P_18_C_R\$1	-2.304	0.041	-55.824	0.000
	P_18_C_R\$2	-1.634	0.024	-68.895	0.000
	P_18_C_R\$3	-0.737	0.016	-47.046	0.000
	P_18_C_R\$4	0.519	0.015	34.874	0.000
	P_18_F_R\$1	-1.964	0.030	-64.701	0.000
	P_18_F_R\$2	-1.209	0.019	-64.701	0.000
	P_18_F_R\$3	-0.372	0.015	-25.567	0.000
	P_18_F_R\$4	0.868	0.016	53.179	0.000
	P_18_G_R\$1	-1.109	0.018	-62.198	0.000
	P_18_G_R\$2	-0.257	0.014	-17.960	0.000
	P_18_G_R\$3	0.580	0.015	38.507	0.000
	P_18_G_R\$4	1.597	0.023	69.011	0.000
	P_18_I_R\$1	-1.665	0.024	-68.822	0.000
	P_18_I_R\$2	-1.023	0.017	-59.448	0.000
	P_18_I_R\$3	-0.314	0.014	-21.819	0.000
	P_18_I_R\$4	0.674	0.015	43.777	0.000
	P_18_J_R\$1	-1.551	0.023	-68.911	0.000
	P_18_J_R\$2	-0.722	0.016	-46.232	0.000
	P_18_J_R\$3	0.127	0.014	8.926	0.000
	P_18_J_R\$4	1.176	0.018	63.935	0.000
	P_18_B_R\$1	-1.872	0.028	-66.598	0.000
	P_18_B_R\$2	-1.287	0.019	-66.506	0.000
	P_18_B_R\$3	-0.534	0.015	-35.832	0.000
	P_18_B_R\$4	0.412	0.015	28.248	0.000
	P_18_D_R\$1	-2.058	0.033	-62.740	0.000
	P_18_D_R\$2	-1.460	0.021	-68.601	0.000
	P_18_D_R\$3	-0.751	0.016	-47.788	0.000
	P_18_D_R\$4	0.229	0.014	15.981	0.000
	P_18_E_R\$1	-2.419	0.046	-52.130	0.000
	P_18_E_R\$2	-1.812	0.027	-67.417	0.000
	P_18_E_R\$3	-1.007	0.017	-58.852	0.000
	P_18_E_R\$4	0.061	0.014	4.316	0.000
	P_18_H_R\$1	-2.145	0.035	-60.491	0.000
	P_18_H_R\$2	-1.472	0.021	-68.664	0.000
	P_18_H_R\$3	-0.720	0.016	-46.172	0.000
	P_18_H_R\$4	0.295	0.014	20.469	0.000
	P_18_K_R\$1	-1.505	0.022	-68.858	0.000
	P_18_K_R\$2	-0.800	0.016	-50.160	0.000

P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000
Variances				
PROP_M1	0.496	0.013	37.709	0.000
PROP_M2	0.677	0.011	63.336	0.000
Residual Variances				
SATPROEV	0.268	0.013	21.030	0.000
FLXEPSNI	0.723	0.015	47.950	0.000
DIRREF	0.553	0.014	40.326	0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.587	0.011	53.979	0.000
P_18_F_R	0.600	0.011	56.239	0.000
P_18_G_R	0.555	0.011	50.944	0.000
P_18_I_R	0.705	0.009	75.418	0.000
P_18_J_R	0.628	0.010	62.203	0.000
PROP_M2 BY				
P_18_B_R	0.772	0.007	114.704	0.000
P_18_D_R	0.823	0.006	126.672	0.000
P_18_E_R	0.586	0.010	60.835	0.000
P_18_H_R	0.651	0.008	77.020	0.000
P_18_K_R	0.682	0.008	84.582	0.000
SATPROEV BY				
P_1_A_R	0.860	0.006	138.838	0.000
P_1_B_R	0.962	0.004	223.244	0.000
P_2_A_R	0.447	0.023	19.219	0.000
P_2_C_R	0.970	0.004	248.948	0.000
P_5_R_D	0.874	0.007	127.348	0.000
P_7_R	0.803	0.009	85.151	0.000
P_8_R	0.953	0.007	140.203	0.000

P_9_A_R	0.360	0.026	13.634	0.000
P_9_B_R	0.317	0.029	10.869	0.000
P_9_D_R	0.644	0.015	41.843	0.000
FLXEPSNI BY				
P_3_R_D	0.542	0.014	37.855	0.000
P_10_R	0.927	0.008	122.884	0.000
P_11_R_D	0.852	0.009	99.348	0.000
P_9_C_R	0.612	0.019	33.070	0.000
P_17_A_R	0.643	0.020	31.409	0.000
P_17_B_R	0.621	0.023	27.459	0.000
P_17_D_R	0.547	0.033	16.556	0.000
DIRREF BY				
P_20_G_R	0.820	0.007	114.720	0.000
P_20_K_R	0.975	0.005	192.444	0.000
P_20_L_R	0.617	0.011	55.066	0.000
P_20_I_R	0.991	0.005	201.165	0.000
P_20_J_R	0.912	0.006	158.569	0.000
P_20_E_R	0.833	0.008	108.521	0.000
FLXEPSNI ON				
PROP_M1	0.343	0.017	20.322	0.000
PROP_M2	0.108	0.017	6.259	0.000
SATPROEV ON				
PROP_M1	0.070	0.015	4.586	0.000
PROP_M2	-0.165	0.014	-11.617	0.000
FLXEPSNI	0.845	0.011	79.695	0.000
DIRREF ON				
SATPROEV	-0.661	0.010	-67.456	0.000
PROP_M2 WITH				
PROP_M1	0.385	0.012	31.899	0.000
Thresholds				
P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000
P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000
P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000
P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000

P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000
P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000
P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000

Variiances

PROP_M1	1.000	0.000	999.000	999.000
PROP_M2	1.000	0.000	999.000	999.000

Residual Variiances

SATPROEV	0.285	0.013	21.367	0.000
FLXEPSNI	0.842	0.011	77.648	0.000
DIRREF	0.563	0.013	43.540	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Residual Variance
P_18_C_R	0.345	0.013	26.990	0.000	0.655
P_18_F_R	0.360	0.013	28.119	0.000	0.640
P_18_G_R	0.308	0.012	25.472	0.000	0.692
P_18_I_R	0.496	0.013	37.709	0.000	0.504
P_18_J_R	0.394	0.013	31.102	0.000	0.606
P_18_B_R	0.597	0.010	57.352	0.000	0.403
P_18_D_R	0.677	0.011	63.336	0.000	0.323
P_18_E_R	0.343	0.011	30.417	0.000	0.657
P_18_H_R	0.424	0.011	38.510	0.000	0.576
P_18_K_R	0.465	0.011	42.291	0.000	0.535
P_1_A_R	0.740	0.011	69.419	0.000	0.260

P_1_B_R	0.926	0.008	111.622	0.000	0.074
P_2_A_R	0.200	0.021	9.610	0.000	0.800
P_2_C_R	0.940	0.008	124.474	0.000	0.060
P_5_R_D	0.764	0.012	63.674	0.000	0.236
P_7_R	0.645	0.015	42.576	0.000	0.355
P_8_R	0.908	0.013	70.102	0.000	0.092
P_9_A_R	0.130	0.019	6.817	0.000	0.870
P_9_B_R	0.100	0.018	5.435	0.000	0.900
P_9_D_R	0.415	0.020	20.922	0.000	0.585
P_3_R_D	0.294	0.016	18.927	0.000	0.706
P_10_R	0.859	0.014	61.442	0.000	0.141
P_11_R_D	0.726	0.015	49.674	0.000	0.274
P_9_C_R	0.375	0.023	16.535	0.000	0.625
P_17_A_R	0.413	0.026	15.705	0.000	0.587
P_17_B_R	0.385	0.028	13.729	0.000	0.615
P_17_D_R	0.299	0.036	8.278	0.000	0.701
P_20_G_R	0.673	0.012	57.360	0.000	0.327
P_20_K_R	0.951	0.010	96.222	0.000	0.049
P_20_L_R	0.381	0.014	27.533	0.000	0.619
P_20_I_R	0.982	0.010	100.583	0.000	0.018
P_20_E_R	0.694	0.013	54.261	0.000	0.306
P_20_J_R	0.832	0.010	79.285	0.000	0.168

Latent Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SATPROEV	0.715	0.013	53.688	0.000
FLXEPSNI	0.158	0.011	14.555	0.000
DIRREF	0.437	0.013	33.728	0.000

#### QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.491E-03  
(ratio of smallest to largest eigenvalue)

#### DIAGRAM INFORMATION

Use View Diagram under the Diagram menu in the Mplus Editor to view the diagram.

If running Mplus from the Mplus Diagrammer, the diagram opens automatically.

#### Diagram output

c:\users\repodrida\documents\nata\documents\doctorado en ciencias sociales\_uam\_x\2016\_capítulos\

Beginning Time: 13:14:58  
Ending Time: 13:15:08  
Elapsed Time: 00:00:10

MUTHEN & MUTHEN  
3463 Stoner Ave.  
Los Angeles, CA 90066

Tel: (310) 391-9971  
Fax: (310) 391-8971  
Web: www.StatModel.com  
Support: Support@StatModel.com

Copyright (c) 1998-2015 Muthen & Muthen

## ANEXO 5: Modelo de Ecuaciones Estructurales Saturado. Salida M-PLUS

Mplus VERSION 7.4  
MUTHEN & MUTHEN  
02/07/2017 12:48 PM

### INPUT INSTRUCTIONS

TITLE: Modelo de Ecuaciones Estructurales saturado para la relación  
entre la evaluación del desempeño individual de investigación de los  
investigadores SNI y la producción de conocimiento.

DATA: FILE IS enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat ;  
FORMAT IS FREE ;

VARIABLE: NAMES ARE id\_user edad form niv\_sni cand niv\_I niv\_II niv\_III area\_sni  
f\_m\_ct b\_q m\_cs h\_cc cs b\_ca ing gen\_sni mujer hombre tip\_inst cpi\_e\_n cpi\_cs\_h  
cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ot\_inst ori\_inst t\_d\_k use\_ori  
p\_1\_a\_r p\_1\_b\_r p\_1\_ab\_r p\_2\_r p\_2\_a\_r p\_2\_b\_r p\_2\_c\_r p\_2\_d\_r p\_3\_r p\_3\_r\_d p\_5\_r  
p\_5\_r\_d p\_6\_a\_r p\_6\_b\_r p\_6\_c\_r p\_6\_d\_r p\_6\_e\_r p\_7\_r p\_8\_r p\_9\_r p\_9\_a\_r p\_9\_b\_r  
p\_9\_c\_r p\_9\_d\_r p\_9\_e\_r p\_10\_r p\_11\_r p\_11\_r\_d p\_13\_r p\_14\_a\_r p\_14\_b\_r p\_14\_c\_r  
p\_14\_d\_r p\_14\_e\_r p\_14\_f\_r p\_14\_g\_r p\_14\_h\_r p\_15\_r p\_15\_a\_r p\_15\_b\_r p\_15\_c\_r  
p\_16\_a\_r p\_16\_b\_r p\_16\_c\_r p\_16\_d\_r p\_16\_e\_r p\_16\_f\_r p\_16\_g\_r p\_17\_a\_r p\_17\_b\_r  
p\_17\_c\_r p\_17\_d\_r p\_17\_e\_r p\_17\_f\_r p\_18\_a\_r p\_18\_b\_r p\_18\_c\_r p\_18\_d\_r p\_18\_e\_r  
p\_18\_f\_r p\_18\_g\_r p\_18\_h\_r p\_18\_i\_r p\_18\_j\_r p\_18\_k\_r p\_19\_r p\_20\_a\_r p\_20\_b\_r  
p\_20\_c\_r p\_20\_d\_r p\_20\_e\_r p\_20\_f\_r p\_20\_g\_r p\_20\_h\_r p\_20\_i\_r p\_20\_j\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_m\_r p\_39\_r ;

USEVARIABLES ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r  
p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r  
!p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r  
p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ; !edad !cand !niv\_I !niv\_II !niv\_III !f\_m\_ct  
!b\_q !m\_cs !h\_cc !cs !b\_ca !ing !mujer !hombre !cpi\_e\_n !cpi\_cs\_h !cpi\_dt\_s !cpi\_sec  
!gobierno !inst\_n\_s !inst\_ut !univ ; !ot\_inst !t\_d\_k !use\_ori ;

MISSING ARE p\_1\_a\_r (2) p\_1\_b\_r (2) p\_1\_ab\_r (2) p\_2\_r (5) p\_2\_a\_r (2) p\_2\_b\_r (2)  
p\_2\_c\_r (2) p\_2\_d\_r (2) p\_3\_r\_d (4 5) p\_5\_r\_d (5) p\_6\_a\_r (2) p\_6\_b\_r (2) p\_6\_c\_r  
(2) p\_6\_d\_r (2) p\_6\_e\_r (2) p\_7\_r (2) p\_8\_r (2) p\_9\_r (6) p\_9\_a\_r (2) p\_9\_b\_r (2)  
p\_9\_c\_r (2) p\_9\_d\_r (2) p\_9\_e\_r (2) p\_10\_r (2) p\_11\_r\_d (5) p\_15\_r (4) p\_15\_a\_r (2)  
p\_15\_b\_r (2) p\_15\_c\_r (2) p\_16\_a\_r (2) p\_16\_b\_r (2) p\_16\_c\_r (2) p\_16\_d\_r (2)  
p\_16\_e\_r (2) p\_16\_f\_r (2) p\_16\_g\_r (2) p\_17\_a\_r (2) p\_17\_b\_r (2) p\_17\_c\_r (2)  
p\_17\_d\_r (2) p\_17\_e\_r (2) p\_17\_f\_r (2) p\_19\_r (2) p\_20\_a\_r (2) p\_20\_b\_r (2) p\_20\_c\_r  
(2) p\_20\_d\_r (2) p\_20\_e\_r (2) p\_20\_f\_r (2) p\_20\_g\_r (2) p\_20\_h\_r (2) p\_20\_i\_r (2)  
p\_20\_j\_r (2) p\_20\_k\_r (2) p\_20\_l\_r (2) p\_20\_m\_r (2) . ;

