

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE PSICOLOGÍA**  
Departamento de Psicobiología y Metodología en Ciencias del  
Comportamiento



**TESIS DOCTORAL**

**Patrones de dominio de atributos cognitivos en comprensión lectora: aplicación de un modelo de diagnóstico cognitivo-psicométrico**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Pablo Cáceres Serrano**

Director

**Jesús M<sup>a</sup> Alvarado Izquierdo**

Madrid  
Ed. electrónica 2019

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE PSICOLOGÍA**

**Departamento de Psicobiología y Metodología en Ciencias del  
Comportamiento**



**TESIS DOCTORAL**

**PATRONES DE DOMINIO DE ATRIBUTOS COGNITIVOS EN  
COMPRESIÓN LECTORA: APLICACIÓN DE UN MODELO DE  
DIAGNÓSTICO COGNITIVO-PSICOMÉTRICO**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR**

**Pablo Cáceres Serrano**

Director: Jesús M<sup>a</sup> Alvarado Izquierdo

Madrid, 2018



## AGRADECIMIENTOS

Por supuesto que los primeros en los que pienso al intentar reconocer a quienes me han ayudado a llegar a esta instancia es en mis padres, porque son quienes me dieron la opción de formarme, de recibir la educación necesaria para llegar a comprender y compartir este desafío académico fuera de mi país.

También agradezco, especialmente en este año 2018, a quien fue mi mentora y guía al finalizar mi licenciatura en psicología: la Dra. Cecilia Quaas. Es a ella a quien le debo haberme dedicado a la Metodología, mi actual pasión. Y la recuerdo especialmente porque apenas hace un par de meses ella nos ha dejado producto de una repentina y agresiva enfermedad. Resulta paradójal, sobre todo para alguien como yo que aún conserva el pensamiento mágico, que haya debido morir para que yo pudiera, simbólicamente, tomar en parte su lugar.

No obstante, hay una persona en particular a la que le debo haber finalizado este proyecto, a mi tutor el Dr. Jesús Alvarado. Ha sido su motivación, su persistencia, su guía y sobretodo su paciencia la que me ha brindado la fuerza suficiente para culminar este proceso. Jesús me ha impulsado, me ha enseñado, se ha enojado conmigo y seguramente en silencio, más de una vez, me habrá mandado a freír monos producto de mi ansiedad y latencia. Pero a pesar de todo, su tesón terminó llevándome por el camino que debí seguir hace años y su paciencia fue la suficiente para comprender que mis exasperantes plazos tenían, por fin, un cierre definitivo.

Gracias Jesús por tu enorme, impagable apoyo.

A fuerza de extenderme un poco más quiero agradecer también a la profesora Carmen Santisteban, hoy retirada, que en virtud de su exigencia, su disciplina y sobretodo su fuerte carácter, me ayudó a comprender y valorar las cosas importantes de la vida.

Al finalizar, no quiero olvidar otro apoyo, que aunque no académico, ha sido tremendamente valioso. Gracias Javier Valera por los buenos momentos y el alojamiento, gracias Pablo Gaviño por las conversaciones y gracias Iwin Leenen, por las enseñanzas al pasar.



# ÍNDICE GENERAL

AGRADECIMIENTOS.....	3
ÍNDICE GENERAL.....	5
ÍNDICE DE TABLAS.....	8
RESUMEN.....	9
ABSTRACT.....	11
PRESENTACIÓN.....	13
INTRODUCCIÓN TEÓRICA.....	19
CAPÍTULO 1: Calidad de la educación y la política del sistema educativo.....	20
Antecedentes económicos de desigualdad.....	23
Antecedentes sociales de desigualdad.....	27
Dimensión cognitiva de la desigualdad.....	34
CAPÍTULO 2: Comprensión lectora y la (des)igualdad.....	37
La influencia de la familia en el desarrollo comprensivo lector.....	40
La influencia de la escuela en el desarrollo comprensivo lector.....	41
CAPÍTULO 3: Un modelo teórico para la comprensión lectora.....	45
Tipos de modelos en comprensión de lectura.....	45
Introducción al modelo comprensivo lector de Kintsch y Van Dijk.....	50
CAPÍTULO 4: Modelos de diagnóstico cognitivo en comprensión lectora.....	57
a) Sistema de diseño cognitivo:.....	61
b) Diseño centrado en evidencia:.....	65
Breve génesis de los modelos de diagnóstico cognitivo.....	70
Etapas de desarrollo del diagnóstico cognitivo.....	73
Construcción de la matriz Q.....	78
Descripción de la matriz Q.....	78
Especificación de la matriz Q.....	81
Tipologías de modelos de diagnóstico cognitivo utilizados en comprensión de lectura.....	86
Criterios de clasificación general de los modelos de diagnóstico cognitivo.....	89
a) Los modelos de diagnóstico cognitivo en función del tipo de relación entre atributos:.....	89
b) Los modelos de diagnóstico cognitivo en función de la escala de la variable respuesta:.....	91
c) Los modelos de diagnóstico cognitivo en función de la escala de los atributos latentes:.....	91
Tipos de modelos de diagnóstico cognitivo aplicados en comprensión de lectura.....	92
a) El modelo Deterministic-Input, Noisy-And-Gate DINA:.....	94
b) El modelo Deterministic-Input, Noisy-OR-Gate DINO:.....	96
c) El modelo Reduced Reparameterized Unified Model (RRUM):.....	97
d) El modelo Generalized Deterministic Inputs, Noisy and Gate (G-DINA):.....	98
e) El modelo Generalized Deterministic-Input, Noisy-And-Gate Model GDM:.....	100

PREGUNTA DE INVESTIGACIÓN .....	104
OBJETIVOS DE INVESTIGACIÓN .....	105
PRESENTACIÓN DE LOS ESTUDIOS .....	106
Estudio 1: El efecto de factores contextuales y económicos sobre el grado de comprensión en lectura.....	107
Introducción.....	107
Metodología.....	111
Muestra.....	111
Diseño .....	112
Instrumentos .....	113
Análisis.....	114
Resultados.....	115
Discusión .....	118
Estudio 2: Evaluación de la comprensión lectora mediante un modelo de diagnóstico cognitivo.....	121
Introducción.....	121
Relevancia de los modelos de diagnóstico cognitivo .....	121
Modelos de diagnóstico cognitivo aplicados en el área de lenguaje .....	123
Método.....	126
Muestra.....	126
Diseño .....	127
Instrumento .....	127
Procedimiento.....	127
Resultados.....	129
Estimación de la dificultad y discriminación de los ítems mediante CDM.....	131
Índices de ajuste de los modelos .....	134
Patrones de atributos obtenidos de acuerdo al modelo.....	134
Discusión .....	135
Estudio 3: Análisis de la Comprensión Lectora en función de la condición pública o privada del centro educativo .....	137
Introducción.....	137
Escuelas vulnerables y efectos sobre la comprensión lectora.....	137
Diferencias en habilidades cognitivas mediante funcionamiento diferencial del ítem.....	139
Método.....	142
Participantes .....	142
Diseño .....	142
Instrumento .....	142
Procedimiento.....	143

Resultados.....	145
Discusión .....	149
DISCUSIÓN FINAL .....	151
CONCLUSIONES.....	161
REFERENCIAS BIBLIOGRÁFICAS .....	162

## ÍNDICE DE TABLAS

Tabla 1. Ejemplo de matriz Q con 4 ítems y 3 atributos cognitivos.....	80
Tabla 2. Distribución de frecuencia conjunta de estudiantes por variables de estratificación.....	112
Tabla 3. Variables de nivel individual y grupal incluidas inicialmente en el modelo.	113
Tabla 4. Estadísticos descriptivos de los factores individuales y la variable resultado organizados por la variable grupal grupo socio-económico.....	115
Tabla 5. Estadísticos descriptivos de los factores individuales y la variable resultado organizados por variable grupo socio-económico.....	117
Tabla 6. Atributos subyacentes en la comprensión de lectura en el test CL-PT. ....	130
Tabla 7. Matriz Q de la prueba de comprensión lectora CL-PT.....	131
Tabla 8. Parámetros de los ítems obtenidos mediante ajuste del modelo DINA.....	132
Tabla 9. Índices de ajuste relativos y absolutos para los modelos de diagnóstico cognitivo. ....	134
Tabla 10. Patrones de atributos en porcentajes estimados por el modelo G-DINA .....	135
Tabla 11. Matriz Q de la prueba de comprensión de lectura CL-PT.....	144
Tabla 12. Parámetros estimados Modelo G-DINA. ....	146
Tabla 13. Análisis general de DIF entre colegios públicos y privados. ....	147
Tabla 14. Patrones de dominio de competencias más frecuentes.....	148
Tabla 15. Probabilidad de dominio de los atributos en centros privados y públicos ...	148

## RESUMEN

La tesis doctoral reúne tres estudios del área psicoeducativa independientes pero vinculados teóricamente entre sí, que buscan aportar evidencia en torno a las diferencias cognitivas que pueden apreciarse entre escuelas cuando el factor social y económico es un determinante del desempeño. El problema proviene de la necesidad de darle un sentido a las habituales y esperables diferencias que se generan por efecto de la condición socioeconómica en diversos grupos humanos. El trabajo se ha centrado en el estudio de la comprensión lectora pues además de ser una habilidad fundamental para el desarrollo futuro de muchas otras disciplinas académicas, representa una de las medidas consideradas para evaluar la calidad de la educación en muchas mediciones internacionales. Se expone la desigualdad en la adquisición y desempeño de esta importante habilidad, fruto de diferencias generadas por el factor socioeconómico, lo que perjudica a los más desfavorecidos, situación que se ha perpetuado en el tiempo en el contexto local chileno.

La pretensión de fondo es tender un puente entre las diferencias que se observan en el nivel social más amplio y a las cuales se les atribuyen las diferencias observadas entre escuelas, y los efectos psicológicos que estas diferencias pueden llegar a tener - específicamente en el ámbito cognitivo- entre estudiantes de los centros comparados.

Para dar cuenta de esta pretensión, el trabajo se ha estructurado en una serie de tres estudios. El primero de ellos se hace cargo de las diferencias entre escuelas a nivel nacional. Mediante análisis multinivel sobre datos secundarios, se ha determinado que factores de orden contextual y económico tienen mayor efecto en las diferencias entre escuelas privadas y públicas en lo relativo a calidad de educación, que ha sido evaluada mediante una prueba estandarizada (Simce) de aplicación nacional. Los resultados indican que aspectos como la educación de los padres, sus expectativas y el ingreso familiar tienen un papel de relevancia al momento de explicar la variabilidad en comprensión de lectura. No obstante lo anterior, hay una proporción de varianza no despreciable que puede ser atribuida al hecho de pertenecer a las escuelas, siendo el factor económico la dimensión que cobija muchos de los aspectos que establecen una diferencia entre ambos tipos de centros. Los resultados hallados en este primer estudio refrendan

muchos otros que fueron realizados en el mismo contexto educativo, dejando establecido que estas diferencias existen y que afectan seriamente la calidad de la educación, en particular cuando se atiende al desarrollo comprensivo lector.