CATEGORICAL ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r  
p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r  
!p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r  
p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r  
p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ;

AUXILIARY ARE id\_user ;

MODEL: PROP\_M1 BY p\_18\_c\_r\* p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r@1 p\_18\_j\_r ;  
PROP\_M2 BY p\_18\_b\_r\* p\_18\_d\_r@1 p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r ;  
SATPROEV BY p\_1\_a\_r\* p\_1\_b\_r p\_2\_a\_r p\_2\_c\_r@1 p\_5\_r\_d p\_7\_r p\_8\_r p\_9\_a\_r  
p\_9\_b\_r p\_9\_d\_r ;

FLXEPSNI BY p\_3\_r\_d\* p\_10\_r@1 p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r;  
DIRREF BY p\_20\_g\_r\* p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r@1 p\_20\_j\_r p\_20\_e\_r;

FLXEPSNI ON PROP\_M1 PROP\_M2 ;  
SATPROEV ON PROP\_M1 PROP\_M2 FLXEPSNI ;  
DIRREF ON PROP\_M1 PROP\_M2 SATPROEV FLXEPSNI ;

OUTPUT: stdyx;

\*\*\* WARNING

Data set contains cases with missing on all variables. These cases were not included in the analysis. Number of cases with missing on all variables: 2

1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

Modelo de Ecuaciones Estructurales saturado para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

#### SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	8053
Number of dependent variables	33
Number of independent variables	0
Number of continuous latent variables	5

#### Observed dependent variables

Binary and ordered categorical (ordinal)

P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R	P_18_B_R
P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R	P_1_A_R	P_1_B_R
P_2_A_R	P_2_C_R	P_5_R_D	P_7_R	P_8_R	P_9_A_R
P_9_B_R	P_9_D_R	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R
P_17_A_R	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_20_I_R	P_20_E_R	P_20_J_R			

#### Observed auxiliary variables

ID\_USER

#### Continuous latent variables

PROP\_M1 PROP\_M2 SATPROEV FLXEPSNI DIRREF

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.100D-03
Parameterization	DELTA

#### Input data file(s)

enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat

Input data format FREE

#### SUMMARY OF DATA

Number of missing data patterns 1292

#### COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

	Covariance Coverage				
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	0.972				
P_18_F_R	0.965	0.968			
P_18_G_R	0.969	0.966	0.972		
P_18_I_R	0.969	0.966	0.971	0.973	
P_18_J_R	0.965	0.963	0.967	0.969	0.970
P_18_B_R	0.970	0.967	0.971	0.971	0.968
P_18_D_R	0.968	0.965	0.970	0.970	0.966
P_18_E_R	0.968	0.966	0.970	0.971	0.967
P_18_H_R	0.967	0.965	0.969	0.970	0.966
P_18_K_R	0.968	0.965	0.969	0.971	0.967
P_1_A_R	0.834	0.831	0.834	0.835	0.832
P_1_B_R	0.849	0.846	0.850	0.851	0.847
P_2_A_R	0.948	0.944	0.949	0.950	0.946
P_2_C_R	0.948	0.944	0.949	0.950	0.946
P_5_R_D	0.849	0.846	0.850	0.851	0.847
P_7_R	0.670	0.667	0.670	0.671	0.668
P_8_R	0.614	0.612	0.614	0.615	0.613
P_9_A_R	0.613	0.612	0.614	0.615	0.613
P_9_B_R	0.613	0.612	0.614	0.615	0.613
P_9_D_R	0.613	0.612	0.614	0.615	0.613
P_3_R_D	0.904	0.901	0.905	0.906	0.902
P_10_R	0.834	0.830	0.834	0.834	0.831
P_11_R_D	0.864	0.860	0.865	0.865	0.862
P_9_C_R	0.613	0.612	0.614	0.615	0.613
P_17_A_R	0.910	0.907	0.911	0.912	0.908
P_17_B_R	0.910	0.907	0.911	0.912	0.908
P_17_D_R	0.910	0.907	0.911	0.912	0.908
P_20_G_R	0.760	0.757	0.761	0.762	0.760
P_20_K_R	0.767	0.765	0.768	0.769	0.766
P_20_L_R	0.729	0.726	0.730	0.730	0.728
P_20_I_R	0.778	0.775	0.779	0.780	0.777
P_20_E_R	0.768	0.766	0.770	0.770	0.768
P_20_J_R	0.728	0.725	0.729	0.730	0.728

	Covariance Coverage				
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_B_R	0.975				
P_18_D_R	0.970	0.972			
P_18_E_R	0.971	0.969	0.973		
P_18_H_R	0.970	0.968	0.969	0.971	
P_18_K_R	0.970	0.969	0.969	0.969	0.972
P_1_A_R	0.836	0.834	0.834	0.833	0.834
P_1_B_R	0.851	0.849	0.850	0.849	0.849
P_2_A_R	0.951	0.948	0.949	0.948	0.948
P_2_C_R	0.951	0.948	0.949	0.948	0.948
P_5_R_D	0.852	0.849	0.850	0.849	0.849
P_7_R	0.672	0.670	0.670	0.669	0.670
P_8_R	0.615	0.614	0.615	0.614	0.614
P_9_A_R	0.615	0.614	0.615	0.614	0.614
P_9_B_R	0.615	0.614	0.615	0.614	0.614
P_9_D_R	0.615	0.614	0.615	0.614	0.614
P_3_R_D	0.907	0.904	0.905	0.904	0.905
P_10_R	0.835	0.833	0.834	0.833	0.833
P_11_R_D	0.866	0.864	0.865	0.864	0.864
P_9_C_R	0.615	0.614	0.615	0.614	0.614
P_17_A_R	0.913	0.910	0.911	0.910	0.910
P_17_B_R	0.913	0.910	0.911	0.910	0.910



P_17_D_R	0.913	0.910	0.911	0.910	0.910
P_20_G_R	0.762	0.761	0.761	0.760	0.761
P_20_K_R	0.769	0.768	0.768	0.768	0.768
P_20_L_R	0.730	0.730	0.730	0.729	0.730
P_20_I_R	0.780	0.779	0.779	0.778	0.779
P_20_E_R	0.770	0.769	0.770	0.769	0.770
P_20_J_R	0.730	0.729	0.729	0.728	0.729

Covariance Coverage

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
P_1_A_R	0.856				
P_1_B_R	0.752	0.873			
P_2_A_R	0.855	0.871	0.975		
P_2_C_R	0.855	0.871	0.975	0.975	
P_5_R_D	0.750	0.775	0.853	0.853	0.871
P_7_R	0.596	0.615	0.675	0.675	0.657
P_8_R	0.544	0.569	0.618	0.618	0.606
P_9_A_R	0.543	0.569	0.618	0.618	0.606
P_9_B_R	0.543	0.569	0.618	0.618	0.606
P_9_D_R	0.543	0.569	0.618	0.618	0.606
P_3_R_D	0.796	0.840	0.911	0.911	0.820
P_10_R	0.733	0.751	0.832	0.832	0.756
P_11_R_D	0.757	0.774	0.860	0.860	0.781
P_9_C_R	0.543	0.569	0.618	0.618	0.606
P_17_A_R	0.790	0.806	0.898	0.898	0.804
P_17_B_R	0.790	0.806	0.898	0.898	0.804
P_17_D_R	0.790	0.806	0.898	0.898	0.804
P_20_G_R	0.666	0.683	0.751	0.751	0.685
P_20_K_R	0.669	0.689	0.758	0.758	0.691
P_20_L_R	0.639	0.657	0.722	0.722	0.658
P_20_I_R	0.678	0.699	0.769	0.769	0.702
P_20_E_R	0.671	0.691	0.760	0.760	0.693
P_20_J_R	0.636	0.658	0.719	0.719	0.668

Covariance Coverage

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
P_7_R	0.685				
P_8_R	0.552	0.628			
P_9_A_R	0.552	0.628	0.628		
P_9_B_R	0.552	0.628	0.628	0.628	
P_9_D_R	0.552	0.628	0.628	0.628	0.628
P_3_R_D	0.647	0.598	0.597	0.597	0.597
P_10_R	0.605	0.560	0.560	0.560	0.560
P_11_R_D	0.621	0.570	0.570	0.570	0.570
P_9_C_R	0.552	0.628	0.628	0.628	0.628
P_17_A_R	0.634	0.581	0.581	0.581	0.581
P_17_B_R	0.634	0.581	0.581	0.581	0.581
P_17_D_R	0.634	0.581	0.581	0.581	0.581
P_20_G_R	0.556	0.520	0.520	0.520	0.520
P_20_K_R	0.560	0.527	0.526	0.526	0.526
P_20_L_R	0.539	0.505	0.505	0.505	0.505
P_20_I_R	0.569	0.534	0.534	0.534	0.534
P_20_E_R	0.565	0.529	0.528	0.528	0.528
P_20_J_R	0.548	0.518	0.518	0.518	0.518

Covariance Coverage

	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
P_3_R_D	0.930				
P_10_R	0.793	0.848			
P_11_R_D	0.822	0.830	0.879		

P_9_C_R	0.597	0.560	0.570	0.628	
P_17_A_R	0.856	0.801	0.829	0.581	0.919
P_17_B_R	0.856	0.801	0.829	0.581	0.919
P_17_D_R	0.856	0.801	0.829	0.581	0.919
P_20_G_R	0.722	0.683	0.699	0.520	0.724
P_20_K_R	0.728	0.690	0.707	0.526	0.729
P_20_L_R	0.693	0.655	0.672	0.505	0.693
P_20_I_R	0.738	0.697	0.715	0.534	0.739
P_20_E_R	0.729	0.691	0.707	0.528	0.730
P_20_J_R	0.693	0.657	0.671	0.518	0.693

Covariance Coverage

	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_B_R	0.919				
P_17_D_R	0.919	0.919			
P_20_G_R	0.724	0.724	0.764		
P_20_K_R	0.729	0.729	0.737	0.771	
P_20_L_R	0.693	0.693	0.704	0.709	0.733
P_20_I_R	0.739	0.739	0.741	0.750	0.714
P_20_E_R	0.730	0.730	0.738	0.742	0.708
P_20_J_R	0.693	0.693	0.701	0.710	0.677

Covariance Coverage

	P_20_I_R	P_20_E_R	P_20_J_R
P_20_I_R	0.782		
P_20_E_R	0.751	0.773	
P_20_J_R	0.717	0.706	0.732

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

P_18_C_R		
Category 1	0.011	83.000
Category 2	0.041	317.000
Category 3	0.180	1405.000
Category 4	0.467	3658.000
Category 5	0.302	2362.000
P_18_F_R		
Category 1	0.025	193.000
Category 2	0.088	690.000
Category 3	0.242	1884.000
Category 4	0.452	3527.000
Category 5	0.193	1503.000
P_18_G_R		
Category 1	0.134	1047.000
Category 2	0.265	2073.000
Category 3	0.321	2511.000
Category 4	0.226	1768.000
Category 5	0.055	432.000
P_18_I_R		
Category 1	0.048	376.000
Category 2	0.105	825.000
Category 3	0.223	1751.000
Category 4	0.373	2925.000
Category 5	0.250	1962.000
P_18_J_R		
Category 1	0.060	472.000
Category 2	0.175	1365.000
Category 3	0.315	2462.000
Category 4	0.330	2574.000
Category 5	0.120	936.000
P_18_B_R		

Category 1	0.031	240.000
Category 2	0.069	538.000
Category 3	0.198	1550.000
Category 4	0.363	2851.000
Category 5	0.340	2669.000
P_18_D_R		
Category 1	0.020	155.000
Category 2	0.052	410.000
Category 3	0.154	1206.000
Category 4	0.364	2850.000
Category 5	0.410	3206.000
P_18_E_R		
Category 1	0.008	61.000
Category 2	0.027	213.000
Category 3	0.122	956.000
Category 4	0.367	2877.000
Category 5	0.476	3725.000
P_18_H_R		
Category 1	0.016	125.000
Category 2	0.055	427.000
Category 3	0.165	1293.000
Category 4	0.380	2973.000
Category 5	0.384	3005.000
P_18_K_R		
Category 1	0.066	518.000
Category 2	0.146	1141.000
Category 3	0.285	2228.000
Category 4	0.334	2613.000
Category 5	0.170	1327.000
P_1_A_R		
Category 1	0.174	1202.000
Category 2	0.826	5692.000
P_1_B_R		
Category 1	0.237	1669.000
Category 2	0.763	5361.000
P_2_A_R		
Category 1	0.027	213.000
Category 2	0.973	7641.000
P_2_C_R		
Category 1	0.196	1542.000
Category 2	0.804	6312.000
P_5_R_D		
Category 1	0.356	2500.000
Category 2	0.644	4514.000
P_7_R		
Category 1	0.219	1208.000
Category 2	0.781	4308.000
P_8_R		
Category 1	0.381	1926.000
Category 2	0.619	3131.000
P_9_A_R		
Category 1	0.044	224.000
Category 2	0.956	4830.000
P_9_B_R		
Category 1	0.045	225.000
Category 2	0.955	4829.000
P_9_D_R		
Category 1	0.137	693.000
Category 2	0.863	4361.000
P_3_R_D		
Category 1	0.506	3786.000
Category 2	0.494	3703.000
P_10_R		

Category 1	0.451	3083.000
Category 2	0.549	3749.000
P_11_R_D		
Category 1	0.530	3753.000
Category 2	0.470	3326.000
P_9_C_R		
Category 1	0.142	719.000
Category 2	0.858	4335.000
P_17_A_R		
Category 1	0.033	242.000
Category 2	0.967	7158.000
P_17_B_R		
Category 1	0.026	189.000
Category 2	0.974	7211.000
P_17_D_R		
Category 1	0.011	85.000
Category 2	0.989	7315.000
P_20_G_R		
Category 1	0.485	2982.000
Category 2	0.515	3172.000
P_20_K_R		
Category 1	0.330	2047.000
Category 2	0.670	4162.000
P_20_L_R		
Category 1	0.701	4134.000
Category 2	0.299	1767.000
P_20_I_R		
Category 1	0.310	1953.000
Category 2	0.690	4345.000
P_20_E_R		
Category 1	0.428	2661.000
Category 2	0.572	3562.000
P_20_J_R		
Category 1	0.439	2588.000
Category 2	0.561	3306.000

WARNING: THE BIVARIATE TABLE OF P\_2\_C\_R AND P\_2\_A\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_A\_R AND P\_8\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_8\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_B\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_8\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_D\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_8\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_A\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_B\_R HAS AN EMPTY CELL.  
WARNING: THE BIVARIATE TABLE OF P\_9\_C\_R AND P\_9\_D\_R HAS AN EMPTY CELL.