El segundo estudio se efectuó a partir de la recolección de datos primarios, incorporando tanto escuelas públicas como privadas. Se midió a través de un instrumento de carácter comercial (CL-PT) -validado para población chilena - el nivel de comprensión de lectura de estudiantes de 9 a 10 años de educación primaria, conservando en lo posible la composición escolar del primer estudio. El propósito fue establecer una evaluación haciendo uso de los modelos de diagnóstico cognitivo que en diversas investigaciones se han aplicado a la habilidad lectora. Con la aplicación de estos modelos se fue más allá de la simple medición de resultados, identificando los atributos específicos que explican el desempeño en lectura y los patrones de dominio de dichos atributos que denotan los estudiantes de ambos tipos de centros. El estudio enfocó su atención en el ajuste de los modelos siendo G-DINA -por su flexibilidad- el que mejor dio cuenta de la varianza de los datos.

El tercer y último estudio se hizo cargo de las diferencias entre escuelas públicas y privadas pero a nivel de atributos cognitivos, justamente aquellos que fueron identificados en el estudio previo. La comparación se llevó a cabo con los mismos datos primarios del segundo estudio, evaluando en primer lugar la presencia o no de DIF. Con base en los hallazgos del primer estudio, que corroboró la existencia de diferencias en el desempeño comprensivo lector fruto de factores económicos y sociales, se esperaba obtener también diferencias a nivel cognitivo, y con ello aportar alguna evidencia en torno a las consecuencias psicológicas atribuibles a las condiciones que influyen en la calidad de vida y de educación de los estudiantes. Los resultados, en efecto, mostraron diferencias en atributos cognitivos específicos, permitiendo establecer que entre estudiantes de escuelas privadas hay un mayor desarrollo de ciertas habilidades cognitivas que explican su mejor desempeño frente a sus pares de escuelas públicas.

La tesis finaliza con una discusión general de los resultados obtenidos, su relación con el marco teórico propuesto y una reflexión en torno a las proyecciones futuras del trabajo, así como las limitaciones que es necesario considerar en la ponderación de los hallazgos obtenidos.

## **ABSTRACT**

This doctoral thesis brings together three independent psychoeducational studies, but linked theoretically with each other. These studies seek to provide evidence about the cognitive differences that can be seen between schools when social and economic factors are a determinant of educational performance. The problem comes from the need to give a sense to the usual and expected differences that are generated by the effect of socioeconomic status in various human groups. The work has focused on the study of reading comprehension because it's a fundamental skill for the future development of many other academic disciplines, and it represents one of the measures considered to evaluate the quality of education in many international measurements. It exposes the inequality in the acquisition and performance of this important skill, the result of differences generated by the socioeconomic factor, which harms the most disadvantaged, a situation that has been perpetuated over time in the Chilean local context.

The underlying aim is to build a bridge between the differences observed in the broader social level and to which the differences observed between schools are attributed, and the psychological effects that these differences may have -specifically in the cognitive domain- between students of the compared centers.

To account for this claim, the work has been structured in a set of three studies. The first of them takes care of the differences between schools nationwide. Through multilevel analysis on secondary data, it has been determined which contextual and economic factors have the greatest effect on the differences between private and public schools in terms of quality of education, which has been evaluated using a standardized test (Simce) of national application. The results indicate that aspects such as parents' education, their expectations and family income play a relevant role when explaining the variability in reading comprehension. Notwithstanding the above, there is a non-negligible proportion of variance that can be attributed to the fact of belonging to schools, with the economic factor being the dimension that covers many of the aspects that establish a difference between both types of centers. The results found in this first study endorse many others that were made in the same educational context, stating that these

differences exist and that they seriously affect the quality of education, particularly when attending to the comprehensive development of the reader.

The second study was carried out based on the collection of primary data, incorporating both public and private schools. The reading comprehension level of students from 9 to 10 years old, belonging to primary education was measured through a commercial instrument validated for the Chilean population (CL-PT), preserving as much as possible the school composition of the first study. The purpose was to establish an evaluation of cognitive character making use of the models of cognitive diagnosis that in diverse investigations have been applied to the reading ability. With the application of these models, it went beyond the simple measurement of final results, identifying the specific attributes that explain the performance in reading and the domain patterns of those attributes that denote the students of both types of schools. The study focused on the adjustment of the models being G-DINA -for its flexibility- the one that best accounted for the variance of the data.

The third and last study took care of the differences between public and private schools but at the level of cognitive attributes, precisely those that were identified in the previous study. The comparison was carried out with the same primary data of the second study, first evaluating the presence or absence of Differential Item Functioning (DIF). Based on the findings of the first study, which corroborated the existence of differences in the reader's comprehensive performance resulting from economic and social factors, it was expected to also obtain differences at the cognitive level, and thereby provide some evidence about the psychological consequences attributable to the conditions that influence the quality of life and education of these students. The results, in effect, showed differences in specific cognitive attributes, making it possible to establish that among private school students there is a greater development of certain cognitive skills that explain their better performance compared to their peers in public schools.

The thesis ends with a general discussion of the results obtained, their relationship with the theoretical framework proposed and a reflection on the future projections of the work, as well as the limitations that need to be considered in the weighting of the findings obtained.

## **PRESENTACIÓN**

El presente proyecto de investigación nace de una inquietud inicial cuyo origen está en la particular situación que caracteriza a un país como Chile, enfrentado desde hace años a la consolidación de un modelo social y económico no exento de controversias y que sin duda ha tenido un impacto relevante en cómo se concibe la educación y la evaluación formal del logro académico.

Dicha inquietud generó un conjunto de interrogantes relacionadas con las diferencias halladas dentro del sistema escolar y como ello afectaba a los estudiantes y sus posibilidades futuras. Áreas como la salud y la educación se vieron influidas por la visión económica a la que el país adhirió. La educación pública que durante años se sostuvo a la par que la iniciativa privada, fue viéndose progresivamente relegada, que comenzó a mejorar su oferta educativa atrayendo y segregando al mismo tiempo a estudiantes y eventualmente a sus familias, que pasaron a formar parte de núcleos sociales homogéneos. De tal suerte, que la educación pública comenzó a acoger principalmente a aquellos niños y jóvenes que provenían de situaciones socialmente vulnerables, mientras que las familias de mayores ingresos optaron por la educación particular o privada. Esta migración trajo insospechadas consecuencias en la estructura educacional que aunque han sido bastante estudiadas, están lejos de constituir un cuerpo sólido de evidencia sobre las consecuencias de este cambio.

No obstante, los estudios que intentan explicar las diferencias entre lo público y lo privado que condicionan el desempeño y futuro académico de los estudiantes, se han aproximado mayoritariamente desde perspectivas socioeconómicas y sociológicas. Fue entonces que emergió la idea de centrar la investigación tomando en consideración estos efectos, pero intentando referirla a aspectos intrapsíquicos. La idea, por supuesto, no es nueva, sin embargo, es menos habitual acceder a trabajos que consideren las diferencias desde variables psicológicas o cognitivas. El interés se sostiene en el supuesto de que este tipo de características pueden ser intervenidas individualmente y a menor plazo que los cambios que podrían requerirse a nivel social y económico.

Entendiendo la educación privada como aquella que es provista por actores no estatales en que la propiedad de los establecimientos es, por lo general, privada y que puede o no recibir subvención de parte de poderes públicos para efectos de otorgar el servicio educacional (Corvalán, Elacqua y Salazar, 2009), y que la educación pública es provista en Chile, no por el Estado, sino por los municipios de cada ciudad, quienes reciben una subvención por estudiante (Oliva, 2008), se puede asumir que los alumnos de escuelas privadas y públicas provienen de contextos socioculturales distintos y que es posible determinar el efecto que tienen los mayores o menores aportes económicos a las escuelas o los municipios, o las mejoras salariales a docentes y sus consecuencias sobre la marcha de los procesos educativos. No obstante, un diagnóstico que considere estos u otros factores no implica que se pondrán en marcha políticas que disminuyan la desigualdad o si se observarán cambios en un plazo definido de activarse dichas políticas.

El esfuerzo a largo plazo por la disminución de la desigualdad debe complementarse con resultados más inmediatos que incidan sobre los individuos y especialmente sobre variables cognitivas que podrían explicar el diferencial entre estudiantes de educación pública y privada. Al actuar sobre estas variables, se está más próximo a un diagnóstico que pueda llevar a modificaciones y posibles cambios de corto o mediano plazo, pues la intervención es directa y no depende en lo inmediato de factores a nivel macro o de contexto.

Un antecedente de esta mirada centrada en lo cognitivo proviene de la ley estadounidense “No Child Left Behind” en cuya labor se aplicaron modelos que intentaban dar cuenta, mediante diagnósticos precisos, de las habilidades concretas (también llamadas atributos cognitivos) que los estudiantes dominaban y que les permitían o no resolver una tarea específica y de ese modo obtener no solo el desempeño educativo general de los estudiantes, sino que también se pudiera obtener información sobre las debilidades específicas que evidenciaban frente a dicha tarea (Tatsuoka, 2009).

La tesis intenta indagar no solo factores sociales y económicos que explican la segregación educativa, sino también cuáles son las consecuencias a nivel individual en el desarrollo cognitivo, estableciendo su efecto en el desempeño académico, que es lo que en definitiva evidencia la segregación.

El desempeño académico, tiene multitud de definiciones, aunque es claro que su operacionalización, comprendida desde la calificación escolar promedio de un estudiante hasta los resultados obtenidos en pruebas estandarizadas, es la más extendida de todas (Agencia Calidad de la Educación, 2015c, 2016; Rivkin, Hanushek y Kain, 2005). Estudios internacionales como TIMSS o PISA, llevan a cabo la evaluación del aprendizaje a través de medidas estandarizadas de áreas de conocimiento que se consideran básicas o claves como lectura, matemáticas y ciencias (OECD, 2014). Dichas áreas son relevantes predictores de éxito académico y laboral y de allí el interés por medirlas (OECD, 2002).