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 106

Chi-Square Test of Model Fit

Value	10890.791*
Degrees of Freedom	485
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV,

and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.052	
90 Percent C.I.	0.051	0.052
Probability RMSEA <= .05	0.001	

CFI/TLI

CFI	0.955
TLI	0.951

Chi-Square Test of Model Fit for the Baseline Model

Value	230934.785
Degrees of Freedom	528
P-Value	0.0000

WRMR (Weighted Root Mean Square Residual)

Value	4.775
-------	-------

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.823	0.019	42.219	0.000
P_18_F_R	0.844	0.019	43.401	0.000
P_18_G_R	0.783	0.019	42.331	0.000
P_18_I_R	1.000	0.000	999.000	999.000
P_18_J_R	0.882	0.020	44.695	0.000
PROP_M2 BY				
P_18_B_R	0.937	0.011	83.477	0.000
P_18_D_R	1.000	0.000	999.000	999.000
P_18_E_R	0.715	0.013	56.636	0.000
P_18_H_R	0.795	0.011	69.377	0.000
P_18_K_R	0.827	0.011	73.047	0.000
SATPROEV BY				
P_1_A_R	0.897	0.007	123.275	0.000
P_1_B_R	0.994	0.008	127.770	0.000
P_2_A_R	0.481	0.025	19.038	0.000
P_2_C_R	1.000	0.000	999.000	999.000
P_5_R_D	0.914	0.008	111.815	0.000
P_7_R	0.844	0.010	82.058	0.000
P_8_R	0.996	0.008	121.154	0.000
P_9_A_R	0.386	0.028	13.737	0.000
P_9_B_R	0.339	0.031	11.018	0.000
P_9_D_R	0.686	0.016	41.716	0.000
FLXEPSNI BY				
P_3_R_D	0.588	0.016	36.599	0.000
P_10_R	1.000	0.000	999.000	999.000
P_11_R_D	0.915	0.013	70.812	0.000
P_9_C_R	0.656	0.020	32.110	0.000
P_17_A_R	0.695	0.023	30.519	0.000
P_17_B_R	0.671	0.025	27.199	0.000
P_17_D_R	0.589	0.036	16.309	0.000
DIRREF BY				

P_20_G_R	0.831	0.008	107.719	0.000
P_20_K_R	0.986	0.008	120.105	0.000
P_20_L_R	0.637	0.011	56.008	0.000
P_20_I_R	1.000	0.000	999.000	999.000
P_20_J_R	0.921	0.008	122.170	0.000
P_20_E_R	0.840	0.009	96.921	0.000
FLXEPSNI ON				
PROP_M1	0.449	0.023	19.592	0.000
PROP_M2	0.121	0.019	6.230	0.000
SATPROEV ON				
PROP_M1	0.012	0.021	0.546	0.585
PROP_M2	-0.085	0.017	-4.941	0.000
FLXEPSNI	0.804	0.015	55.262	0.000
DIRREF ON				
PROP_M1	-0.206	0.025	-8.365	0.000
PROP_M2	0.308	0.019	15.866	0.000
SATPROEV	-0.140	0.024	-5.718	0.000
FLXEPSNI	-0.614	0.030	-20.641	0.000
PROP_M2 WITH				
PROP_M1	0.225	0.008	28.700	0.000
Thresholds				
P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000
P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000
P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000
P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000
P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000
P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000

P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000
Variances				
PROP_M1	0.503	0.013	37.966	0.000
PROP_M2	0.677	0.011	63.785	0.000
Residual Variances				
SATPROEV	0.404	0.013	30.278	0.000
FLXEPSNI	0.725	0.015	47.719	0.000
DIRREF	0.464	0.015	30.065	0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.584	0.011	53.301	0.000
P_18_F_R	0.599	0.011	55.848	0.000
P_18_G_R	0.556	0.011	51.084	0.000
P_18_I_R	0.710	0.009	75.932	0.000
P_18_J_R	0.626	0.010	62.047	0.000
PROP_M2 BY				
P_18_B_R	0.771	0.007	114.790	0.000
P_18_D_R	0.823	0.006	127.569	0.000
P_18_E_R	0.589	0.010	61.525	0.000
P_18_H_R	0.654	0.008	77.680	0.000
P_18_K_R	0.680	0.008	84.493	0.000
SATPROEV BY				
P_1_A_R	0.872	0.006	143.130	0.000
P_1_B_R	0.967	0.004	222.082	0.000
P_2_A_R	0.468	0.024	19.822	0.000
P_2_C_R	0.972	0.004	246.812	0.000
P_5_R_D	0.889	0.007	127.724	0.000
P_7_R	0.821	0.009	87.335	0.000

P_8_R	0.968	0.007	141.639	0.000
P_9_A_R	0.376	0.027	13.769	0.000
P_9_B_R	0.329	0.030	11.043	0.000
P_9_D_R	0.667	0.016	42.503	0.000
FLXEPSNI BY				
P_3_R_D	0.546	0.014	38.337	0.000
P_10_R	0.928	0.008	121.983	0.000
P_11_R_D	0.849	0.009	98.691	0.000
P_9_C_R	0.609	0.018	33.089	0.000
P_17_A_R	0.645	0.020	31.618	0.000
P_17_B_R	0.623	0.022	27.784	0.000
P_17_D_R	0.547	0.033	16.414	0.000
DIRREF BY				
P_20_G_R	0.822	0.007	115.740	0.000
P_20_K_R	0.976	0.005	195.047	0.000
P_20_L_R	0.630	0.011	56.745	0.000
P_20_I_R	0.989	0.005	202.846	0.000
P_20_J_R	0.911	0.006	157.781	0.000
P_20_E_R	0.831	0.008	108.178	0.000
FLXEPSNI ON				
PROP_M1	0.343	0.017	20.358	0.000
PROP_M2	0.108	0.017	6.253	0.000
SATPROEV ON				
PROP_M1	0.008	0.016	0.546	0.585
PROP_M2	-0.072	0.015	-4.945	0.000
FLXEPSNI	0.768	0.011	67.534	0.000
DIRREF ON				
PROP_M1	-0.148	0.017	-8.445	0.000
PROP_M2	0.256	0.016	15.967	0.000
SATPROEV	-0.138	0.024	-5.719	0.000
FLXEPSNI	-0.576	0.027	-21.444	0.000
PROP_M2 WITH				
PROP_M1	0.385	0.012	31.809	0.000
Thresholds				
P_18_C_R\$1	-2.304	0.041	-55.824	0.000
P_18_C_R\$2	-1.634	0.024	-68.895	0.000
P_18_C_R\$3	-0.737	0.016	-47.046	0.000
P_18_C_R\$4	0.519	0.015	34.874	0.000
P_18_F_R\$1	-1.964	0.030	-64.701	0.000
P_18_F_R\$2	-1.209	0.019	-64.701	0.000
P_18_F_R\$3	-0.372	0.015	-25.567	0.000
P_18_F_R\$4	0.868	0.016	53.179	0.000
P_18_G_R\$1	-1.109	0.018	-62.198	0.000
P_18_G_R\$2	-0.257	0.014	-17.960	0.000
P_18_G_R\$3	0.580	0.015	38.507	0.000
P_18_G_R\$4	1.597	0.023	69.011	0.000
P_18_I_R\$1	-1.665	0.024	-68.822	0.000
P_18_I_R\$2	-1.023	0.017	-59.448	0.000
P_18_I_R\$3	-0.314	0.014	-21.819	0.000
P_18_I_R\$4	0.674	0.015	43.777	0.000
P_18_J_R\$1	-1.551	0.023	-68.911	0.000
P_18_J_R\$2	-0.722	0.016	-46.232	0.000
P_18_J_R\$3	0.127	0.014	8.926	0.000
P_18_J_R\$4	1.176	0.018	63.935	0.000
P_18_B_R\$1	-1.872	0.028	-66.598	0.000
P_18_B_R\$2	-1.287	0.019	-66.506	0.000



P_18_B_R\$3	-0.534	0.015	-35.832	0.000
P_18_B_R\$4	0.412	0.015	28.248	0.000
P_18_D_R\$1	-2.058	0.033	-62.740	0.000
P_18_D_R\$2	-1.460	0.021	-68.601	0.000
P_18_D_R\$3	-0.751	0.016	-47.788	0.000
P_18_D_R\$4	0.229	0.014	15.981	0.000
P_18_E_R\$1	-2.419	0.046	-52.130	0.000
P_18_E_R\$2	-1.812	0.027	-67.417	0.000
P_18_E_R\$3	-1.007	0.017	-58.852	0.000
P_18_E_R\$4	0.061	0.014	4.316	0.000
P_18_H_R\$1	-2.145	0.035	-60.491	0.000
P_18_H_R\$2	-1.472	0.021	-68.664	0.000
P_18_H_R\$3	-0.720	0.016	-46.172	0.000
P_18_H_R\$4	0.295	0.014	20.469	0.000
P_18_K_R\$1	-1.505	0.022	-68.858	0.000
P_18_K_R\$2	-0.800	0.016	-50.160	0.000
P_18_K_R\$3	-0.008	0.014	-0.599	0.549
P_18_K_R\$4	0.956	0.017	56.939	0.000
P_1_A_R\$1	-0.937	0.018	-52.739	0.000
P_1_B_R\$1	-0.715	0.016	-43.520	0.000
P_2_A_R\$1	-1.925	0.029	-65.705	0.000
P_2_C_R\$1	-0.855	0.016	-52.798	0.000
P_5_R_D\$1	-0.368	0.015	-23.992	0.000
P_7_R\$1	-0.776	0.019	-41.132	0.000
P_8_R\$1	-0.303	0.018	-16.919	0.000
P_9_A_R\$1	-1.703	0.031	-55.068	0.000
P_9_B_R\$1	-1.700	0.031	-55.081	0.000
P_9_D_R\$1	-1.093	0.022	-49.589	0.000
P_3_R_D\$1	0.014	0.014	0.959	0.338
P_10_R\$1	-0.122	0.015	-8.056	0.000
P_11_R_D\$1	0.076	0.015	5.075	0.000
P_9_C_R\$1	-1.070	0.022	-49.008	0.000
P_17_A_R\$1	-1.842	0.028	-65.118	0.000
P_17_B_R\$1	-1.951	0.031	-63.294	0.000
P_17_D_R\$1	-2.274	0.041	-55.200	0.000
P_20_G_R\$1	-0.039	0.016	-2.422	0.015
P_20_K_R\$1	-0.441	0.016	-26.747	0.000
P_20_L_R\$1	0.526	0.017	30.648	0.000
P_20_I_R\$1	-0.496	0.017	-30.002	0.000
P_20_E_R\$1	-0.182	0.016	-11.416	0.000
P_20_J_R\$1	-0.153	0.016	-9.349	0.000
Variances				
PROP_M1	1.000	0.000	999.000	999.000
PROP_M2	1.000	0.000	999.000	999.000
Residual Variances				
SATPROEV	0.427	0.014	31.320	0.000
FLXEPSNI	0.842	0.011	77.689	0.000
DIRREF	0.474	0.015	31.261	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Residual Variance
P_18_C_R	0.341	0.013	26.651	0.000	0.659
P_18_F_R	0.358	0.013	27.924	0.000	0.642
P_18_G_R	0.309	0.012	25.542	0.000	0.691
P_18_I_R	0.503	0.013	37.966	0.000	0.497
P_18_J_R	0.391	0.013	31.023	0.000	0.609
P_18_B_R	0.594	0.010	57.395	0.000	0.406
P_18_D_R	0.677	0.011	63.785	0.000	0.323

P_18_E_R	0.346	0.011	30.763	0.000	0.654
P_18_H_R	0.427	0.011	38.840	0.000	0.573
P_18_K_R	0.463	0.011	42.247	0.000	0.537
P_1_A_R	0.761	0.011	71.565	0.000	0.239
P_1_B_R	0.935	0.008	111.041	0.000	0.065
P_2_A_R	0.219	0.022	9.911	0.000	0.781
P_2_C_R	0.946	0.008	123.406	0.000	0.054
P_5_R_D	0.791	0.012	63.862	0.000	0.209
P_7_R	0.673	0.015	43.667	0.000	0.327
P_8_R	0.937	0.013	70.819	0.000	0.063
P_9_A_R	0.141	0.020	6.884	0.000	0.859
P_9_B_R	0.109	0.020	5.521	0.000	0.891
P_9_D_R	0.446	0.021	21.252	0.000	0.554
P_3_R_D	0.298	0.016	19.169	0.000	0.702
P_10_R	0.861	0.014	60.991	0.000	0.139
P_11_R_D	0.721	0.015	49.345	0.000	0.279
P_9_C_R	0.371	0.022	16.544	0.000	0.629
P_17_A_R	0.417	0.026	15.809	0.000	0.583
P_17_B_R	0.388	0.028	13.892	0.000	0.612
P_17_D_R	0.299	0.036	8.207	0.000	0.701
P_20_G_R	0.675	0.012	57.870	0.000	0.325
P_20_K_R	0.952	0.010	97.523	0.000	0.048
P_20_L_R	0.396	0.014	28.372	0.000	0.604
P_20_I_R	0.978	0.010	101.423	0.000	0.022
P_20_E_R	0.690	0.013	54.089	0.000	0.310
P_20_J_R	0.830	0.011	78.891	0.000	0.170

Latent Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SATPROEV	0.573	0.014	42.003	0.000
FLXEPSNI	0.158	0.011	14.561	0.000
DIRREF	0.526	0.015	34.700	0.000

#### QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.607E-03  
 (ratio of smallest to largest eigenvalue)

#### DIAGRAM INFORMATION

Use View Diagram under the Diagram menu in the Mplus Editor to view the diagram.

If running Mplus from the Mplus Diagrammer, the diagram opens automatically.

Diagram output

c:\users\repodrida\documents\nata\documents\doctorado en ciencias  
 sociales\_uam\_x\2016\_capítulos\

Beginning Time: 12:48:13  
 Ending Time: 12:48:23  
 Elapsed Time: 00:00:10

MUTHEN & MUTHEN  
 3463 Stoner Ave.  
 Los Angeles, CA 90066

Tel: (310) 391-9971  
 Fax: (310) 391-8971  
 Web: www.StatModel.com  
 Support: Support@StatModel.com

## ANEXO 6: Modelo de Ecuaciones Estructurales Condicional. Salida M-PLUS

Mplus VERSION 7.4  
MUTHEN & MUTHEN  
03/01/2017 2:14 PM

### INPUT INSTRUCTIONS

TITLE: Modelo de Ecuaciones Estructurales condicional para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

DATA: FILE IS enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat ;  
FORMAT IS FREE ;

VARIABLE: NAMES ARE id\_user edad form niv\_sni cand niv\_I niv\_II niv\_III area\_sni f\_m\_ct b\_q m\_cs h\_cc cs b\_ca ing gen\_sni mujer hombre tip\_inst cpi\_e\_n cpi\_cs\_h cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ot\_inst ori\_inst t\_d\_k use\_ori p\_1\_a\_r p\_1\_b\_r p\_1\_ab\_r p\_2\_r p\_2\_a\_r p\_2\_b\_r p\_2\_c\_r p\_2\_d\_r p\_3\_r p\_3\_r\_d p\_5\_r p\_5\_r\_d p\_6\_a\_r p\_6\_b\_r p\_6\_c\_r p\_6\_d\_r p\_6\_e\_r p\_7\_r p\_8\_r p\_9\_r p\_9\_a\_r p\_9\_b\_r p\_9\_c\_r p\_9\_d\_r p\_9\_e\_r p\_10\_r p\_11\_r p\_11\_r\_d p\_13\_r p\_14\_a\_r p\_14\_b\_r p\_14\_c\_r p\_14\_d\_r p\_14\_e\_r p\_14\_f\_r p\_14\_g\_r p\_14\_h\_r p\_15\_r p\_15\_a\_r p\_15\_b\_r p\_15\_c\_r p\_16\_a\_r p\_16\_b\_r p\_16\_c\_r p\_16\_d\_r p\_16\_e\_r p\_16\_f\_r p\_16\_g\_r p\_17\_a\_r p\_17\_b\_r p\_17\_c\_r p\_17\_d\_r p\_17\_e\_r p\_17\_f\_r p\_18\_a\_r p\_18\_b\_r p\_18\_c\_r p\_18\_d\_r p\_18\_e\_r p\_18\_f\_r p\_18\_g\_r p\_18\_h\_r p\_18\_i\_r p\_18\_j\_r p\_18\_k\_r p\_19\_r p\_20\_a\_r p\_20\_b\_r p\_20\_c\_r p\_20\_d\_r p\_20\_e\_r p\_20\_f\_r p\_20\_g\_r p\_20\_h\_r p\_20\_i\_r p\_20\_j\_r p\_20\_k\_r p\_20\_l\_r p\_20\_m\_r p\_39\_r ;

USEVARIABLES ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r edad cand niv\_I niv\_II !niv\_III f\_m\_ct b\_q m\_cs !h\_cc cs b\_ca ing mujer !hombre cpi\_e\_n cpi\_cs\_h cpi\_dt\_s cpi\_sec gobierno inst\_n\_s inst\_ut univ ; !ot\_inst !t\_d\_k !use\_ori ;

MISSING ARE p\_1\_a\_r (2) p\_1\_b\_r (2) p\_1\_ab\_r (2) p\_2\_r (5) p\_2\_a\_r (2) p\_2\_b\_r (2) p\_2\_c\_r (2) p\_2\_d\_r (2) p\_3\_r\_d (4 5) p\_5\_r\_d (5) p\_6\_a\_r (2) p\_6\_b\_r (2) p\_6\_c\_r (2) p\_6\_d\_r (2) p\_6\_e\_r (2) p\_7\_r (2) p\_8\_r (2) p\_9\_r (6) p\_9\_a\_r (2) p\_9\_b\_r (2) p\_9\_c\_r (2) p\_9\_d\_r (2) p\_9\_e\_r (2) p\_10\_r (2) p\_11\_r\_d (5) p\_15\_r (4) p\_15\_a\_r (2) p\_15\_b\_r (2) p\_15\_c\_r (2) p\_16\_a\_r (2) p\_16\_b\_r (2) p\_16\_c\_r (2) p\_16\_d\_r (2) p\_16\_e\_r (2) p\_16\_f\_r (2) p\_16\_g\_r (2) p\_17\_a\_r (2) p\_17\_b\_r (2) p\_17\_c\_r (2) p\_17\_d\_r (2) p\_17\_e\_r (2) p\_17\_f\_r (2) p\_19\_r (2) p\_20\_a\_r (2) p\_20\_b\_r (2) p\_20\_c\_r (2) p\_20\_d\_r (2) p\_20\_e\_r (2) p\_20\_f\_r (2) p\_20\_g\_r (2) p\_20\_h\_r (2) p\_20\_i\_r (2) p\_20\_j\_r (2) p\_20\_k\_r (2) p\_20\_l\_r (2) p\_20\_m\_r (2) . ;

CATEGORICAL ARE p\_18\_c\_r p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r p\_18\_j\_r p\_18\_b\_r p\_18\_d\_r p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r p\_1\_a\_r p\_1\_b\_r p\_2\_a\_r !p\_2\_b\_r p\_2\_c\_r p\_5\_r\_d !p\_6\_a\_r !p\_6\_b\_r !p\_6\_c\_r !p\_6\_d\_r p\_7\_r p\_8\_r p\_9\_a\_r p\_9\_b\_r p\_9\_d\_r p\_3\_r\_d p\_10\_r p\_11\_r\_d p\_9\_c\_r p\_17\_a\_r p\_17\_b\_r p\_17\_d\_r !p\_20\_c\_r !p\_20\_d\_r p\_20\_g\_r p\_20\_k\_r p\_20\_l\_r p\_20\_i\_r p\_20\_e\_r p\_20\_j\_r ;

AUXILIARY ARE id\_user ;

MODEL: PROP\_M1 BY p\_18\_c\_r\* p\_18\_f\_r p\_18\_g\_r p\_18\_i\_r@1 p\_18\_j\_r ;  
PROP\_M2 BY p\_18\_b\_r\* p\_18\_d\_r@1 p\_18\_e\_r p\_18\_h\_r p\_18\_k\_r ;

```

SATPROEV BY p_1_a_r* p_1_b_r p_2_a_r p_2_c_r@1 p_5_r_d p_7_r p_8_r p_9_a_r
p_9_b_r p_9_d_r ;
FLXEPSNI BY p_3_r_d* p_10_r@1 p_11_r_d p_9_c_r p_17_a_r p_17_b_r p_17_d_r;
DIRREF BY p_20_g_r* p_20_k_r p_20_l_r p_20_i_r@1 p_20_j_r p_20_e_r;

```

```

PROP_M1 ON edad cand niv_I niv_II f_m_ct b_q m_cs cs b_ca ing mujer cpi_e_n
cpi_cs_h cpi_dt_s cpi_sec gobierno inst_n_s inst_ut univ ;
PROP_M2 ON edad cand niv_I niv_II f_m_ct b_q m_cs cs b_ca ing mujer cpi_e_n
cpi_cs_h cpi_dt_s cpi_sec gobierno inst_n_s inst_ut univ ;
FLXEPSNI ON PROP_M1 PROP_M2 ;
SATPROEV ON PROP_M1 PROP_M2 FLXEPSNI ;
DIRREF ON SATPROEV ;

```

OUTPUT: stdyx;

\*\*\* WARNING

Data set contains cases with missing on all variables except x-variables. These cases were not included in the analysis. Number of cases with missing on all variables except x-variables: 2

1 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

Modelo de Ecuaciones Estructurales condicional para la relación entre la evaluación del desempeño individual de investigación de los investigadores SNI y la producción de conocimiento.

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	8053
Number of dependent variables	33
Number of independent variables	19
Number of continuous latent variables	5

Observed dependent variables

Binary and ordered categorical (ordinal)

P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R	P_18_B_R
P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R	P_1_A_R	P_1_B_R
P_2_A_R	P_2_C_R	P_5_R_D	P_7_R	P_8_R	P_9_A_R
P_9_B_R	P_9_D_R	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R
P_17_A_R	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_20_I_R	P_20_E_R	P_20_J_R			

Observed independent variables

EDAD	CAND	NIV_I	NIV_II	F_M_CT	B_Q
M_CS	CS	B_CA	ING	MUJER	CPI_E_N
CPI_CS_H	CPI_DT_S	CPI_SEC	GOBIERNO	INST_N_S	INST_UT
UNIV					

Observed auxiliary variables

ID\_USER

Continuous latent variables

PROP_M1	PROP_M2	SATPROEV	FLXEPSNI	DIRREF
---------	---------	----------	----------	--------

Estimator	WLSMV
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.100D-03
Parameterization	DELTA

Input data file(s)  
 enc\_sni\_vivos\_vigentes19022014\_recod4\_mplus.dat

Input data format FREE

SUMMARY OF DATA

Number of missing data patterns 1292  
 COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT					
Covariance Coverage					
	P_18_C_R	P_18_F_R	P_18_G_R	P_18_I_R	P_18_J_R
P_18_C_R	0.972				
P_18_F_R	0.965	0.968			
P_18_G_R	0.969	0.966	0.972		
P_18_I_R	0.969	0.966	0.971	0.973	
P_18_J_R	0.965	0.963	0.967	0.969	0.970
P_18_B_R	0.970	0.967	0.971	0.971	0.968
P_18_D_R	0.968	0.965	0.970	0.970	0.966
P_18_E_R	0.968	0.966	0.970	0.971	0.967
P_18_H_R	0.967	0.965	0.969	0.970	0.966
P_18_K_R	0.968	0.965	0.969	0.971	0.967
P_1_A_R	0.834	0.831	0.834	0.835	0.832
P_1_B_R	0.849	0.846	0.850	0.851	0.847
P_2_A_R	0.948	0.944	0.949	0.950	0.946
P_2_C_R	0.948	0.944	0.949	0.950	0.946
P_5_R_D	0.849	0.846	0.850	0.851	0.847
P_7_R	0.670	0.667	0.670	0.671	0.668
P_8_R	0.614	0.612	0.614	0.615	0.613
P_9_A_R	0.613	0.612	0.614	0.615	0.613
P_9_B_R	0.613	0.612	0.614	0.615	0.613
P_9_D_R	0.613	0.612	0.614	0.615	0.613
P_3_R_D	0.904	0.901	0.905	0.906	0.902
P_10_R	0.834	0.830	0.834	0.834	0.831
P_11_R_D	0.864	0.860	0.865	0.865	0.862
P_9_C_R	0.613	0.612	0.614	0.615	0.613
P_17_A_R	0.910	0.907	0.911	0.912	0.908
P_17_B_R	0.910	0.907	0.911	0.912	0.908
P_17_D_R	0.910	0.907	0.911	0.912	0.908
P_20_G_R	0.760	0.757	0.761	0.762	0.760
P_20_K_R	0.767	0.765	0.768	0.769	0.766
P_20_L_R	0.729	0.726	0.730	0.730	0.728
P_20_I_R	0.778	0.775	0.779	0.780	0.777
P_20_E_R	0.768	0.766	0.770	0.770	0.768
P_20_J_R	0.728	0.725	0.729	0.730	0.728
Covariance Coverage					
	P_18_B_R	P_18_D_R	P_18_E_R	P_18_H_R	P_18_K_R
P_18_B_R	0.975				
P_18_D_R	0.970	0.972			
P_18_E_R	0.971	0.969	0.973		
P_18_H_R	0.970	0.968	0.969	0.971	
P_18_K_R	0.970	0.969	0.969	0.969	0.972
P_1_A_R	0.836	0.834	0.834	0.833	0.834
P_1_B_R	0.851	0.849	0.850	0.849	0.849
P_2_A_R	0.951	0.948	0.949	0.948	0.948
P_2_C_R	0.951	0.948	0.949	0.948	0.948
P_5_R_D	0.852	0.849	0.850	0.849	0.849

P_7_R	0.672	0.670	0.670	0.669	0.670
P_8_R	0.615	0.614	0.615	0.614	0.614
P_9_A_R	0.615	0.614	0.615	0.614	0.614
P_9_B_R	0.615	0.614	0.615	0.614	0.614
P_9_D_R	0.615	0.614	0.615	0.614	0.614
P_3_R_D	0.907	0.904	0.905	0.904	0.905
P_10_R	0.835	0.833	0.834	0.833	0.833
P_11_R_D	0.866	0.864	0.865	0.864	0.864
P_9_C_R	0.615	0.614	0.615	0.614	0.614
P_17_A_R	0.913	0.910	0.911	0.910	0.910
P_17_B_R	0.913	0.910	0.911	0.910	0.910
P_17_D_R	0.913	0.910	0.911	0.910	0.910
P_20_G_R	0.762	0.761	0.761	0.760	0.761
P_20_K_R	0.769	0.768	0.768	0.768	0.768
P_20_L_R	0.730	0.730	0.730	0.729	0.730
P_20_I_R	0.780	0.779	0.779	0.778	0.779
P_20_E_R	0.770	0.769	0.770	0.769	0.770
P_20_J_R	0.730	0.729	0.729	0.728	0.729