En Chile ocurre algo similar pero aplicado a nivel local, siendo el Simce (Sistema de Medición de la Calidad de la Educación) el mecanismo por el cual se establece el desempeño académico, siempre considerando las áreas claves en el aprendizaje, en particular comprensión de lectura, matemática y ciencias naturales (Agencia de Calidad de la Educación, 2016a).

En virtud de que es comprensión de lectura una de las áreas curriculares de mayor importancia en la educación chilena (MINISTERIO DE EDUCACIÓN, 2013; Pérez, 2009) y que su desarrollo tiene un fuerte impacto en los aprendizajes de muchas otras áreas, es que resulta de interés considerar los atributos cognitivos que rigen su dominio, pues un buen desarrollo comprensivo lector incide en la comprensión de muchas otras áreas (Marchant, Lucchini y Cuadrado, 2007; OECD, 2002), sin olvidar mencionar la importancia que tiene para una mejor adaptación e inserción social (Guerra y Guevara, 2017; OECD, 2002).

Por supuesto, surge una particularidad: Los modelos cognitivos que se determinen para comprensión lectora, dependen de cuáles atributos sean puestos en juego en dicha tarea. Hay una importante y vasta literatura dedicada al tema, con diversos enfoques y énfasis (Follmer, 2017; Gómez-Veiga, Vila, García-Madruga, Contreras y Elosúa, 2013; Guzmán, Véliz y Reyes, 2017; Kendeou, van den Broek, Helder y Karlsson, 2014; Ravand, 2016; Svetina, Gorin y Tatsuoka, 2011; Tatsuoka, 2009; Tijero, 2009; Vergara, Strasser y del Río, 2016; Yi, 2017) Algunas de estas propuestas adoptan una perspectiva psicométrica que se integra en particular con la evaluación educacional, pero centrada en

“la evaluación para el aprendizaje” y no solamente en la “evaluación del aprendizaje” (Ravand y Robintzsch, 2015).

Una ventaja de los modelos “para el aprendizaje” reside en que los atributos cognitivos que son necesarios para evidenciar comprensión de lectura, no provienen de evaluaciones adicionales, como se informa en algunos estudios (Gómez-Veiga, Vila, García-Madruga, Contreras y Elosúa, 2013; Guzmán, Véliz y Reyes, 2017) sino que son establecidos directamente a partir de la tarea propuesta en el test aplicado.

Algunos de los modelos cognitivos que buscan determinar los atributos cognitivos a la base de la medida de comprensión lectora son: a) El denominado Rule Space Method (Tatsuoka, 2009), que intenta explicar el desempeño en esta área a través de la propuesta de atributos detrás del fenómeno de comprensión (Buck, Tatsuoka y Kostin, 1997; Sventina, Gorin y Tatsuoka, 2011); los modelos Deterministic-Input, Noisy-And-Gate, DINA y Deterministic-Input, Noisy-OR-Gate, DINO (Liao, Kuo y Deenang, 2015; Park y Cho, 2017); c) el modelo Generalized Deterministic-Input, Noisy-And-Gate Model, G-DINA, que es general y permite aplicar modelos subordinados para distintos ítems de un mismo test (Chen y Chen, 2016; Ravand, 2016); c) el modelo General Diagnostic Model, GDM (von Davier, 2005); el modelo Reduced Reparameterized Unified Model, o RRUM, (Jang, 2009; Jang, Dunlop, Wagner, Kim y Gu, 2013; Li, Hunter y Lei, 2015) también reconocido como Noncompesatory Reparameterized Unified Model, NC-RUM (Ravand y Robitzsch, 2015) y el modelo Attribute Hierarchy Method o AHM, que establece un procedimiento para la clasificación de las respuestas que examinados dan a los ítems de un test, de acuerdo a un conjunto de patrones de dominio de atributos que se asocian con distintos componentes de un modelo cognitivo (Wang y Gierl, 2011). Todos ellos serán tratados con algún detalle posteriormente.

Entonces, el interés fundamental de este trabajo es considerar no solo las diferencias entre escuelas públicas y privadas que pueden ser establecidas a nivel social y económico, sino también cómo dichas diferencias se reflejan a nivel cognitivo y de desempeño en los estudiantes que forman parte de una y otra, evidenciando segregación. El supuesto es que la segregación tiene efectos no solo en la composición socioeconómica, sino también en aspectos cognitivos tan relevantes como aquellos que son determinantes para enfrentarse a cualquier evaluación, pues se asume que la

comprensión de textos es una de las áreas fundamentales de adaptación y progreso escolar.

El concepto segregación tiene especial relevancia en el marco teórico de este trabajo y se entiende como “el grado en que los individuos de algún grupo social o económico se encuentran distribuidos en los distintos espacios geográficos o sociales” (Treviño, Salazar y Donoso, 2011, p. 36). Un aspecto particular de la segregación que se menciona en este trabajo -por su relevancia en los diferentes estudios citados- es el de homogeneidad, que se refiere a “la sub o sobre-representación de grupos minoritarios en distintos espacios. La distribución de los grupos minoritarios influye en la exposición de estos a grupos mayoritarios o dominantes. También los grupos más desfavorecidos pueden estar concentrados, ocupando un menor espacio per cápita que los mayoritarios” (Treviño, Salazar y Donoso, 2011, p. 36).

En consecuencia, el presente estudio de investigación implica un diagnóstico psicométrico de los patrones cognitivos de estudiantes de escuelas privadas y públicas que se suponen distintos en virtud de la segregación y los factores que la generan. Se intenta establecer las diferencias individuales entre estudiantes, sin perder de vista que estas se encuentran anidadas en uno u otro modelo educativo (privado-público) y que aunque es relevante que los sujetos estén subordinados al diseño de la estructura socioeducativa, es más productivo diagnosticar sus fortalezas y debilidades en orden a proveer un perfil diferencial de las capacidades de estudiantes del sistema privado y del sistema público con objeto de tomar medidas futuras de apoyo con base en dicha información.

Aunque en el aspecto técnico el estudio es fundamentalmente psicométrico, el trasfondo es socioeducativo y a diferencia de un enfoque meramente analítico, el trabajo pretende brindar las bases de un camino que sustente futuras intervenciones destinadas a compensar las carencias cognitivas que determinan buen o mal desempeño académico.

La estructura del presente trabajo requiere una breve introducción, pues en este se combina un fuerte componente teórico junto con una sección de estudios que responden a los objetivos del estudio.

La tesis comienza con una Introducción, que da cuenta del contexto teórico en que se inserta este trabajo y donde fundamentan las principales nociones que guían la propuesta de este estudio. Esta Introducción es eminentemente teórica y está dividida en capítulos: El capítulo 1 presenta los antecedentes relativos a la calidad de la educación y los fundamentos de la política educativa aplicada en Chile, así como una breve revisión histórica de los cambios que esta política y sus procedimientos han tenido a lo largo de los años desde el retorno a la democracia. Se pone énfasis en la desigualdad y la segregación que esta ha generado hasta llevarla al estado actual de las cosas, en el ámbito económico, social y, de modo relevante, en el cognitivo; El capítulo 2 es breve pero muy importante para la vinculación de los aspectos educativos y los modelos cognitivos que se mencionan posteriormente. En este se pone énfasis en las consecuencias de la desigualdad a nivel cognitivo y su relación con las habilidades de comprensión de lectura, vinculando su desempeño con algunos efectos críticos que esto tiene en otras disciplinas; El capítulo 3 expone argumentos en favor de la composición cognitiva de la comprensión lectora en virtud de los atributos o habilidades constituyentes; El cuarto capítulo expone el procedimiento general de establecimiento de diseño cognitivo en tests y posteriormente describe la génesis de los modelos de diagnóstico cognitivo, así como sus etapas de elaboración y las especificaciones técnicas de dicho desarrollo.

En tercer lugar, se dispone un apartado específico para la Pregunta de Investigación y otro para los Objetivos. Estos guían la presentación ulterior de las investigaciones efectuadas en función de los fines trazados para el estudio.

En cuarto lugar, se presenta el apartado Presentación de los Estudios, en los que se insertan los trabajos de investigación que se hacen cargo de los distintos objetivos y que en conjunto intentan dar cuenta del problema que guía la tesis.

El trabajo finaliza con un apartado de Discusión en las que se sintetizan los hallazgos, su relación con los fundamentos teóricos y una reflexión sobre las implicancias de estos estudios.

## INTRODUCCIÓN TEÓRICA

El propósito de todo trabajo de investigación es hacer una aportación relevante al conocimiento existente y en lo posible contribuir a la solución de problemas que deriven de la falta de este. En Chile, en el campo de la psicología de la educación, ambas pretensiones suelen ser un proceso en desarrollo. Ni el conocimiento ni las soluciones son suficientes. Las razones estructurales de por qué esto ocurre escapan a la argumentación que podría darse en este espacio, sin embargo, el presente estudio tiene este doble sentido.

En efecto, el propósito del presente trabajo busca no solo resolver una serie de cuestiones teórico-metodológicas que emergen producto de la desafiante complejidad del problema a tratar, sino también proveer una vía –no necesariamente una solución- de intervención para las apremiantes consecuencias de la ya reconocida desigualdad presente en el sistema educativo chileno.

No obstante, a diferencia de la vasta literatura que puede citarse al respecto (Cabalin, 2012; Mayol, 2011; Paz, 2015; Oliva, 2008; Valenzuela, Bellei y de los Ríos, 2013) este estudio no tiene una orientación meramente sociológica o centrada en políticas públicas. Su énfasis principal es entender la desigualdad desde las diferencias cognitivas. No pueden ser consideradas causas de la segregación, sino más bien consecuencias, pues como se expondrá a continuación, la evidencia ha establecido que los factores sociales, económicos y políticos contribuyen a conformar grupos, aventajados y desventajados (Valenzuela, Bellei y de los Ríos, 2013), que se distinguen en muchos aspectos, y entre ellos los psicológicos.

La diferencia está en que esas particularidades psicológicas en algunos casos, además de ser parte del problema, pueden también ser parte del cambio y, bajo un proceso de intervención, más asequible que la intervención de factores políticos, económicos y/o sociales.

En esta introducción se busca establecer que la desigualdad está presente y que las diferencias entre grupos con o sin ventajas, son reales y pueden ser identificadas a nivel psicológico y específicamente a nivel cognitivo.