Covariance Coverage

	P_1_A_R	P_1_B_R	P_2_A_R	P_2_C_R	P_5_R_D
P_1_A_R	0.856				
P_1_B_R	0.752	0.873			
P_2_A_R	0.855	0.871	0.975		
P_2_C_R	0.855	0.871	0.975	0.975	
P_5_R_D	0.750	0.775	0.853	0.853	0.871
P_7_R	0.596	0.615	0.675	0.675	0.657
P_8_R	0.544	0.569	0.618	0.618	0.606
P_9_A_R	0.543	0.569	0.618	0.618	0.606
P_9_B_R	0.543	0.569	0.618	0.618	0.606
P_9_D_R	0.543	0.569	0.618	0.618	0.606
P_3_R_D	0.796	0.840	0.911	0.911	0.820
P_10_R	0.733	0.751	0.832	0.832	0.756
P_11_R_D	0.757	0.774	0.860	0.860	0.781
P_9_C_R	0.543	0.569	0.618	0.618	0.606
P_17_A_R	0.790	0.806	0.898	0.898	0.804
P_17_B_R	0.790	0.806	0.898	0.898	0.804
P_17_D_R	0.790	0.806	0.898	0.898	0.804
P_20_G_R	0.666	0.683	0.751	0.751	0.685
P_20_K_R	0.669	0.689	0.758	0.758	0.691
P_20_L_R	0.639	0.657	0.722	0.722	0.658
P_20_I_R	0.678	0.699	0.769	0.769	0.702
P_20_E_R	0.671	0.691	0.760	0.760	0.693
P_20_J_R	0.636	0.658	0.719	0.719	0.668

Covariance Coverage

	P_7_R	P_8_R	P_9_A_R	P_9_B_R	P_9_D_R
P_7_R	0.685				
P_8_R	0.552	0.628			
P_9_A_R	0.552	0.628	0.628		
P_9_B_R	0.552	0.628	0.628	0.628	
P_9_D_R	0.552	0.628	0.628	0.628	0.628
P_3_R_D	0.647	0.598	0.597	0.597	0.597
P_10_R	0.605	0.560	0.560	0.560	0.560
P_11_R_D	0.621	0.570	0.570	0.570	0.570
P_9_C_R	0.552	0.628	0.628	0.628	0.628
P_17_A_R	0.634	0.581	0.581	0.581	0.581
P_17_B_R	0.634	0.581	0.581	0.581	0.581
P_17_D_R	0.634	0.581	0.581	0.581	0.581
P_20_G_R	0.556	0.520	0.520	0.520	0.520
P_20_K_R	0.560	0.527	0.526	0.526	0.526
P_20_L_R	0.539	0.505	0.505	0.505	0.505
P_20_I_R	0.569	0.534	0.534	0.534	0.534

P_20_E_R	0.565	0.529	0.528	0.528	0.528
P_20_J_R	0.548	0.518	0.518	0.518	0.518
Covariance Coverage					
	P_3_R_D	P_10_R	P_11_R_D	P_9_C_R	P_17_A_R
P_3_R_D	0.930				
P_10_R	0.793	0.848			
P_11_R_D	0.822	0.830	0.879		
P_9_C_R	0.597	0.560	0.570	0.628	
P_17_A_R	0.856	0.801	0.829	0.581	0.919
P_17_B_R	0.856	0.801	0.829	0.581	0.919
P_17_D_R	0.856	0.801	0.829	0.581	0.919
P_20_G_R	0.722	0.683	0.699	0.520	0.724
P_20_K_R	0.728	0.690	0.707	0.526	0.729
P_20_L_R	0.693	0.655	0.672	0.505	0.693
P_20_I_R	0.738	0.697	0.715	0.534	0.739
P_20_E_R	0.729	0.691	0.707	0.528	0.730
P_20_J_R	0.693	0.657	0.671	0.518	0.693
Covariance Coverage					
	P_17_B_R	P_17_D_R	P_20_G_R	P_20_K_R	P_20_L_R
P_17_B_R	0.919				
P_17_D_R	0.919	0.919			
P_20_G_R	0.724	0.724	0.764		
P_20_K_R	0.729	0.729	0.737	0.771	
P_20_L_R	0.693	0.693	0.704	0.709	0.733
P_20_I_R	0.739	0.739	0.741	0.750	0.714
P_20_E_R	0.730	0.730	0.738	0.742	0.708
P_20_J_R	0.693	0.693	0.701	0.710	0.677
Covariance Coverage					
	P_20_I_R	P_20_E_R	P_20_J_R		
P_20_I_R	0.782				
P_20_E_R	0.751	0.773			
P_20_J_R	0.717	0.706	0.732		

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

P_18_C_R		
Category 1	0.011	83.000
Category 2	0.041	317.000
Category 3	0.180	1405.000
Category 4	0.467	3658.000
Category 5	0.302	2362.000
P_18_F_R		
Category 1	0.025	193.000
Category 2	0.088	690.000
Category 3	0.242	1884.000
Category 4	0.452	3527.000
Category 5	0.193	1503.000
P_18_G_R		
Category 1	0.134	1047.000
Category 2	0.265	2073.000
Category 3	0.321	2511.000
Category 4	0.226	1768.000
Category 5	0.055	432.000
P_18_I_R		
Category 1	0.048	376.000
Category 2	0.105	825.000
Category 3	0.223	1751.000
Category 4	0.373	2925.000
Category 5	0.250	1962.000
P_18_J_R		

Category 1	0.060	472.000
Category 2	0.175	1365.000
Category 3	0.315	2462.000
Category 4	0.330	2574.000
Category 5	0.120	936.000
P_18_B_R		
Category 1	0.031	240.000
Category 2	0.069	538.000
Category 3	0.198	1550.000
Category 4	0.363	2851.000
Category 5	0.340	2669.000
P_18_D_R		
Category 1	0.020	155.000
Category 2	0.052	410.000
Category 3	0.154	1206.000
Category 4	0.364	2850.000
Category 5	0.410	3206.000
P_18_E_R		
Category 1	0.008	61.000
Category 2	0.027	213.000
Category 3	0.122	956.000
Category 4	0.367	2877.000
Category 5	0.476	3725.000
P_18_H_R		
Category 1	0.016	125.000
Category 2	0.055	427.000
Category 3	0.165	1293.000
Category 4	0.380	2973.000
Category 5	0.384	3005.000
P_18_K_R		
Category 1	0.066	518.000
Category 2	0.146	1141.000
Category 3	0.285	2228.000
Category 4	0.334	2613.000
Category 5	0.170	1327.000
P_1_A_R		
Category 1	0.174	1202.000
Category 2	0.826	5692.000
P_1_B_R		
Category 1	0.237	1669.000
Category 2	0.763	5361.000
P_2_A_R		
Category 1	0.027	213.000
Category 2	0.973	7641.000
P_2_C_R		
Category 1	0.196	1542.000
Category 2	0.804	6312.000
P_5_R_D		
Category 1	0.356	2500.000
Category 2	0.644	4514.000
P_7_R		
Category 1	0.219	1208.000
Category 2	0.781	4308.000
P_8_R		
Category 1	0.381	1926.000
Category 2	0.619	3131.000
P_9_A_R		
Category 1	0.044	224.000
Category 2	0.956	4830.000
P_9_B_R		
Category 1	0.045	225.000
Category 2	0.955	4829.000
P_9_D_R		



Category 1	0.137	693.000
Category 2	0.863	4361.000
P_3_R_D		
Category 1	0.506	3786.000
Category 2	0.494	3703.000
P_10_R		
Category 1	0.451	3083.000
Category 2	0.549	3749.000
P_11_R_D		
Category 1	0.530	3753.000
Category 2	0.470	3326.000
P_9_C_R		
Category 1	0.142	719.000
Category 2	0.858	4335.000
P_17_A_R		
Category 1	0.033	242.000
Category 2	0.967	7158.000
P_17_B_R		
Category 1	0.026	189.000
Category 2	0.974	7211.000
P_17_D_R		
Category 1	0.011	85.000
Category 2	0.989	7315.000
P_20_G_R		
Category 1	0.485	2982.000
Category 2	0.515	3172.000
P_20_K_R		
Category 1	0.330	2047.000
Category 2	0.670	4162.000
P_20_L_R		
Category 1	0.701	4134.000
Category 2	0.299	1767.000
P_20_I_R		
Category 1	0.310	1953.000
Category 2	0.690	4345.000
P_20_E_R		
Category 1	0.428	2661.000
Category 2	0.572	3562.000
P_20_J_R		
Category 1	0.439	2588.000
Category 2	0.561	3306.000

WARNING: THE SAMPLE CORRELATION OF P\_9\_D\_R AND P\_8\_R IS 0.986  
DUE TO ONE OR MORE ZERO CELLS IN THEIR BIVARIATE TABLE.  
INFORMATION FROM THESE VARIABLES CAN BE USED TO CREATE ONE NEW VARIABLE.

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 140

Chi-Square Test of Model Fit

Value	15784.388*
Degrees of Freedom	1078
P-Value	0.0000

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.041	
90 Percent C.I.	0.041	0.042
Probability RMSEA <= .05	1.000	

CFI/TLI

CFI	0.925
TLI	0.920

Chi-Square Test of Model Fit for the Baseline Model

Value	197262.432
Degrees of Freedom	1155
P-Value	0.0000

WRMR (Weighted Root Mean Square Residual)

Value	4.581
-------	-------

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.641	0.017	37.523	0.000
P_18_F_R	0.730	0.017	41.884	0.000
P_18_G_R	0.730	0.017	43.741	0.000
P_18_I_R	1.000	0.000	999.000	999.000
P_18_J_R	0.846	0.018	47.644	0.000
PROP_M2 BY				
P_18_B_R	0.909	0.012	75.971	0.000
P_18_D_R	1.000	0.000	999.000	999.000
P_18_E_R	0.719	0.013	55.424	0.000
P_18_H_R	0.784	0.012	66.172	0.000
P_18_K_R	0.744	0.012	63.107	0.000
SATPROEV BY				
P_1_A_R	0.885	0.007	119.525	0.000
P_1_B_R	0.993	0.008	128.842	0.000
P_2_A_R	0.481	0.023	20.992	0.000
P_2_C_R	1.000	0.000	999.000	999.000
P_5_R_D	0.896	0.008	110.234	0.000
P_7_R	0.825	0.010	79.717	0.000
P_8_R	0.977	0.008	118.600	0.000
P_9_A_R	0.373	0.026	14.105	0.000
P_9_B_R	0.340	0.029	11.788	0.000
P_9_D_R	0.672	0.016	43.037	0.000
FLXEPSNI BY				
P_3_R_D	0.614	0.016	37.522	0.000
P_10_R	1.000	0.000	999.000	999.000
P_11_R_D	0.923	0.013	70.198	0.000
P_9_C_R	0.642	0.021	30.639	0.000
P_17_A_R	0.696	0.022	31.199	0.000
P_17_B_R	0.668	0.025	27.022	0.000
P_17_D_R	0.567	0.035	16.289	0.000
DIRREF BY				
P_20_G_R	0.827	0.008	106.436	0.000
P_20_K_R	0.984	0.008	118.651	0.000

P_20_L_R	0.650	0.011	59.454	0.000
P_20_I_R	1.000	0.000	999.000	999.000
P_20_J_R	0.915	0.008	119.134	0.000
P_20_E_R	0.841	0.009	95.119	0.000
FLXEPSNI ON				
PROP_M1	0.474	0.018	26.137	0.000
PROP_M2	0.272	0.017	16.217	0.000
SATPROEV ON				
PROP_M1	0.007	0.019	0.393	0.694
PROP_M2	-0.154	0.016	-9.544	0.000
FLXEPSNI	0.889	0.016	54.284	0.000
DIRREF ON				
SATPROEV	-0.668	0.011	-59.898	0.000
PROP_M1 ON				
EDAD	0.004	0.001	2.595	0.009
CAND	-0.178	0.057	-3.132	0.002
NIV_I	-0.149	0.046	-3.261	0.001
NIV_II	-0.201	0.048	-4.235	0.000
F_M_CT	0.146	0.041	3.540	0.000
B_Q	0.252	0.039	6.512	0.000
M_CS	0.371	0.050	7.367	0.000
CS	0.026	0.039	0.647	0.518
B_CA	0.160	0.043	3.700	0.000
ING	-0.055	0.045	-1.224	0.221
MUJER	-0.108	0.024	-4.455	0.000
CPI_E_N	-0.093	0.085	-1.093	0.274
CPI_CS_H	-0.277	0.093	-2.977	0.003
CPI_DT_S	-0.149	0.137	-1.093	0.275
CPI_SEC	0.031	0.084	0.369	0.712
GOBIERNO	0.257	0.081	3.156	0.002
INST_N_S	0.139	0.089	1.558	0.119
INST_UT	0.152	0.073	2.087	0.037
UNIV	0.110	0.067	1.645	0.100
PROP_M2 ON				
EDAD	0.004	0.001	2.823	0.005
CAND	0.570	0.054	10.528	0.000
NIV_I	0.401	0.044	9.092	0.000
NIV_II	0.171	0.046	3.688	0.000
F_M_CT	-0.137	0.040	-3.435	0.001
B_Q	-0.084	0.037	-2.252	0.024
M_CS	0.303	0.047	6.466	0.000
CS	0.300	0.039	7.610	0.000
B_CA	0.302	0.042	7.256	0.000
ING	0.265	0.043	6.136	0.000
MUJER	0.033	0.023	1.427	0.154
CPI_E_N	0.018	0.081	0.226	0.821
CPI_CS_H	-0.252	0.094	-2.690	0.007
CPI_DT_S	0.452	0.119	3.797	0.000
CPI_SEC	0.232	0.083	2.793	0.005
GOBIERNO	0.170	0.079	2.158	0.031
INST_N_S	0.077	0.085	0.904	0.366
INST_UT	0.003	0.072	0.037	0.970
UNIV	-0.006	0.066	-0.090	0.928
Thresholds				
P_18_C_R\$1	-1.749	0.131	-13.386	0.000
P_18_C_R\$2	-1.074	0.126	-8.530	0.000
P_18_C_R\$3	-0.168	0.125	-1.341	0.180

P_18_C_R\$4	1.098	0.126	8.730	0.000
P_18_F_R\$1	-1.431	0.128	-11.181	0.000
P_18_F_R\$2	-0.665	0.126	-5.275	0.000
P_18_F_R\$3	0.185	0.125	1.476	0.140
P_18_F_R\$4	1.441	0.126	11.453	0.000
P_18_G_R\$1	-0.265	0.125	-2.129	0.033
P_18_G_R\$2	0.601	0.125	4.822	0.000
P_18_G_R\$3	1.452	0.125	11.617	0.000
P_18_G_R\$4	2.483	0.126	19.763	0.000
P_18_I_R\$1	-1.578	0.129	-12.259	0.000
P_18_I_R\$2	-0.922	0.128	-7.204	0.000
P_18_I_R\$3	-0.196	0.128	-1.531	0.126
P_18_I_R\$4	0.818	0.128	6.391	0.000
P_18_J_R\$1	-1.189	0.125	-9.524	0.000
P_18_J_R\$2	-0.338	0.125	-2.711	0.007
P_18_J_R\$3	0.548	0.125	4.382	0.000
P_18_J_R\$4	1.662	0.126	13.240	0.000
P_18_B_R\$1	-1.213	0.130	-9.296	0.000
P_18_B_R\$2	-0.593	0.128	-4.634	0.000
P_18_B_R\$3	0.201	0.128	1.569	0.117
P_18_B_R\$4	1.187	0.128	9.258	0.000
P_18_D_R\$1	-1.439	0.134	-10.754	0.000
P_18_D_R\$2	-0.813	0.131	-6.189	0.000
P_18_D_R\$3	-0.071	0.130	-0.544	0.587
P_18_D_R\$4	0.946	0.131	7.238	0.000
P_18_E_R\$1	-2.166	0.145	-14.978	0.000
P_18_E_R\$2	-1.540	0.139	-11.066	0.000
P_18_E_R\$3	-0.706	0.138	-5.129	0.000
P_18_E_R\$4	0.392	0.137	2.856	0.004
P_18_H_R\$1	-1.615	0.136	-11.892	0.000
P_18_H_R\$2	-0.925	0.132	-7.028	0.000
P_18_H_R\$3	-0.149	0.131	-1.145	0.252
P_18_H_R\$4	0.893	0.131	6.824	0.000
P_18_K_R\$1	-0.649	0.126	-5.145	0.000
P_18_K_R\$2	0.075	0.125	0.599	0.549
P_18_K_R\$3	0.887	0.126	7.053	0.000
P_18_K_R\$4	1.871	0.126	14.807	0.000
P_1_A_R\$1	-1.523	0.200	-7.626	0.000
P_1_B_R\$1	-0.951	0.182	-5.226	0.000
P_2_A_R\$1	-1.687	0.294	-5.734	0.000
P_2_C_R\$1	-1.546	0.188	-8.230	0.000
P_5_R_D\$1	-0.472	0.169	-2.801	0.005
P_7_R\$1	-1.592	0.226	-7.045	0.000
P_8_R\$1	-0.287	0.202	-1.420	0.156
P_9_A_R\$1	-0.997	0.334	-2.987	0.003
P_9_B_R\$1	-1.814	0.414	-4.379	0.000
P_9_D_R\$1	-0.992	0.258	-3.838	0.000
P_3_R_D\$1	-0.283	0.160	-1.771	0.077
P_10_R\$1	-0.588	0.168	-3.504	0.000
P_11_R_D\$1	0.239	0.159	1.497	0.134
P_9_C_R\$1	-1.512	0.251	-6.022	0.000
P_17_A_R\$1	-1.967	0.300	-6.562	0.000
P_17_B_R\$1	-2.632	0.368	-7.160	0.000
P_17_D_R\$1	-2.011	0.548	-3.674	0.000
P_20_G_R\$1	0.473	0.174	2.723	0.006
P_20_K_R\$1	0.004	0.178	0.021	0.983
P_20_L_R\$1	0.782	0.191	4.098	0.000
P_20_I_R\$1	0.007	0.179	0.041	0.967
P_20_E_R\$1	-0.250	0.176	-1.417	0.156
P_20_J_R\$1	0.514	0.179	2.868	0.004
Residual Variances				
PROP_M1	0.608	0.015	41.278	0.000

PROP_M2	0.689	0.011	61.263	0.000
SATPROEV	0.294	0.013	22.709	0.000
FLXEPSNI	0.671	0.017	40.163	0.000
DIRREF	0.563	0.014	40.437	0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1 BY				
P_18_C_R	0.510	0.012	44.286	0.000
P_18_F_R	0.580	0.011	52.794	0.000
P_18_G_R	0.580	0.011	54.573	0.000
P_18_I_R	0.788	0.009	85.571	0.000
P_18_J_R	0.670	0.009	73.100	0.000
PROP_M2 BY				
P_18_B_R	0.771	0.007	111.494	0.000
P_18_D_R	0.843	0.006	132.126	0.000
P_18_E_R	0.616	0.010	64.160	0.000
P_18_H_R	0.670	0.009	78.751	0.000
P_18_K_R	0.636	0.009	74.007	0.000
SATPROEV BY				
P_1_A_R	0.859	0.006	137.879	0.000
P_1_B_R	0.964	0.004	224.619	0.000
P_2_A_R	0.468	0.021	21.893	0.000
P_2_C_R	0.970	0.004	249.289	0.000
P_5_R_D	0.870	0.007	126.031	0.000
P_7_R	0.802	0.010	84.305	0.000
P_8_R	0.949	0.007	137.392	0.000
P_9_A_R	0.363	0.026	14.143	0.000
P_9_B_R	0.331	0.028	11.820	0.000
P_9_D_R	0.653	0.015	43.918	0.000
FLXEPSNI BY				
P_3_R_D	0.572	0.014	39.807	0.000
P_10_R	0.928	0.008	121.384	0.000
P_11_R_D	0.857	0.009	100.549	0.000
P_9_C_R	0.597	0.019	31.651	0.000
P_17_A_R	0.648	0.020	32.627	0.000
P_17_B_R	0.621	0.022	27.823	0.000
P_17_D_R	0.528	0.032	16.470	0.000
DIRREF BY				
P_20_G_R	0.821	0.007	114.286	0.000
P_20_K_R	0.976	0.005	192.645	0.000
P_20_L_R	0.645	0.011	60.882	0.000
P_20_I_R	0.991	0.005	202.967	0.000
P_20_J_R	0.908	0.006	151.954	0.000
P_20_E_R	0.834	0.008	107.078	0.000
FLXEPSNI ON				
PROP_M1	0.407	0.014	28.990	0.000
PROP_M2	0.254	0.015	16.649	0.000
SATPROEV ON				
PROP_M1	0.006	0.016	0.394	0.694
PROP_M2	-0.138	0.014	-9.561	0.000
FLXEPSNI	0.852	0.013	65.137	0.000
DIRREF ON				

SATPROEV	-0.655	0.010	-65.390	0.000
PROP_M1 ON				
EDAD	0.047	0.018	2.600	0.009
CAND	-0.089	0.028	-3.134	0.002
NIV_I	-0.092	0.028	-3.261	0.001
NIV_II	-0.093	0.022	-4.240	0.000
F_M_CT	0.066	0.018	3.547	0.000
B_Q	0.125	0.019	6.562	0.000
M_CS	0.141	0.019	7.439	0.000
CS	0.011	0.017	0.647	0.518
B_CA	0.066	0.018	3.710	0.000
ING	-0.023	0.019	-1.224	0.221
MUJER	-0.065	0.015	-4.470	0.000
CPI_E_N	-0.024	0.022	-1.093	0.274
CPI_CS_H	-0.055	0.018	-2.981	0.003
CPI_DT_S	-0.018	0.017	-1.093	0.275
CPI_SEC	0.008	0.021	0.369	0.712
GOBIERNO	0.074	0.023	3.162	0.002
INST_N_S	0.037	0.024	1.559	0.119
INST_UT	0.061	0.029	2.089	0.037
UNIV	0.066	0.040	1.646	0.100
PROP_M2 ON				
EDAD	0.045	0.016	2.825	0.005
CAND	0.262	0.025	10.643	0.000
NIV_I	0.228	0.025	9.165	0.000
NIV_II	0.073	0.020	3.692	0.000
F_M_CT	-0.056	0.016	-3.438	0.001
B_Q	-0.038	0.017	-2.253	0.024
M_CS	0.106	0.016	6.498	0.000
CS	0.120	0.016	7.664	0.000
B_CA	0.115	0.016	7.303	0.000
ING	0.102	0.017	6.164	0.000
MUJER	0.018	0.013	1.427	0.154
CPI_E_N	0.004	0.020	0.226	0.821
CPI_CS_H	-0.046	0.017	-2.691	0.007
CPI_DT_S	0.050	0.013	3.802	0.000
CPI_SEC	0.053	0.019	2.796	0.005
GOBIERNO	0.045	0.021	2.159	0.031
INST_N_S	0.019	0.021	0.904	0.366
INST_UT	0.001	0.026	0.037	0.970
UNIV	-0.003	0.037	-0.090	0.928
Thresholds				
P_18_C_R\$1	-1.737	0.130	-13.374	0.000
P_18_C_R\$2	-1.066	0.125	-8.526	0.000
P_18_C_R\$3	-0.167	0.125	-1.341	0.180
P_18_C_R\$4	1.090	0.125	8.733	0.000
P_18_F_R\$1	-1.419	0.127	-11.169	0.000
P_18_F_R\$2	-0.659	0.125	-5.272	0.000
P_18_F_R\$3	0.183	0.124	1.476	0.140
P_18_F_R\$4	1.428	0.125	11.464	0.000
P_18_G_R\$1	-0.263	0.123	-2.128	0.033
P_18_G_R\$2	0.595	0.123	4.823	0.000
P_18_G_R\$3	1.439	0.124	11.626	0.000
P_18_G_R\$4	2.461	0.124	19.786	0.000
P_18_I_R\$1	-1.552	0.127	-12.234	0.000
P_18_I_R\$2	-0.906	0.126	-7.196	0.000
P_18_I_R\$3	-0.192	0.126	-1.531	0.126
P_18_I_R\$4	0.804	0.126	6.397	0.000
P_18_J_R\$1	-1.174	0.123	-9.516	0.000
P_18_J_R\$2	-0.334	0.123	-2.710	0.007

P_18_J_R\$3	0.541	0.123	4.383	0.000
P_18_J_R\$4	1.642	0.124	13.249	0.000
P_18_B_R\$1	-1.178	0.127	-9.261	0.000
P_18_B_R\$2	-0.576	0.125	-4.626	0.000
P_18_B_R\$3	0.195	0.124	1.570	0.116
P_18_B_R\$4	1.153	0.124	9.292	0.000
P_18_D_R\$1	-1.390	0.130	-10.698	0.000
P_18_D_R\$2	-0.785	0.127	-6.171	0.000
P_18_D_R\$3	-0.068	0.126	-0.543	0.587
P_18_D_R\$4	0.913	0.126	7.263	0.000
P_18_E_R\$1	-2.127	0.143	-14.916	0.000
P_18_E_R\$2	-1.512	0.137	-11.033	0.000
P_18_E_R\$3	-0.693	0.135	-5.122	0.000
P_18_E_R\$4	0.385	0.135	2.858	0.004
P_18_H_R\$1	-1.580	0.133	-11.842	0.000
P_18_H_R\$2	-0.905	0.129	-7.011	0.000
P_18_H_R\$3	-0.146	0.128	-1.145	0.252
P_18_H_R\$4	0.874	0.128	6.839	0.000
P_18_K_R\$1	-0.636	0.124	-5.138	0.000
P_18_K_R\$2	0.074	0.123	0.599	0.549
P_18_K_R\$3	0.869	0.123	7.065	0.000
P_18_K_R\$4	1.834	0.123	14.859	0.000
P_1_A_R\$1	-1.519	0.199	-7.624	0.000
P_1_B_R\$1	-0.948	0.181	-5.225	0.000
P_2_A_R\$1	-1.685	0.294	-5.734	0.000
P_2_C_R\$1	-1.541	0.187	-8.227	0.000
P_5_R_D\$1	-0.471	0.168	-2.801	0.005
P_7_R\$1	-1.589	0.226	-7.044	0.000
P_8_R\$1	-0.286	0.201	-1.420	0.156
P_9_A_R\$1	-0.997	0.334	-2.987	0.003
P_9_B_R\$1	-1.814	0.414	-4.379	0.000
P_9_D_R\$1	-0.990	0.258	-3.838	0.000
P_3_R_D\$1	-0.282	0.159	-1.771	0.077
P_10_R\$1	-0.585	0.167	-3.502	0.000
P_11_R_D\$1	0.238	0.159	1.497	0.134
P_9_C_R\$1	-1.508	0.250	-6.021	0.000
P_17_A_R\$1	-1.961	0.299	-6.561	0.000
P_17_B_R\$1	-2.626	0.367	-7.159	0.000
P_17_D_R\$1	-2.008	0.547	-3.674	0.000
P_20_G_R\$1	0.473	0.174	2.723	0.006
P_20_K_R\$1	0.004	0.178	0.021	0.983
P_20_L_R\$1	0.782	0.191	4.098	0.000
P_20_I_R\$1	0.007	0.179	0.041	0.967
P_20_E_R\$1	-0.250	0.176	-1.417	0.156
P_20_J_R\$1	0.513	0.179	2.868	0.004