## **CAPÍTULO 1: Calidad de la educación y la política del sistema educativo**

El éxito académico del alumnado ha acaparado una parte importante de la atención tanto de instituciones internacionales como nacionales orientadas al ámbito de la educación, como ocurre a nivel internacional con el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), el International Activities Program del National Center for Education Statistics y su estudio Trends in International Mathematics and Science Study (TIMSS) o el muy conocido Programme for International Student Assessment (PISA) de la Organisation for Economic Cooperation and Development (OECD), o en el caso nacional, con la implementación -desde los primitivos intentos de evaluación en 1982- del Sistema de Calidad de la Educación (SIMCE), que cada año mide a estudiantes de 8 a 10 años (segundo y tercer ciclo de primaria de acuerdo a su equivalente español), además de estudiantes de 12 años de edad (equivalente a secundaria en España) para determinar si han alcanzado los objetivos fundamentales y los contenidos mínimos obligatorios del currículo nacional.

El interés de las instituciones internacionales y nacionales está justificado: Mucho del desempeño futuro, ya sea logro profesional o simplemente un mejor nivel de vida se fundamenta en haber tenido las adecuadas oportunidades escolares y una educación de calidad por un período mínimo de tiempo (Comisión Económica para América Latina, 2008; Ministerio de Educación, 2003). En efecto, algunos estudios son concluyentes en que a mayor número de años de estudio mejores salarios y menor precarización del trabajo (Fuentes, 2009; Pérez, 2009; Silva-Peña, Chelme-Bustos y Salgado-Labra, 2003).

En este sentido, la calidad de la educación, se impone como un derecho para la igualdad de oportunidades (Eyzaguirre y Le Foulon, 2003). De larga data es el debate nacional respecto de la educación y la (des)igualdad de oportunidades, y de cómo afecta la dificultad de acceso y una mala calidad de educación a las posibilidades futuras del alumnado (Casassus, 2003a).

En Chile, la situación ha sido abordada repetidas veces, incluso a nivel de gobierno y las conclusiones tienden a demostrar que la desigualdad social se replica en la educación

(Donoso, 2011; Donoso y Hawes, 2002; Mayol, 2011; Ministerio de Educación, 2003; Oliva, 2008; Valenzuela, Bellei y de los Ríos, 2013), lo que afecta tanto a profesores y estudiantes (Riquelme, 2010). Este hecho cobra especial relevancia cuando se alude al concepto de equidad, entendida como la igualdad de oportunidades, pero que comprendida en sentido negativo se refiere a la discriminación generada porque quien tiene más recibe más (Casassus, 2003a; Donoso, 2011; Donoso y Hawes, 2002). Esta inequidad tiene su reflejo más evidente en la diferenciación entre educación privada y educación pública.

En efecto, hay múltiples estudios que aterrizan la desigualdad social en los establecimientos educacionales, de manera que habrá importantes diferencias sociales, económicas y académicas entre el estudiantado. De ellas, el aspecto académico es crucial, si se considera que es la llave para otros logros futuros (Aguirre, Castro y Adasme, 2009)

A inicios de los años 80, Chile impuso un modelo neoliberal a la sociedad, bajo influjo de un conjunto de nuevos economistas, denominados comúnmente “Chicago Boys” (debido a que la gran mayoría provenía de la Escuela de Economía de la Universidad de Chicago, siendo su principal mentor el catedrático Milton Friedman), quienes propusieron, entre un conjunto de medidas una mayor apertura al mercado, una menor participación del Estado y la promoción de la competencia y la búsqueda del éxito en función del logro individual, lo que por cierto, también aplicaba a grupos sociales. Este nuevo modelo se constituyó en una alternativa de desarrollo no solo en Chile, sino también en su país de origen, Estados Unidos, y llevó a los grandes cambios en las esferas política, económica y social. El cambio, afectó también al sistema educacional, pues se comenzaron a enfatizar las prácticas de mercado en el diseño, implementación y evaluación de las políticas educativas (Cabalin, 2012; Robertson, Mundy, Verger y Menashy, 2012).

La influencia de esta visión de mercado sobre buena parte del funcionamiento económico y social del país no significó, sin embargo, el fin de la educación pública, sino que más bien la relegó a un segundo plano, pues a partir de 1981, como parte de un proceso de descentralización, se traspasó la gestión de los establecimientos públicos desde el Estado a los municipios y además se fortaleció la subvención escolar por cada niño/a que asistía a clases. Se pasó de un modelo de provisión directa por parte del Estado,

a un modelo de financiación en función de la demanda; al decir de algunos, un modelo subsidiario (Donoso, 2006; UNESCO, 2011).

Este cambio abrió la puerta a una mayor participación de dineros privados en la provisión de educación, permitiendo que surja la figura de la educación particular subvencionada, que implicaba que un responsable, usualmente el dueño de un/os establecimiento/os privados, asumiera el papel de “sostenedor” ante el Estado, velando por el funcionamiento de un establecimiento educacional en las condiciones exigidas por la ley (Ministerio de Educación, 1998).

De este modo, la educación en Chile quedó estructurada en función de tres dependencias: a) los municipios, que se hicieron cargo de los establecimientos públicos que estaban originalmente a cargo del Estado b) los privados con subvención del estado vía subsidio a la demanda (que se lleva a cabo a través de un “voucher”, una cantidad fija por estudiante), y c) la educación privada tradicional que no recibe subsidios y es financiada completamente por los padres o familias de los estudiantes (Bravo, Contreras y Sanhueza, 1999; Sapelli y Vial, 2002). Cabe especificar que en la educación pública (agrupada en conjunto como “municipal”) se pueden distinguir dos tipos: la educación municipal que es sostenida por los Departamentos de Educación Municipal (DAEM) y la que está bajo la tutela de una Corporación Municipal, entidades autónomas que tienen personalidad jurídica propia y separada de la administración del municipio. Por su parte, los privados con subvención, también llamada educación particular subvencionada, está conformada por escuelas a cargo de instituciones religiosas, como los obispos católicos, escuelas sin fines de lucro, y escuelas particulares subvencionadas con fines de lucro (Sapelli y Vial, 2002).

A partir de estos cambios, en educación se impuso una lógica de oferta y demanda, pues los padres y las familias se enfrentaban ahora a la necesidad de optar por una variedad de instituciones escolares que ofrecían disciplina, protección y variados estándares académicos en función de sus capacidades de pago. Los padres y apoderados, en este contexto, llegaron a ser consumidores (Paz, 2015), pues aquellos que cuentan con mayores recursos están en condiciones de comprar mejores servicios. Se sustenta así, un modelo en que las instituciones escolares pueden satisfacer necesidades y expectativas específicas de sus usuarios (Paz, 2015; Robertson, Mundy, Verger y Menashy, 2012).

Esta aparente diversidad en algunas ocasiones ha sido valorada como positiva, pues dado que la subvención se otorga en función de la asistencia de los estudiantes, parece ser que más tiempo en aula es igual a más aprendizaje y por ende mayor calidad de educación (Donoso, 2006; Sapelli y Vial, 2002). Sin embargo, las visiones opuestas no tardaron en aparecer, pues se ha criticado duramente que no se haya logrado relacionar la entrega de subsidios con la calidad de educación (entendida como rendimiento académico) (Donoso, 2011; Donoso y Hawes, 2002; Bravo, Contreras y Sanhueza, 1999; Mizala, Romaguera y Urquiola, 2007; Puga, 2011).

La situación ha llevado a que muchas de las investigaciones en educación, intenten determinar si el subsidio que brinda el Estado a las instituciones educacionales genera mayor calidad de educación y, por otro lado, si la dependencia administrativa y económica de las escuelas evidencia diferencias económicas, sociales y académicas entre estudiantes, buscando en algunos casos la atribución de causalidad (Paz, 2016; Silva-Peña, Chelme-Bustos y Salgado-Labra, 2003; Mizala, Romaguera y Urquiola, 2007; Puga, 2011). A pesar de los cambios que se han introducido con el fin de mejorar la eficiencia y eficacia del sistema educacional, como la Ley Orgánica Constitucional de 1990 que posteriormente fue derogada en favor de la Ley General de Educación promulgada en 2009 (UNESCO, 2011), la percepción de desigualdad no ha mejorado sustancialmente (Paz, 2016; Silva-Peña, Chelme-Bustos y Salgado-Labra, 2003; Mizala, Romaguera y Urquiola, 2007; Puga, 2011; Valenzuela, Bellei y de los Ríos, 2013).

### **Antecedentes económicos de desigualdad**

Un de las aristas más comúnmente estudiadas en cuanto a las diferencias entre dependencias, tiene relación con el aporte del subsidio a la demanda, en que se busca determinar si los aportes (usualmente llamados “vouchers”, que implica una contribución económica fija por estudiante) que el Estado hace a las escuelas representa una política que vaya en beneficio de la calidad en educación o bien otros criterios (Cox, 2012; Sapelli y Vial, 2002).

La discusión ha estado centrada en buena medida en los recursos otorgados y si estos han tenido algún impacto en promover mejoras. No obstante, ya en 1998, se señalaba que cundía la decepción producto de que los fondos aportados realmente no estaban generando los resultados esperados. La situación tendió a ser más evidente

cuando a partir de la descentralización de 1980, se permitió la participación de establecimientos privados que podían ofertar sus servicios a los estudiantes que hasta ese momento habían sido parte en exclusiva de la educación pública provista por el Estado. Esta posibilidad abrió las puertas a que las familias pudiesen optar por diferentes establecimientos en función de los servicios que estos ostentaban (proceso al que se le denomina “choice”). Teóricamente se supone que la política de “choice” daría lugar a una competencia que irremediablemente mejoraría los servicios de las escuelas y a la vez, las malas instituciones verían mermada su matrícula pues los padres solo optarían por aquellas que ofrecieran una educación de mejor calidad (Mizala y Romaguera, 1998).

La entrega de “vouchers” es la materialización de esta política de “choice” que supuestamente dinamiza el mercado educativo. Nótese que para que esto funcione, a la mencionada política se le debe adicionar algún criterio objetivo de medición de dicha calidad, propósito que encarnó en la prueba aplicada por el Sistema de Medición de la Calidad de la Educación (SIMCE), que permite juzgar, mediante un ranking de escuelas creado especialmente a partir de los resultados que obtienen los estudiantes en esta prueba, la calidad de la educación impartida por una institución en particular (Eyzaguirre y Fontaine, 1999; Mizala y Romaguera, 1998).

Aunque la política de elección parece ser de sentido común (se premia a los mejores y se castiga a los peores), hay algunos factores que llevaron a varias situaciones negativas que cargaron la balanza hacia las escuelas particulares subvencionadas, en desmedro de las llamadas municipales. En primer lugar, en 1991 cambian las condiciones de contratación de los docentes; en las instituciones subvencionadas depende de la dirección del establecimiento que muchas veces puede ofrecer mejores condiciones salariales para sus profesores, mientras que la contratación en las escuelas públicas puede depender, en última instancia, del municipio. En segundo lugar, en 1993 se abrió la posibilidad a lo que se llamó financiamiento compartido, es decir, que los apoderados pueden realizar un aporte por sus pupilos además de la subvención recibida, lo que fue más favorable para las escuelas particulares subvencionadas. Tercero, los establecimientos privados tienen completa libertad para seleccionar a sus estudiantes, mientras que los establecimientos públicos deben aceptar a todos los que deseen inscribirse en ellos (Mizala y Romaguera, 1998).