Residual Variances

PROP_M1	0.947	0.006	154.278	0.000
PROP_M2	0.905	0.007	129.691	0.000
SATPROEV	0.310	0.013	23.099	0.000
FLXEPSNI	0.771	0.015	52.014	0.000
DIRREF	0.571	0.013	43.452	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	Residual Variance
P_18_C_R	0.260	0.012	22.143	0.000	0.750
P_18_F_R	0.336	0.013	26.397	0.000	0.676
P_18_G_R	0.336	0.012	27.287	0.000	0.676
P_18_I_R	0.621	0.015	42.786	0.000	0.392
P_18_J_R	0.448	0.012	36.550	0.000	0.565
P_18_B_R	0.594	0.011	55.747	0.000	0.430

P_18_D_R	0.710	0.011	66.063	0.000	0.311
P_18_E_R	0.379	0.012	32.080	0.000	0.644
P_18_H_R	0.448	0.011	39.375	0.000	0.576
P_18_K_R	0.405	0.011	37.004	0.000	0.619
P_1_A_R	0.738	0.011	68.939	0.000	0.263
P_1_B_R	0.929	0.008	112.310	0.000	0.072
P_2_A_R	0.219	0.020	10.947	0.000	0.782
P_2_C_R	0.942	0.008	124.645	0.000	0.059
P_5_R_D	0.757	0.012	63.016	0.000	0.244
P_7_R	0.643	0.015	42.152	0.000	0.359
P_8_R	0.900	0.013	68.696	0.000	0.101
P_9_A_R	0.132	0.019	7.071	0.000	0.869
P_9_B_R	0.109	0.019	5.910	0.000	0.891
P_9_D_R	0.427	0.019	21.959	0.000	0.575
P_3_R_D	0.327	0.016	19.904	0.000	0.676
P_10_R	0.861	0.014	60.692	0.000	0.141
P_11_R_D	0.734	0.015	50.275	0.000	0.269
P_9_C_R	0.357	0.023	15.826	0.000	0.646
P_17_A_R	0.419	0.026	16.314	0.000	0.584
P_17_B_R	0.386	0.028	13.912	0.000	0.617
P_17_D_R	0.279	0.034	8.235	0.000	0.724
P_20_G_R	0.673	0.012	57.143	0.000	0.327
P_20_K_R	0.952	0.010	96.322	0.000	0.048
P_20_L_R	0.416	0.014	30.441	0.000	0.585
P_20_I_R	0.983	0.010	101.484	0.000	0.017
P_20_E_R	0.696	0.013	53.539	0.000	0.304
P_20_J_R	0.824	0.011	75.977	0.000	0.177

Latent Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROP_M1	0.053	0.006	8.689	0.000
PROP_M2	0.095	0.007	13.621	0.000
SATPROEV	0.690	0.013	51.434	0.000
FLXEPSNI	0.229	0.015	15.410	0.000
DIRREF	0.429	0.013	32.695	0.000

#### QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.172E-05  
(ratio of smallest to largest eigenvalue)

#### DIAGRAM INFORMATION

Use View Diagram under the Diagram menu in the Mplus Editor to view the diagram.

If running Mplus from the Mplus Diagrammer, the diagram opens automatically.

#### Diagram output

c:\users\repodrida\documents\nata\documents\doctorado en ciencias sociales\_uam\_x\2016\_capitulos\

Beginning Time: 14:14:34

Ending Time: 14:15:21

Elapsed Time: 00:00:47

#### MUTHEN & MUTHEN

3463 Stoner Ave.

Los Angeles, CA 90066

Tel: (310) 391-9971

Fax: (310) 391-8971

Web: www.StatModel.com

Support: Support@StatModel.com









<b>hombre</b>		<b>- .3285599</b>	<b>.1441886</b>	<b>-2.28</b>	<b>0.023</b>	<b>- .6111644</b>	<b>- .0459554</b>
cpi_exac_nat		.0871043	.4434015	0.20	0.844	-.7819467	.9561552
cpi_ccss_hum		.1736343	.5080927	0.34	0.733	-.8222091	1.169478
cpi_destec_serv		-1.109656	1.071173	-1.04	0.300	-3.209115	.9898041
cpi_sectorizado		-.4233465	.4951841	-0.85	0.393	-1.39389	.5471965
gobierno		-.4814556	.4774694	-1.01	0.313	-1.417278	.4543672
inst_nac_sal		-.1689586	.4886069	-0.35	0.729	-1.12661	.7886933
inst_univ_tec		-.1547477	.4052335	-0.38	0.703	-.9489908	.6394955
universidades		-.5017107	.3766743	-1.33	0.183	-1.239979	.2365573
_cons		-4.635292	.6984567	-6.64	0.000	-6.004242	-3.266342

-----

Logistic regression

Number of obs = 8055  
LR chi2(19) = 26.76  
Prob > chi2 = 0.1104  
Pseudo R2 = 0.0131

Log likelihood = -1006.3784

i_missp_18_e_rec	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
<b>edad</b>		<b>.020944</b>	<b>.0081405</b>	<b>2.57</b>	<b>0.010</b>	<b>.0049889</b>	<b>.0368992</b>
candidato		.0577653	.3495978	0.17	0.869	-.6274338	.7429644
nivel_I		.0942013	.2747939	0.34	0.732	-.4443848	.6327874
nivel_II		.1385606	.2862563	0.48	0.628	-.4224914	.6996126
fismat_cctierra		.3066461	.2764407	1.11	0.267	-.2351677	.8484599
biol_quimica		.4540314	.2486127	1.83	0.068	-.0332406	.9413034
med_ccsalud		-.0205389	.3341958	-0.06	0.951	-.6755506	.6344727
ccss		.2711388	.2628829	1.03	0.302	-.2441022	.7863798
biotec_ccagro		.3439283	.2800923	1.23	0.219	-.2050425	.892899
ingenieria		.4455965	.2864547	1.56	0.120	-.1158443	1.007037
<b>hombre</b>		<b>-.3351591</b>	<b>.1450082</b>	<b>-2.31</b>	<b>0.021</b>	<b>-.6193698</b>	<b>-.0509483</b>
cpi_exac_nat		-.1498153	.4610378	-0.32	0.745	-1.053433	.7538023
cpi_ccss_hum		.0831707	.5077507	0.16	0.870	-.9120024	1.078344
cpi_destec_serv		-1.066509	1.071157	-1.00	0.319	-3.165937	1.03292
cpi_sectorizado		-.3083178	.4859339	-0.63	0.526	-1.260731	.6440951
gobierno		-.4238312	.4772382	-0.89	0.374	-1.359201	.5115384
inst_nac_sal		-.1021234	.4882232	-0.21	0.834	-1.059023	.8547764
inst_univ_tec		-.1197756	.4050601	-0.30	0.767	-.9136788	.6741275
universidades		-.5294585	.3767704	-1.41	0.160	-1.267915	.208998
_cons		-4.357121	.698739	-6.24	0.000	-5.726624	-2.987617

-----

Logistic regression

Number of obs = 8055  
LR chi2(19) = 31.11  
Prob > chi2 = 0.0392  
Pseudo R2 = 0.0148

Log likelihood = -1036.0458

i_missp_18_h_rec	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
<b>edad</b>		<b>.018969</b>	<b>.0080859</b>	<b>2.35</b>	<b>0.019</b>	<b>.003121</b>	<b>.034817</b>
candidato		.1667043	.3478292	0.48	0.632	-.5150283	.848437
nivel_I		.1430586	.2798223	0.51	0.609	-.4053829	.6915002
nivel_II		.2037352	.2904766	0.70	0.483	-.3655884	.7730589
<b>fismat_cctierra</b>		<b>.7128257</b>	<b>.2849383</b>	<b>2.50</b>	<b>0.012</b>	<b>.1543569</b>	<b>1.271294</b>
<b>biol_quimica</b>		<b>.701069</b>	<b>.2655155</b>	<b>2.64</b>	<b>0.008</b>	<b>.1806683</b>	<b>1.22147</b>
med_ccsalud		.357832	.3321241	1.08	0.281	-.2931193	1.008783
<b>ccss</b>		<b>.6055378</b>	<b>.2746987</b>	<b>2.20</b>	<b>0.027</b>	<b>.0671382</b>	<b>1.143937</b>
<b>biotec_ccagro</b>		<b>.5906295</b>	<b>.2955824</b>	<b>2.00</b>	<b>0.046</b>	<b>.0112985</b>	<b>1.16996</b>
<b>ingenieria</b>		<b>.758202</b>	<b>.2981867</b>	<b>2.54</b>	<b>0.011</b>	<b>.1737668</b>	<b>1.342637</b>
<b>hombre</b>		<b>-.4154516</b>	<b>.1418074</b>	<b>-2.93</b>	<b>0.003</b>	<b>-.6933889</b>	<b>-.1375143</b>
cpi_exac_nat		-.0687273	.4555634	-0.15	0.880	-.9616151	.8241606
cpi_ccss_hum		.2006553	.5078802	0.40	0.693	-.7947716	1.196082
cpi_destec_serv		-1.068592	1.071103	-1.00	0.318	-3.167772	1.030587
cpi_sectorizado		-.3361003	.4949901	-0.68	0.497	-1.306263	.6340625
gobierno		-.194751	.4561398	-0.43	0.669	-1.088769	.6992665
inst_nac_sal		.0220167	.4742205	0.05	0.963	-.9074385	.9514718
inst_univ_tec		-.0797495	.4038894	-0.20	0.843	-.8713581	.7118592
universidades		-.4700945	.3764889	-1.25	0.212	-1.207999	.2678101
_cons		-4.574265	.7028407	-6.51	0.000	-5.951807	-3.196722











edad	-.0069264	.0029449	-2.35	0.019	-.0126984	-.0011544
candidato	1.066359	.1333135	8.00	0.000	.805069	1.327648
nivel_I	.8752159	.1145359	7.64	0.000	.6507297	1.099702
nivel_II	.4627046	.1209729	3.82	0.000	.225602	.6998072
fismat_cctierra	.197381	.0896295	2.20	0.028	.0217105	.3730515
biol_quimica	.0628494	.0835491	0.75	0.452	-.1009039	.2266027
med_ccsalud	.0186096	.1044436	0.18	0.859	-.186096	.2233152
ccss	-.2591256	.0892267	-2.90	0.004	-.4340066	-.0842445
biotec_ccagro	-.0691971	.0935688	-0.74	0.460	-.2525885	.1141943
ingenieria	-.0153598	.0954001	-0.16	0.872	-.2023407	.171621
hombre	-.2063075	.0504657	-4.09	0.000	-.3052185	-.1073966
cpi_exac_nat	-.1896156	.1743535	-1.09	0.277	-.5313421	.1521109
cpi_ccss_hum	-.2309307	.2009597	-1.15	0.250	-.6248045	.162943
cpi_destec_serv	-.5098018	.2742096	-1.86	0.063	-1.047243	.0276391
cpi_sectorizado	-.3811026	.1804201	-2.11	0.035	-.7347194	-.0274858
gobierno	-.4543671	.167039	-2.72	0.007	-.7817574	-.1269767
inst_nac_sal	.004813	.176257	0.03	0.978	-.3406443	.3502704
inst_univ_tec	-.5116009	.1519461	-3.37	0.001	-.8094098	-.2137919
universidades	-.3143196	.1380046	-2.28	0.023	-.5848036	-.0438356
_cons	-.5392862	.2560104	-2.11	0.035	-1.041057	-.037515

Logistic regression

Number of obs = 8055  
 LR chi2(19) = 273.40  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.0257

Log likelihood = -5182.0642

i_missp_9_d_rec_m	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
edad	-.0069264	.0029449	-2.35	0.019	-.0126984	-.0011544
candidato	1.066359	.1333135	8.00	0.000	.805069	1.327648
nivel_I	.8752159	.1145359	7.64	0.000	.6507297	1.099702
nivel_II	.4627046	.1209729	3.82	0.000	.225602	.6998072
fismat_cctierra	.197381	.0896295	2.20	0.028	.0217105	.3730515
biol_quimica	.0628494	.0835491	0.75	0.452	-.1009039	.2266027
med_ccsalud	.0186096	.1044436	0.18	0.859	-.186096	.2233152
ccss	-.2591256	.0892267	-2.90	0.004	-.4340066	-.0842445
biotec_ccagro	-.0691971	.0935688	-0.74	0.460	-.2525885	.1141943
ingenieria	-.0153598	.0954001	-0.16	0.872	-.2023407	.171621
hombre	-.2063075	.0504657	-4.09	0.000	-.3052185	-.1073966
cpi_exac_nat	-.1896156	.1743535	-1.09	0.277	-.5313421	.1521109
cpi_ccss_hum	-.2309307	.2009597	-1.15	0.250	-.6248045	.162943
cpi_destec_serv	-.5098018	.2742096	-1.86	0.063	-1.047243	.0276391
cpi_sectorizado	-.3811026	.1804201	-2.11	0.035	-.7347194	-.0274858
gobierno	-.4543671	.167039	-2.72	0.007	-.7817574	-.1269767
inst_nac_sal	.004813	.176257	0.03	0.978	-.3406443	.3502704
inst_univ_tec	-.5116009	.1519461	-3.37	0.001	-.8094098	-.2137919
universidades	-.3143196	.1380046	-2.28	0.023	-.5848036	-.0438356
_cons	-.5392862	.2560104	-2.11	0.035	-1.041057	-.037515

Logistic regression

Number of obs = 7978  
 LR chi2(18) = 152.98  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.0374