Córdova, Rojas y Azócar (2016) en un reciente y particular estudio refrendan algunos de los factores que promueven la segregación, destacando que es un hecho consolidado que el sistema educacional chileno refleja amplias diferencias entre la educación privada y la pública. Pero además, aporta el interesante hecho que las familias también tienen un papel en la segregación observada, pues aquellas familias con mayores ingresos optan por mejores escuelas aun cuando estas estén más lejos de su zona geográfica, separando incluso físicamente a los estudiantes de las comunidades en las cuales residen. Para estos autores, el hecho que los padres puedan pagar por sus hijos/as en las escuelas con financiación compartida, agrava la segregación económica, pues las escuelas, además de poder solicitar un pago adicional al “voucher” que puede llegar a ser 1.6 veces mayor que este último (lo que de por sí discrimina contra los que son incapaces de pagarlo), pueden llegar a realizar evaluaciones destinadas a seleccionar a los mejores estudiantes en función de su proyecto educativo (Bellei, 2007; Donoso, 2012; Elacqua, 2012).

Para Elacqua (2012) queda claro que los estudiantes más vulnerables se distribuyen de manera dispar entre los distintos tipos de escuelas. Según el autor hay que considerar el hecho que hay escuelas particulares con subsidio y sin fines de lucro, mientras que otras pueden declarar fines de lucro, llevando a que mientras las escuelas públicas y las particulares con subsidio sin fines de lucro, están más abiertas y acogen la diversidad socioeconómica entre sus estudiantes, las escuelas particulares con subsidio y con fines de lucro tienden a estratificarlos en función de su capacidad de pago.

Para Donoso (2012) bajo las reglas del lucro, fin sostenido por la educación privada y la subsidiada con fines de lucro, la educación no es propiamente un bien público, sino un bien de consumo, en que el Estado asegura estándares mínimos para la educación primaria y secundaria, aspecto al que queda relegada la educación pública. Quien aspire a mejor educación, debe asumir los costos que ello implica, aunque con la posibilidad de acceder a apoyos financieros de privados que, por supuesto, representan también una opción de rentabilización.

Quizás lo más duro de esta situación descrita, es que a pesar de los años el sistema educativo chileno sigue generando, al decir de algunos, más segregación. El sistema en sí reproduce la diferenciación y la exclusión (Bellei, 2007; Cabalin, 2012)

No todas las visiones han sido tan drásticas como las expuestas. McEwan y Carnoy (2000) evaluaron la efectividad de los establecimientos particulares con subsidio, diferenciando entre los religiosos y los no religiosos y los establecimientos públicos o municipales. La conclusión fue que los públicos eran más efectivos en el logro matemático y lenguaje, siempre y cuando los particulares con subsidio no fuesen religiosos (católicos) y fuesen de fuera de la capital. Sin embargo, en términos de uso de recursos, los particulares con subsidio logran los mismos resultados que los públicos pero con un uso más eficiente de los recursos. Una corroboración de estos resultados se obtuvieron en el estudio de Sapelli y Vial (2002) en que se señala que cuando a los establecimientos municipales se les asignaban recursos cuantiosos, hasta un 200% más que un establecimiento particular subvencionado, el rendimiento en Lenguaje era mejor en los primeros. Por el contrario, si las ayudas económicas para la educación municipal representaban menos del 27% de los recursos requeridos, son los establecimientos particulares con subsidio los que surgen como efectivos. Por su parte, Tokman (2002) señaló que las escuelas públicas en Chile no son intrínsecamente malas o buenas, sino que son “funcionales” a la población desfavorecida que atienden (lo que no deja de sonar a eufemismo).

A pesar de las visiones positivas, un determinante estudio que parece fortalecer la visión más negativa de la segregación proviene de la OECD (2012) que acudiendo a los resultados de la prueba PISA, muestra que los padres con mayores ventajas socioeconómicas tienden a enviar a sus niños a escuelas privadas. Aunque se dan varias razones de por qué esto puede ocurrir, como el manejo de mayor información, el deseo de que sus niños compartan con estudiantes de similares características o la búsqueda de mejores ambientes de aprendizaje, entre otros aspectos, lo cierto es que las escuelas privadas tienen mayor autonomía, mejores recursos y claramente se desempeñan mejor en PISA que las escuelas públicas.

En la suma final y a pesar de la falta de consenso en los efectos de este modelo financiación a la educación (Correa, Inostroza, Parro, Reyes y Ugarte, 2013), se puede declarar que la situación de la educación pública está en desmedro respecto de la privada. Varios estudios revelan que este modelo basado en subsidio fomentó la segregación de las familias por ingreso y se vio afectado por el apoyo financiero que los municipios podían dar a sus escuelas, sin importar si esta es más o menos vulnerable por contexto o

zona geográfica y requiere más recursos que otras (Paz, 2015). En la práctica esto significó reproducir un modelo segregado socialmente dejando las escuelas municipales para el grupo económico más vulnerable, las escuelas particulares subvencionadas para el grupo económico considerado clase media y media-baja, y las escuelas privadas para el grupo de mayores ingresos, pudiendo abarcar la clase media de mayores recursos (Brandt, 2010; Campos, 2010; Paz, 2015; Puga, 2011).

Los efectos de esta segregación económica no son triviales. Se sabe que las escuelas privadas sin subsidio, que acogen a estudiantes de mayor poder económico tienen mejores resultados académicos y son más equitativos en términos de la distribución social del logro académico, comparadas con las privadas con subsidio y las públicas (Mizala, Romaguera y Ostoic, 2004).

### **Antecedentes sociales de desigualdad**

Sin poder hacer una separación radical de los aspectos económicos ya señalados, se pueden destacar varios factores que cobran particular importancia en el contexto nacional. Campos (2010) inicia su presentación señalando que abordar la desigualdad en Chile no se puede realizar sin reconocer que ha sido la sociedad chilena en su conjunto la que ha fracasado en el logro de la equidad y que ello no es sino un reflejo actualizado de algo que Bordieu y Passeron señalaron en 1964: Que la institucionalidad escolar era solo un mecanismo de reproducción y profundización de las diferencias sociales entre los grupos sociales y los individuos que los componen.

En el caso chileno, parece ser que la perpetuación de las diferencias educacionales tiende a estar más presentes que nunca a pesar de los intentos por generar condiciones de competitividad que motiven tanto a las instituciones educativas tanto privadas como públicas a buscar caminos de mejora desde sus respectivos ámbitos de intervención. El problema sin embargo, sigue residiendo en la composición estudiantil de las escuelas que Mizala, Romaguera y Ostoic (2004) identificaron en su momento. La composición no es aleatoria, y por tanto la tan reconocida provisión de subsidios a la educación privada por sobre la pública no puede representar su efecto positivo en la calidad de la educación. Si las escuelas privadas atraen a estudiantes aventajados, que además pueden ser previamente seleccionados (OECD, 2012) y si las escuelas públicas, por su parte, reciben aportes económicos que no necesariamente están relacionados con sus necesidades (en

particular si están en contextos vulnerables) y además no pueden hacer selección de sus alumnos, entonces hay múltiples variables enmascaradas que podrían actuar como explicaciones alternativas a la calidad de educación en estas instituciones (Bravo, Contreras y Sanhueza, 1999).

Estos factores emergen en virtud de sus consecuencias. Del mismo modo que se discuten los factores económicos y su influencia en la calidad de la educación, estas otras variables explicativas pueden incidir desde el ámbito social e inclusive del psicológico. Son variables que han sido tempranamente reconocidas en variados estudios (Mizala y Romaguera, 1998; Schiefelbein, 1992) y que en algún grado se expresan a partir de la influencia de los factores económicos expuestos.

La inquietud por estos otros factores en buena medida proviene de los resultados que se aprecian al comparar las escuelas privadas sin subsidio, las que tienen subsidio y las públicas, pues las diferencias no se explicarían únicamente por la entrega de recursos económicos, panorama que sería el ideal para quienes defienden el sistema de “vouchers” y la política del “choice” (Mizala y Romaguera, 1998; Sapelli y Vial, 2002), sino que en la práctica se ven influenciados por un conjunto de atributos de carácter social propios de las escuelas, las familias y los estudiantes.

Si se tiene en consideración que los estudiantes no se distribuyen de manera aleatoria en los distintos tipos de dependencia (Bravo, Contreras y Sanhueza, 1999), y en consecuencia es posible vincular la calidad de la educación con otras variables, es de importancia conocer cuáles son las que parecen jugar un papel determinante en las diferencias.

La composición de las escuelas públicas llevó a Jadue (1997) a plantear que los niños y niñas que llegan al sistema escolar y que tienen una baja condición socioeconómica y cultural, tendrán un mayor grado de fracaso escolar. El informe que toma como base los resultados de PISA 2006 (Ministerio de Educación, 2007) señala que las escuelas chilenas se caracterizan por reunir a estudiantes con atributos similares en los aspectos socioeconómicos y culturales. Los factores asociados son múltiples, contando entre ellos aspectos sociales, como la falta de condiciones apropiadas de estudio en el hogar (hacinamiento, falta de privacidad, falta de espacios para el estudio, entre otros), falta de preocupación y comunicación con los padres o el bajo nivel educacional de estos

que a la larga afecta los aspectos cognitivos, como carencia de estimulación temprana y aprendizajes preescolares básicos, o dificultades de atención y concentración.

El estudio de Mizala y Romaguera (1998) que reconocía la importancia de estas otras variables, consideró su influencia al controlar sus efectos y determinar que, por ejemplo, entre escuelas con subsidio y escuelas públicas no había diferencias estadísticamente significativas cuando se trataba del rendimiento académico (al nivel de 8 a 10 años de edad o segundo y tercer ciclo primaria en su equivalente español). Por supuesto, al efectuar el control se mantenía libre de efectos de confusión la relación de entrega de recursos y desempeño de los estudiantes en la escuela, lo que revelaba que el sistema de “voucher” no era una contribución a la calidad. No obstante, tan importante como asegurar esta relación explicativa, es el apreciar cuáles son aquellas variables que generaban confusión y podían brindar explicaciones alternativas a las diferencias de rendimiento, pues dichas diferencias son las que resultan de relevancia cuando se hace un juicio sobre la calidad de la educación pública versus la privada.