Log likelihood = -1966.4929

i_missp_3_r_d_m	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
edad	.0002415	.005742	0.04	0.966	-.0110126	.0114956
candidato	.887691	.2339708	3.79	0.000	.4291166	1.346265
nivel_I	.1390571	.2015226	0.69	0.490	-.25592	.5340341
nivel_II	-.6287564	.2394671	-2.63	0.009	-1.098103	-.1594095
fismat_cctierra	-.3459338	.1647057	-2.10	0.036	-.6687511	-.0231166
biol_quimica	-.1467928	.1474084	-1.00	0.319	-.4357079	.1421224
med_ccsalud	-.2678547	.190791	-1.40	0.160	-.6417982	.1060888
ccss	-.1285714	.1532679	-0.84	0.402	-.428971	.1718282
biotec_ccagro	-.7483937	.1866328	-4.01	0.000	-1.114187	-.3826001
ingenieria	-.4961554	.1785184	-2.78	0.005	-.846045	-.1462658

hombre		.0104085	.0940248	0.11	0.912	-1.1738767	.1946936
<b>cpi_exac_nat</b>		<b>-1.023254</b>	<b>.3105162</b>	<b>-3.30</b>	<b>0.001</b>	<b>-1.631854</b>	<b>-.4146531</b>
<b>cpi_ccss_hum</b>		<b>-1.20753</b>	<b>.3765925</b>	<b>-3.21</b>	<b>0.001</b>	<b>-1.945637</b>	<b>-.4694219</b>
<b>cpi_destec_serv</b>		<b>0</b>	<b>(omitted)</b>				
<b>cpi_sectorizado</b>		<b>-.6611541</b>	<b>.2916815</b>	<b>-2.27</b>	<b>0.023</b>	<b>-1.232839</b>	<b>-.0894688</b>
<b>gobierno</b>		<b>-.8197951</b>	<b>.2544309</b>	<b>-3.22</b>	<b>0.001</b>	<b>-1.31847</b>	<b>-.3211197</b>
<b>inst_nac_sal</b>		<b>-1.074365</b>	<b>.2948879</b>	<b>-3.64</b>	<b>0.000</b>	<b>-1.652335</b>	<b>-.4963952</b>
<b>inst_univ_tec</b>		<b>-.956316</b>	<b>.2269031</b>	<b>-4.21</b>	<b>0.000</b>	<b>-1.401038</b>	<b>-.5115942</b>
<b>universidades</b>		<b>-.7944323</b>	<b>.187827</b>	<b>-4.23</b>	<b>0.000</b>	<b>-1.162566</b>	<b>-.4262982</b>
<b>_cons</b>		<b>-1.757498</b>	<b>.450966</b>	<b>-3.90</b>	<b>0.000</b>	<b>-2.641375</b>	<b>-.8736211</b>

-----

Logistic regression	Number of obs	=	8055
	LR chi2(19)	=	34.48
	Prob > chi2	=	0.0161
Log likelihood = -3413.1596	Pseudo R2	=	0.0050

i_missp_10_rec_m	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
<b>edad</b>	<b>.016144</b>	<b>.0037782</b>	<b>4.27</b>	<b>0.000</b>	<b>.0087389</b> <b>.0235491</b>
candidato	.2497878	.1603925	1.56	0.119	-.0645758 .5641515
nivel_I	.2195192	.1294626	1.70	0.090	-.034223 .4732613
nivel_II	.0051428	.1383026	0.04	0.970	-.2659254 .276211
<b>fismat_cctierra</b>	<b>.2783668</b>	<b>.117889</b>	<b>2.36</b>	<b>0.018</b>	<b>.0473085</b> <b>.509425</b>
biol_quimica	.1311512	.1119154	1.17	0.241	-.0881989 .3505014
med_ccsalud	.2396212	.1379926	1.74	0.082	-.0308393 .5100818
ccss	.065656	.1165879	0.56	0.573	-.1628522 .2941642
biotec_ccagro	.0264699	.1265448	0.21	0.834	-.2215533 .274493
ingenieria	.059256	.1299398	0.46	0.648	-.1954213 .3139332
hombre	-.077155	.0673599	-1.15	0.252	-.2091779 .0548679
cpi_exac_nat	-.1891603	.2313055	-0.82	0.413	-.6425108 .2641902
cpi_ccss_hum	-.1514703	.2652395	-0.57	0.568	-.6713302 .3683897
cpi_destec_serv	-.3916028	.397211	-0.99	0.324	-1.170122 .3869164
cpi_sectorizado	-.1718991	.2358433	-0.73	0.466	-.6341434 .2903452
<b>gobierno</b>	<b>-.4698386</b>	<b>.2295888</b>	<b>-2.05</b>	<b>0.041</b>	<b>-.9198244</b> <b>-.0198527</b>
inst_nac_sal	-.2289302	.2351731	-0.97	0.330	-.689861 .2320006
inst_univ_tec	-.179195	.1998093	-0.90	0.370	-.570814 .2124241
universidades	-.192012	.1822343	-1.05	0.292	-.5491847 .1651607
<b>_cons</b>	<b>-2.536287</b>	<b>.3271865</b>	<b>-7.75</b>	<b>0.000</b>	<b>-3.177561</b> <b>-1.895014</b>

-----

Logistic regression	Number of obs	=	8055
	LR chi2(19)	=	37.82
	Prob > chi2	=	0.0063
Log likelihood = -2955.3483	Pseudo R2	=	0.0064

i_missp_11_r_d_m	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
<b>edad</b>	<b>.0169299</b>	<b>.0041402</b>	<b>4.09</b>	<b>0.000</b>	<b>.0088153</b> <b>.0250445</b>
candidato	.0637274	.1727379	0.37	0.712	-.2748327 .4022876
nivel_I	.0409701	.1366193	0.30	0.764	-.2267987 .3087389
nivel_II	-.0915448	.145721	-0.63	0.530	-.3771527 .1940631
fismat_cctierra	.1235633	.1304825	0.95	0.344	-.1321777 .3793043
biol_quimica	.1080306	.1212798	0.89	0.373	-.1296735 .3457347
med_ccsalud	.1501455	.1513056	0.99	0.321	-.146408 .4466989
ccss	.0507109	.1262592	0.40	0.688	-.1967525 .2981743
biotec_ccagro	.0360127	.1367949	0.26	0.792	-.2321003 .3041257
ingenieria	-.0632399	.1448337	-0.44	0.662	-.3471087 .220629
<b>hombre</b>	<b>-.1453773</b>	<b>.0735804</b>	<b>-1.98</b>	<b>0.048</b>	<b>-.2895923</b> <b>-.0011623</b>
cpi_exac_nat	.2656454	.2662448	1.00	0.318	-.2561848 .7874757
cpi_ccss_hum	.289259	.2958164	0.98	0.328	-.2905305 .8690486
cpi_destec_serv	-.1012223	.452811	-0.22	0.823	-.9887156 .786271
cpi_sectorizado	.1242053	.2756637	0.45	0.652	-.4160856 .6644962
gobierno	-.129464	.2695536	-0.48	0.631	-.6577794 .3988514
inst_nac_sal	.1289267	.2734399	0.47	0.637	-.4070057 .6648591
inst_univ_tec	.131742	.2394288	0.55	0.582	-.3375298 .6010138
universidades	.0472117	.222309	0.21	0.832	-.3885058 .4829293
<b>_cons</b>	<b>-2.859054</b>	<b>.3694316</b>	<b>-7.74</b>	<b>0.000</b>	<b>-3.583127</b> <b>-2.134982</b>





hombre		-.1363402	.0569651	-2.39	0.017	-.2479898	-.0246906
cpi_exac_nat		-.2231028	.1893274	-1.18	0.239	-.5941778	.1479721
cpi_ccss_hum		-.6354944	.2339754	-2.72	0.007	-1.094078	-.1769112
cpi_destec_serv		-.9438323	.3571585	-2.64	0.008	-1.64385	-.2438145
cpi_sectorizado		-.4206793	.1983364	-2.12	0.034	-.8094115	-.0319471
gobierno		-.460074	.184786	-2.49	0.013	-.8222479	-.0979
inst_nac_sal		-.339413	.1959993	-1.73	0.083	-.7235646	.0447386
inst_univ_tec		-.3328138	.1642864	-2.03	0.043	-.6548092	-.0108184
universidades		-.3599732	.1491964	-2.41	0.016	-.6523929	-.0675536
_cons		-1.661591	.2779424	-5.98	0.000	-2.206348	-1.116834

Logistic regression  
 Number of obs = 8055  
 LR chi2(19) = 61.10  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.0070

i_missp_20_k_rec_m	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
edad		.0089528	.0033023	2.71	0.007	.0024803	.0154252
candidato		.5106179	.1391904	3.67	0.000	.2378097	.7834262
nivel_I		.2295983	.1156479	1.99	0.047	.0029326	.456264
nivel_II		.1057879	.1223541	0.86	0.387	-.1340217	.3455976
fismat_cctierra		.2766249	.0991578	2.79	0.005	.0822792	.4709705
biol_quimica		-.0395579	.0952906	-0.42	0.678	-.2263242	.1472083
med_ccsalud		.0502046	.1184259	0.42	0.672	-.181906	.2823152
ccss		-.1449307	.10079	-1.44	0.150	-.3424756	.0526141
biotec_ccagro		-.1068484	.1073535	-1.00	0.320	-.3172573	.1035606
ingenieria		.0195065	.1083016	0.18	0.857	-.1927607	.2317736
hombre		-.1190647	.057555	-2.07	0.039	-.2318705	-.0062589
cpi_exac_nat		-.263184	.191635	-1.37	0.170	-.6387816	.1124136
cpi_ccss_hum		-.6591933	.2373259	-2.78	0.005	-1.124344	-.194043
cpi_destec_serv		-.8137054	.347016	-2.34	0.019	-1.493844	-.1335666
cpi_sectorizado		-.4812347	.201794	-2.38	0.017	-.8767436	-.0857258
gobierno		-.5204639	.1868797	-2.79	0.005	-.8867414	-.1541864
inst_nac_sal		-.4371648	.1987618	-2.20	0.028	-.8267308	-.0475987
inst_univ_tec		-.364566	.1654872	-2.20	0.028	-.6889148	-.0402171
universidades		-.3313842	.149795	-2.21	0.027	-.624977	-.0377914
_cons		-1.472778	.2803425	-5.25	0.000	-2.022239	-.9233165

Logistic regression  
 Number of obs = 8055  
 LR chi2(19) = 110.13  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.0118

i_missp_20_l_rec_m	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
edad		.0104977	.003138	3.35	0.001	.0043473	.0166481
candidato		.6084019	.1335684	4.55	0.000	.3466126	.8701912
nivel_I		.3284787	.1109316	2.96	0.003	.1110568	.5459007
nivel_II		.1376798	.1174764	1.17	0.241	-.0925697	.3679293
fismat_cctierra		.0399603	.093094	0.43	0.668	-.1425006	.2224213
biol_quimica		-.273411	.0886867	-3.08	0.002	-.4472337	-.0995883
med_ccsalud		-.2588547	.1124097	-2.30	0.021	-.4791737	-.0385358
ccss		-.218486	.0917628	-2.38	0.017	-.3983377	-.0386343
biotec_ccagro		-.3821877	.1010734	-3.78	0.000	-.580288	-.1840874
ingenieria		-.3565386	.1037189	-3.44	0.001	-.5598239	-.1532532
hombre		-.1775233	.0544455	-3.26	0.001	-.2842345	-.0708122
cpi_exac_nat		-.3571021	.1841478	-1.94	0.052	-.7180251	.003821
cpi_ccss_hum		-.4708004	.2115681	-2.23	0.026	-.8854664	-.0561345
cpi_destec_serv		-1.022567	.3445142	-2.97	0.003	-1.697802	-.3473312
cpi_sectorizado		-.5972403	.1933264	-3.09	0.002	-.9761531	-.2183275
gobierno		-.6440343	.1795906	-3.59	0.000	-.9960255	-.2920431
inst_nac_sal		-.4620922	.1892247	-2.44	0.015	-.8329657	-.0912187
inst_univ_tec		-.5891729	.1595669	-3.69	0.000	-.9019182	-.2764275
universidades		-.4213243	.1431773	-2.94	0.003	-.7019466	-.140702
_cons		-1.083564	.2664206	-4.07	0.000	-1.605739	-.5613896



edad	.0101846	.0033224	3.07	0.002	.0036727	.0166964
candidato	.6962093	.1433567	4.86	0.000	.4152353	.9771833
nivel_I	.4466815	.1203298	3.71	0.000	.2108395	.6825235
nivel_II	.1538156	.1282425	1.20	0.230	-.097535	.4051663
fismat_cctierra	.5202139	.1013825	5.13	0.000	.3215078	.7189201
biol_quimica	.1498718	.097766	1.53	0.125	-.0417461	.3414898
med_ccsalud	.1939686	.1206236	1.61	0.108	-.0424493	.4303866
ccss	.0375596	.1031986	0.36	0.716	-.1647059	.2398251
biotec_ccagro	.1441498	.1085666	1.33	0.184	-.0686369	.3569364
ingenieria	.153451	.1112628	1.38	0.168	-.0646201	.3715221
hombre	-.1544586	.0577199	-2.68	0.007	-.2675876	-.0413296
cpi_exac_nat	-.2895129	.1904563	-1.52	0.128	-.6628004	.0837747
cpi_ccss_hum	-.9225192	.2525507	-3.65	0.000	-1.417509	-.4275289
cpi_destec_serv	-.9752466	.3571609	-2.73	0.006	-1.675269	-.2752241
cpi_sectorizado	-.5651689	.2022685	-2.79	0.005	-.9616079	-.1687299
gobierno	-.4530698	.1840872	-2.46	0.014	-.8138741	-.0922656
inst_nac_sal	-.4838622	.1984321	-2.44	0.015	-.872782	-.0949424
inst_univ_tec	-.4750696	.1654768	-2.87	0.004	-.7993982	-.150741
universidades	-.3599996	.1488626	-2.42	0.016	-.6517649	-.0682344
_cons	-1.814951	.2835337	-6.40	0.000	-2.370666	-1.259235