Estas variables de carácter social incluyeron el índice de vulnerabilidad de las familias (corresponde, grosso modo, al porcentaje de estudiantes en condición de vulnerabilidad que una escuela acoge), características de los establecimientos (ubicación geográfica, número de docentes, la experiencia docente, la presencia de enseñanza preescolar y la presencia de enseñanza mixta), condición de ruralidad (se ha observado que escuelas rurales presentan rendimientos menores a los urbanos) (Mizala y Romaguera, 1998). Los autores también rescatan como variable adicional el tamaño de las escuelas, en función del número de profesores que laboran en ellas. Escuelas pequeñas denotan mejores desempeños. Otra variable corresponde al grado de homogeneidad, pues esto puede influir en los resultados que las escuelas obtienen. Las escuelas privadas sin subsidio son las más homogéneas en la composición de sus alumnos, la siguen los privados con subsidio y por último los públicos.

Bravo, Contreras y Sanhueza (1999) citan algunos estudios en donde, con el propósito de ajustar modelos econométricos, se controlan otras variables de confusión propias del entorno de la escuela: El cociente docente/estudiante, el tamaño de la clase, la educación promedio de la madre, el ingreso mensual de la familia, las tasas de reprobación y aprobación de los estudiantes. Controlando estas variables, los resultados

que se obtienen son poco relevantes, aunque los privados subvencionados obtienen un desempeño levemente mejor que los públicos.

Pérez (2009) destaca varias características de interés que tienen relación con el desempeño académico. En primer lugar señala que hay diferencias notorias en la condición urbano-rural de los establecimientos, incluso al considerar la dependencia. Así, escuelas públicas rurales evidenciarían menores desempeños que las públicas urbanas. En segundo lugar, indica que las diferencias no se observan entre géneros y que las que se presentan van desapareciendo con los años de escolaridad. Para este autor la diferenciación cualitativa de las escuelas se ha ido acentuando con los años, lo que debe ser interpretado como mayores logros académicos para centros privados y menores para los centros públicos, llevando a una “diferenciación horizontal” (Pérez, 2009, p. 10) que da cuenta de una segmentación sectorial, en donde la expansión de la educación privada con subsidio ha crecido principalmente en zonas pobladas de clase media, capturando a la población residente, quienes no tienen los recursos para pagar educación privada sin subsidio, pero que pueden al menos alejarse de la educación municipal, que intuyen de menor calidad. Y aunque no se puede asumir que esto es una conducta deliberada, es explicable pues basta asumir que las familias desean la mejor educación disponible para sus hijos (Atria, 2009; Treviño, Salazar y Donoso, 2011).

Esta segmentación sectorial lleva a una composición social claramente reconocible entre escuelas privadas, privadas con subsidio y públicas que acentúa las diferencias que de por sí existen en función de la calidad de servicio que estas pueden prestar. Algunos han llamado a este fenómeno un verdadero “apartheid educativo” (Campos, 2010). En la escuela pública, de cada 100 alumnos que asisten, 86 forman parte de los tres primeros quintiles de la población, es decir, la población más vulnerable.

Aunque parezca de menor importancia, lo cierto es que esta segregación sectorial tiene sus bases en la organización territorial de la población, pues los estudiantes de condición más vulnerable usualmente asisten a las escuelas cercanas a sus residencias, lo que determina la vinculación social que estos pueden tener y las redes con las que pueden asociarse. Se comprenderá que esta situación representa un mecanismo que facilita la homogeneización de la población que compone las escuelas (Matus y Quezada, 2007). El posicionamiento territorial de las escuelas tiene como consecuencia la prestación de

servicios a quienes pueden pagarlos. No es coincidencia que las escuelas privadas se posicionen territorialmente en zonas acomodadas y las municipales usualmente queden relegadas a las zonas más vulnerables, o bien, se ubiquen en espacios céntricos, recibiendo al colectivo más vulnerable de dichos espacios territoriales. Las distintas dependencias de escuela concentran población con características similares y que pertenecen a grupos sociales homogéneos, con escasa integración entre sí (Borsdorf, Hidalgo y Vidal-Koppmann, 2016; Campos, 2010).

La falta de integración, o desvinculación social que se aprecia, es la fuente de los procesos de estigmatización, exclusión y discriminación que ataca a los sectores más vulnerables, asociando la escuela pública y la población a la que atienden, con problemáticas sociales (Campos, 2010).

Hay otro efecto asociado con la segregación social, y es el llamado “efecto pares” que implica que si los estudiantes forman parte de un contexto homogéneo, poco diverso, tienen menos probabilidades de lograr aprendizajes, pues los compañeros de curso presentarán fortalezas y debilidades parecidas. Hay un efecto social relacionado, y es que las prácticas de los profesores están de acuerdo a la interpretación que hacen de la composición social y las capacidades que esperan de sus estudiantes (Campos, 2010; Ortíz, 2015; Valenzuela, Bellei y de los Ríos, 2008).

En el contexto escolar, los estudiantes de escuelas públicas se ven enfrentados con contexto que no ayuda a su condición social. De acuerdo a (Ortúzar, Flores, Milesi y Cox, 2009) los profesores con menor formación académica suelen enseñar en colegios con alumnos de nivel socioeconómico bajo (menor porcentaje con mención o especialidad), por lo que profesores de formación precaria terminan en escuelas de mayor pobreza con estudiantes de mayor vulnerabilidad. A ello se suma, según estos autores, el que además los profesores con mala formación pertenecen en su mayoría a niveles socioeconómicos bajos.

Dicho de otro modo, se trata de un fenómeno de segregación que no afecta solo a los estudiantes y sus familias, sino incluso a los profesores que componen la planta profesional de la escuela y que contribuye a configurar núcleos sociales desvinculados y con características distintivas. Los estudiantes más pobres reciben educación de peor

calidad, y por tanto su aprendizaje es menor y sus capacidades son insuficientemente desarrolladas (Campos, 2010).

Las condiciones de pobreza quedan entonces retratadas y reflejadas también en el sistema educativo. Las diferencias que obtienen las escuelas en sus resultados académicos están determinadas por el origen social de los estudiantes (Cisterna, 2007).

La segregación de las escuelas es la segregación de los estudiantes y las reformas impulsadas no han cambiado, en lo sustantivo, este panorama. Treviño, Salazar y Donoso (2011) aportan una precisión interesante: Existen también diferencias entre las escuelas privadas con subsidio que tienen fines de lucro, o bien una orientación religiosa o bien cobro de financiación compartida (el monto pagado por la familia además del subsidio entregado por el Estado). Las escuelas públicas tienen menores niveles de estratificación socioeconómica, mientras que en los privados mencionados se puede apreciar un mayor nivel de segregación de sus estudiantes vulnerables. Si se segmenta por tipo de escuela, aquellas públicas no religiosas y sin fines de lucro presentan niveles de homogeneidad similares a las escuelas públicas.

Un estudio previo, de Elacqua (2007) resulta de interés pues pone cifras a la situación descrita por Treviño y coautores. De acuerdo a dicho autor, internamente la segregación escolar es mayor en las escuelas privadas con subsidio. Se establece que estas atienden un 7% menos de estudiantes vulnerables en comparación con escuelas públicas, también se reafirman resultados previos, que señalan que las escuelas privadas católicas con subsidio atienden un 12% menos de estudiantes vulnerables que las públicas. Por su parte, son las escuelas privadas con subsidio y que además solicitan un pago adicional a las familias (financiación compartida), las que denotan más estratificación y segregación interna en función de la capacidad de pago de los padres. Por último, las escuelas rurales serían más propensas a recibir estudiantes vulnerables que los otros tipos de escuelas.

Ortiz (2015) agrega una interpretación de interés a la falta de diversidad que involucra a la escuela en las que se tiende a homogeneizar a los estudiantes por nivel socioeconómico; No solo se segregan grupos sociales y las características de los estudiantes, sino que la falta de oportunidades para compartir con alumnos distintos, lo que impediría la identificación de aspectos de convivencia social en común. Es decir, se pierde la experiencia formativa de inclusión.

En este sentido, los efectos composicionales de la escuela, entendidos como efectos asociados a la agregación de características individuales de un grupo de estudiantes sobre su desempeño académico, habiendo descontado el efecto de dichas características consideradas a nivel individual, más “los efectos de pares”, ya discutidos, son factores que en conjunto pueden tener influencia sobre el rendimiento (Wilkinson, 2002).

Queda claro, a partir de la exposición realizada en esta introducción, que aunque las visiones son variadas hay cierto consenso en que, en efecto, en el sistema escolar chileno se está frente a una situación de fuerte segregación que condiciona la calidad de la educación y que los desempeños de distintas dependencias reflejan diferencias de importancia.

Calidad de la educación es algo que no se suele definir en los trabajos que han sido citados, pero que este caso tiene una definición operacional de facto, pues ha sido usada en muchos de los estudios ya mencionados. Se corresponde a la evaluación en el logro de estándares de aprendizaje de los estudiantes en los diferentes niveles de enseñanza (Comisión para el Desarrollo y Uso del Sistema de Medición de la Calidad de la Educación, 2003). El desempeño de este se inserta en el contexto nacional, pues este logro se establece a través de los resultados de una prueba estandarizada denominada genéricamente “Simce”<sup>1</sup> y se aplica anualmente a ciertos niveles, en particular 4° de Primaria, 8° Básico (2° ESO) y 2° Medio (4° ESO). En algunas ocasiones, también se evalúa a estudiantes de 2° de Primaria y 3° Medio (1° Bachillerato) (Información que puede ser consultada en <http://www.agenciaeducacion.cl/simce/>).

En consecuencia, los factores económicos y sociales caracterizan diferencialmente los tipos de escuela y acumulan evidencia como posibles variables explicativas del desempeño académico. En Chile, la mejor aproximación al desempeño reside en SIMCE y es por eso que se le utiliza profusamente como variable criterio.

---

<sup>1</sup> La prueba, llamada por su abreviatura, prueba SIMCE (Sistema de Medición de la Calidad en Educación) permite obtener información sobre el logro de los Estándares de Aprendizaje por los estudiantes en los diferentes niveles de enseñanza. La Agencia de Calidad de la Educación, a cargo de SIMCE, también administra las evaluaciones estandarizadas internacionales como PISA o TIMMS (Agencia de Calidad de la Educación, 2011, 2014).

## **Dimensión cognitiva de la desigualdad**

Las diferencias no son solo económicas y sociales, también hay atributos psicológicos, en particular cognitivos, asociados con la vulnerabilidad y la segregación. Las consecuencias de esta situación, expuesta en el apartado previo, son especialmente visibles cuando se considera como criterio la variable aprendizaje.

Si los estudiantes más vulnerables deben asistir a establecimientos de educación pública (o bien aquellos privados con subsidio sin fines de lucro, que acogen población vulnerable cuya composición social y económica es mayormente homogénea), y los privados acogen a estudiantes con mayores recursos, es esperable que se tenga un grupo favorecido y otro no. Las consecuencias de estas diferencias van más allá de lo meramente socioeconómico.

En efecto, en el informe PISA dedicado a estudiantes con bajos desempeños (OECD, 2016), se sostiene que este no está asociado a características meramente individuales o a las peculiaridades de una escuela. Más bien, se asocia a una combinación de factores: Experiencias familiares, la escuela y el sistema educacional, que en conjunto pueden restringir las oportunidades de aprendizaje y llevar a desempeños insatisfactorios.

Estudiantes cuyos padres tienen altos niveles de educación y tienen trabajos mejor pagados, tienen acceso a mejores oportunidades educativas y por tanto tienen más opciones de obtener los recursos que contribuyan a su desarrollo cognitivo. Familias con bajos ingresos, afectadas por desempleo crónico o malos empleos, tienden a evidenciar menores oportunidades y peor acceso a educación, y en consecuencia, menos opciones de desarrollo (OECD, 2016; Pérez, Betancort y Cabrera, 2011).

No son los únicos determinantes. Un factor sociodemográfico como el género también tiene un papel en las diferencias en desempeño, pero depende del área evaluada. Menores desempeños se esperan para mujeres si el área evaluada es matemáticas, pero la situación es inversa si se considera la comprensión de lectura y las ciencias (OECD, 2016). Otro factor sociodemográfico de relevancia se asigna a población inmigrante, sin embargo, este fenómeno es reciente en Chile y no se observan aun grandes diferencias en desempeño.

Es interesante resaltar que, de acuerdo a los resultados OECD (OECD, 2016) las diferencias en bajo desempeño son explicadas en un 15% por los factores socioeconómicos, demográficos y de base educativa entre los países miembros.

La base educativa familiar encuentra apoyo en el contexto nacional chileno, pues el desempeño de los estudiantes ciertamente también está relacionado con el entorno educativo del que provengan. Aspectos tan concretos como el número de libros en el hogar o el nivel de estudio del padre o de la madre representan factores de relevancia (Mella, 2006; Pérez, Betancort y Cabrera, 2011).

Estas diferencias encuentran consistente respaldo en varios estudios tanto nacionales como extranjeros, que resaltan que un bajo desarrollo cognitivo tiene explicación, al menos en parte, en las condiciones de origen de los estudiantes. Jadue (1997) ya advertía entonces que un bajo nivel socioeconómico es un buen predictor de bajo rendimiento escolar. El bajo nivel educativo de los padres, pobreza y dificultades escolares son elementos relacionados.

En los hogares con bajo nivel socioeconómico, no hay ambientes propicios para que los estudiantes puedan desarrollarse social y psicobiológicamente. Son ambientes que limitan la experiencia cognitiva de los jóvenes. Ello puede explicar en parte el por qué los estudiantes de establecimientos públicos (con menos recursos) presentan claras desventajas cognitivas frente a sus pares de clase media. Las experiencias que involucran aprendizaje vienen, de hecho, mucho antes de la experiencia escolar y son por tanto, claves para la adquisición de habilidades como la lectura y la escritura (Jadue, 1997). Más aún, el posible hacinamiento (que es propio de hogares pobres), la falta de materiales para el estudio en el hogar o la carencia de formación de los padres, pueden representar un mal espacio de trabajo, afectando la concentración, la capacidad de retención y la discriminación de estímulos auditivos y visuales (Agencia de Calidad de la Educación, 2013b; Jadue, 1997; Ruíz de Miguel, 2001). Asimismo, las familias con bajo nivel socioeconómico suelen utilizar estrategias de enseñanza negligentes o poco adecuadas para enseñar a sus hijos, aun cuando valoren la educación y les parezca importante que sus hijos se eduquen. Además, en estas familias puede haber poca interacción entre los padres con los hijos, de modo que tienen pocas oportunidades de fomentar el estudio en sus hogares o responder a las demandas académicas que vienen de la escuela. En estas

familias tampoco se acostumbra a leer a los niños, lo que desde una mirada cognitiva, restringe sus opciones de aprendizaje (Jadue, 1997).

Este influjo de la familia es relevante pues también juega un papel al momento de mejorar el desempeño, incluso cuando las condiciones de pobreza se mantengan, pues la familia proporciona acciones formadoras, que estimulan y favorecen el compromiso con el rendimiento académico (Gil, 2009; Ramírez, Devia, Salazar y Ramón, 2011).

## **CAPÍTULO 2: Comprensión lectora y la (des)igualdad**

Aun cuando se puede pensar en términos abstractos sobre el aprendizaje y su aparentemente ingenua relación solo con mermas en acceso a la información y en los niveles de conocimiento e inserción cultural, lo cierto es que algunos aprendizajes son determinantes en el desarrollo futuro del estudiante. En efecto, en los apartados anteriores se ha mencionado varias veces que las evaluaciones que llevan a las mayores diferencias se hacen en función de aprendizajes claves. Tanto para la OECD (2012, 2014, 2016), TIMMS (Mullis, Martin, Foy y Hooper, 2016), o las mediciones a nivel nacional representadas por el Simce (Agencia Calidad de Educación, 2012, 2016), el foco está puesto en estas áreas clave, pues su desarrollo está vinculado al futuro progreso académico y técnico-profesional de los estudiantes (Agencia de Calidad de la Educación, 2015a).

Matemáticas y Lenguaje son áreas clave fundamentales, y han sido incluidas en todas las evaluaciones internacionales mencionadas. La razón ya ha sido expresada desde hace tiempo por la UNICEF: “Entre las claves de los buenos resultados que tienen raíz en las prácticas pedagógicas en el aula, destacan clases motivadoras y cercanas a la vida cotidiana de los alumnos, con propósitos claros, estructura, ritmo, alto aprovechamiento del tiempo, exigentes, con predominio de refuerzos positivos, evaluación y retroalimentación regular del aprendizaje de los alumnos, dan prioridad a la comprensión lectora, expresión de ideas, razonamiento lógico, autonomía y creatividad de los niños” (Fondo de Naciones Unidas para la Infancia, 2004, p. 14). En la cita se menciona explícitamente dos factores esencialmente cognitivos: la comprensión lectora y el razonamiento lógico, usualmente encarnado en la disciplina matemática. De ellas sin embargo, la comprensión lectora ha sido tema recurrente de investigación por ser una de las variables a la base del logro académico en muchas disciplinas: “En los alumnos, la aplicación de la evaluación estandarizada de comprensión lectora en enseñanza básica permite medir los avances y vacíos de un área que impacta fuertemente los logros alcanzables en las asignaturas restantes.” (Fondo de Naciones Unidas para la Infancia, 2004, p. 297). La comprensión lectora es una de las áreas que puede representar un cambio en el desarrollo futuro y una de las habilidades más importantes que un estudiante

adquiere. Es la base de su aprendizaje posterior y se encuentra presente en todos los contenidos y/o disciplinas que este enfrente (Agencia de Calidad de la Educación, 2015a).

En efecto, la comprensión lectora tiene un sitio destacado en la explicación del logro académico como siguen demostrándolo numerosas investigaciones, particularmente en niños de educación primaria y secundaria (Brunning, Schraw, Norby y Ronning, 2004; Cunningham y Stanovich, 2007; Fuentes, 2009; Gaviria, Biencinto y Navarro, 2009; Gil, 2009; Kendeou, van den Broek, Helder y Karlsson, 2014; Marchant, Lucchini y Cuadrado, 2007; Núñez y Donoso, 2000; Ramos, 2006; Ugarriza, 2006).

Se ha mencionado que los logros en competencia lectora representan un requisito de participación exitosa en la mayoría de las áreas de vida de una persona, pues comprender lo que se lee permite no solo mejorar los aprendizajes, sino también el desarrollo del pensamiento crítico, pues mientras más se domina la lectura, más se comprende y se conoce el mundo y la sociedad en que se vive (Agencia de Calidad de la Educación, 2015a).

Señalar que la comprensión lectora pudiera ser un factor de desigualdad a futuro quizás sea extender demasiado su influencia, sin embargo, este relevante papel proviene del hecho de sobra conocido por el cual el éxito o fracaso del aprendizaje depende del ambiente en que el niño se desenvuelve, visión conocida como teoría sociocultural (McIntyre, 2011).

En particular depende de las interacciones que establece el estudiante con otros, en el contexto de su base histórica y cultural. Es el ambiente el que interactúa con el niño e influye en el desarrollo de su aprendizaje. Dicho ambiente está conformado por muchos agentes, pero de ellos es la escuela y la familia las que ocupan roles principales. Es necesario, no obstante, no sobreestimar ni subestimar la influencia de estas, pues si bien el contexto incide en la transformación del alumno, el alumno también incide activamente en dicho contexto. En consecuencia, no se puede pensar en una relación contexto-estudiante meramente unidireccional, hay que tomar en cuenta que el éxito o fallo en el logro de la comprensión de lectura, pasa por implicar la historia y la cultura del niño, así como las interacciones instruccionales de este con su escuela (McIntyre, 2011). Esta perspectiva descansa en la noción “vigotskyana” acerca de la zona de desarrollo próximo (ZDP), entendida como la distancia entre el actual nivel de desarrollo determinado por la

resolución de problemas de modo independiente y un nivel de desarrollo potencial determinado por la resolución de problemas bajo la guía de un adulto o el acompañamiento de pares. La ZDP asume que lo que un niño puede hacer con la ayuda de otros puede ser un mejor indicador de su desarrollo mental, que lo que pueden hacer solos. Un proceso de operación que se inicia de modo interpersonal, se vuelve posteriormente un proceso intrapersonal (Vygotsky, 1978). En la práctica, la tarea de aprender por medio de asistencia es algo que ocurre en todas las culturas, y aunque es más evidente a nivel familiar, es menos notorio, pero no menos importante, a nivel escolar.

En esta relación multidireccional contexto-estudiante, la comprensión en lectura pasa por la construcción sistemática de un vocabulario vinculado indisolublemente con el conocimiento de mundo que a su vez el niño debe ir desarrollando. De nada sirve conocer una palabra si no conozco, además, las diferentes connotaciones y contextos en los que se puede usar dicha palabra. Un amplio vocabulario y un conocimiento profundo del entorno van de la mano (Hirsch, 2003). El lenguaje, en este sentido, no es una esfera aislada de la experiencia humana es, de hecho, la mejor herramienta para tratar con ella. Un mayor conocimiento de mundo, una mayor inmersión en experiencias también posibilita comprender mejor el significado de las palabras y su aplicación. Aunque hay algún disenso sobre el momento en que esta influencia inicia, no está en duda la relevancia que tiene la inmersión como apoyo pre-lectura (Luo, Pace, Masek, Hirsh-Pasek y Golinkoff, 2016; Sugate, Schaughency y Reese, 2013).

En contextos más pobres, donde el conocimiento de mundo tiende a estar menos enriquecido y las oportunidades de vincularse con pares es menor pues la población tiende a ser más homogénea, las opciones de empaparse de experiencia de mundo son más limitadas.

Entonces, comprendiendo que el entorno social tiene un papel de relevancia en cómo se desarrolla el aprendizaje, y en particular la lectura, es posible explicar, si bien solo sea preliminarmente, las variables de la familia y la escuela que determinan este aprendizaje.

## **La influencia de la familia en el desarrollo comprensivo lector**

La mayoría de los estudios sobre familia, concluyen que su influencia está fuertemente relacionada con el nivel socioeconómico de la misma, las prácticas culturales, el acceso a educación y recursos educativos, tienden a ser buenos predictores de logro (McIntyre, 2011). En efecto, hay una relación positiva entre familias con bajo nivel socioeconómico y el número de horas que los padres invierten leyendo, o bien con el número de libros disponibles en casa. Y aunque estos indicadores pueden parecer burdos, denotan el influjo sustancial de una cultura familiar preocupada por el desarrollo educativo de sus hijos. Más libros significa mayor acceso a la lectura, mayor acceso a la lectura implica mayor alfabetización y evidentemente mayor manejo de conocimiento, en particular conocimiento de mundo así como también conocimiento de temas específicos que en ocasiones no están al alcance de la experiencia directa (Krashen y Brown, 2005) De modo que las familias con entornos educativos empobrecidos en donde incluso la riqueza del lenguaje es mediocre, tienen una mayor probabilidad de que sus hijos o pupilos evidencien menor capacidad para comprender el lenguaje escrito o desarrollar adecuadamente las habilidades de escritura (Chall y Jacobs, 1983).

También se sabe que familias con mayor alfabetización tienden a valorar mucho más las oportunidades de conocimiento, sean estas formales o informales, como por ejemplo asistir a museos o bibliotecas o leer mucho más y valorar la lectura (Noble, Farah y McCandliss, 2006; Piccolo, Arteche, Fonseca y Grassi-Oliveira, 2016). En familias de esta clase, los padres juegan un papel fundamental, pues suelen apreciar más las instancias educativas y son capaces de ver a largo plazo los beneficios de una educación sólida en sus hijos. Es por ello que muchas investigaciones incluyen la educación de los padres y sus años de estudio como un factor relevante en el desempeño futuro de los estudiantes (Leserman y de Jong, 1998; McEwan, 2003; Somers, McEwan y Willms, 2004).

En este punto se puede señalar que el nivel socioeconómico como tal no es un factor explicativo por sí mismo, sino que enmascara una serie de variables que favorecen o no los logros de los estudiantes. En este sentido la pobreza, por sí misma, no representa necesariamente la explicación de las dificultades en la lectura, pero las consecuencias de esta, como la carencia o la falta de oportunidades sí pueden ser indicativos de ello. Bajo condiciones de privación que muchas familias en condición de pobreza enfrentan, es

difícil que tengan la capacidad de proveer un ambiente enriquecedor para mejorar la lecto-escritura y con ellos, los aprendizajes (Jadue, 1997).

Otro factor tiene relación con la presencia del factor familiar histórico, pues se estableció que algunas familias funcionan como “bancos de conocimiento”, en donde la lectura se considera parte de la adquisición y desarrollo de este banco, y por tanto, a través de este se promueven relaciones y se enseñan valiosos conceptos y habilidades relativos a éxito y prosperidad (Moll, Amanti, Neff y González, 1992).

### **La influencia de la escuela en el desarrollo comprensivo lector**

La escuela, por su parte, representa un segundo contexto a considerar. Desde un inicio se tomó el aprendizaje de la lectura como un proceso social, en el que las interacciones profesor-estudiante proveen el contexto para que estos últimos puedan aprender. Un aspecto clave son las expectativas del profesor, que al parecer tienen un importante efecto sobre el logro del alumno: Docentes que piensan que sus estudiantes deben ser rescatados de la influencia negativa de sus comunidades y/o familias, o bien docentes que creen que alumnos en condiciones de pobreza tienen pocas probabilidades de salir adelante. Es interesante notar que esta clase de expectativas ha llevado, en el contexto anglosajón americano, a que los estudiantes de escuelas pobres reciban educación estructurada, programada, mientras que los de escuelas pudientes reciben instrucción enfocada en pensamiento crítico y creatividad (McIntyre, 2011). No es difícil imaginar el impacto que esta clase de educación que solo atiende a habilidades cognitivas básicas tiene en el desarrollo de la lectura, en comparación con una educación centrada en habilidades cognitivas avanzadas.

Además de las expectativas, otro aspecto es la formación de los profesores. En escuelas vulnerables, y especialmente a nivel nacional, pueden presentarse obstáculos como menores recursos, uso metodologías inadecuadas o atrasadas y profesores menos capacitados pueden tratar las dificultades presentes en la enseñanza de la lecto-escritura.

Un tercer aspecto es el efecto pares, pues en estas escuelas, con una composición social homogénea, la mayoría de los estudiantes denotan un vocabulario reducido, lo que también tiene efecto sobre el aprendizaje inicial de la lecto-escritura (Compton et al., 2009; Marchant, Lucchini y Cuadrado, 2007).

Los padres también tienen algo que decir, pues son bastante conscientes de que las escuelas económicamente más fuertes pueden proveer la producción de desempeños académicos mucho más competitivos para sus hijos, lo que a la larga les permite acceder a profesiones de mayor prestigio, ascendiendo en la escala social. Hay de fondo, una especie de promesa de protección social de los pupilos, pues los salvaguarda de vincularse o mezclarse con individuos que pertenecen a contextos sociales desfavorecidos, lo que se asume como un obstáculo para alcanzar metas más altas a nivel social y profesional. En cierto modo, se compra segregación social, pues se piensa que es beneficiosa a largo plazo (Canales, Bellei y Orellana, 2016).

En virtud de lo señalado, no es difícil inferir que un estudiante que no tenga las oportunidades adecuadas se encuentra en franca desventaja no solo para nivelar sus aprendizajes, sino también para enfrentar su desarrollo futuro en muchas otras áreas, sean académicas, organizacionales o sociales en general. Esto es más evidente si se tiene presente la distancia descrita entre la educación de la escuela privada y la pública y cómo esta diferencia puede representar un factor de desigualdad, en particular si se tiene presente los mecanismos que generan dichas desigualdades.

La investigación ha brindado un consistente aval al dominio lector como elemento fundante y vital de todos los ámbitos de aprendizaje. Y aunque esta noción está sólidamente asentada en la literatura, debe agregarse a la preocupación por su dominio, el que no es una habilidad que se pueda adquirir en cualquier momento, sino que se obtiene fruto de un proceso que se desarrolla en la escuela gradualmente desde los primeros cursos de primaria, avanzando por textos cada vez más complejos. Si la habilidad no se adquiere en etapas tempranas del desarrollo, entonces su dominio se verá mermado y afectará los resultados de otras áreas académicas (Marchant, Lucchini y Cuadrado, 2007). Las consecuencias, desde la perspectiva de la vulnerabilidad son severas, pues ello implica que en escuelas más pobres, los estudiantes que no logran a tiempo un dominio lector suficiente, no solo se enfrentan a dificultades futuras evidentes, sino que también hay menos posibilidades de nivelarlos.

Las diferencias que haya entre estudiantes con recursos y aquellos desfavorecidos tenderán a perpetuarse más allá de la escuela (OECD, 2016). En el estudio de Chall y Jacobs (1983) se determinó que en los primeros años de lectura, los estudiantes de bajo

nivel socioeconómico no se diferenciaban tanto de aquellos con nivel medio, pero a medida que pasaba el tiempo y los niveles de lectura tendían a ser más complejos, aquellos de bajo nivel mostraron claros indicios de desaceleración en su proceso de aprendizaje. No solo eso, esta desaceleración también se relacionó con desaceleración en otras aptitudes como matemáticas o escritura y con menores habilidades de lenguaje, lo que no resulta extraño dada las consideraciones previas pues el lenguaje juega el papel de mayor influencia, pues es un reflejo poderoso de la influencia sociocultural en el que el estudiante se inserta (Völkel, Seabi, Cockcroft y Goldschagg, 2016). Peor aún, estudiantes con dificultades de lectura al inicio de su escolaridad, tienden a seguir siendo malos lectores a futuro, al contrario que los que son desde el inicio buenos lectores (Juel, 1988).

El nivel socioeconómico tiene, claramente, efectos sobre los niños y niñas en diferentes aspectos, en particular en aquellos que son muy jóvenes (hasta los nueve años de edad). Esto incluye otras aptitudes de carácter cognitivo como el cociente intelectual, la memoria y algunas funciones ejecutivas (Piccolo, Arteche, Fonseca y Grassi-Oliveira, 2016).

Si se considera, además, que para comprender lo que se lee se requiere el concurso de muchas otras habilidades cognitivas, entonces el retraso en la lectura revela también las dificultades del estudiante para construir una representación mental coherente de los textos en la memoria del lector (Kendeou, van den Broek, Helder y Karlsson, 2014).

El desarrollo de la comprensión lectora posiblemente no sería considerado un determinante de desigualdad a futuro de no ser porque pone en evidencia la relevante relación entre el contexto social (familiar y escolar) sobre el desarrollo de las aptitudes cognitivas. El aspecto cultural da énfasis sobre los aspectos sociales y su influencia en dicho desarrollo. El factor social y en particular la interacción social, afecta el desarrollo de la cognición de modo que el aprendizaje ocurre a través de la participación en un contexto sociocultural (Dehqan y Samar, 2014). Si cada persona es parte de un universo cultural específico, entonces se ve influido por este y a la vez, intenta influir o contribuir al mismo. Las personas recrean, pero también crean mundos culturales. Dependiendo del grado en que ese contexto cultural está empobrecido o enriquecido, es el grado en que un individuo recrea y crea (Gómez, 1997a). El caso de un estudiante refleja esta situación

muy bien considerando que este viene con una visión de mundo y un dominio verbal establecido desde el seno familiar (Gómez, 1997b).

La comprensión lectora, entonces, refleja las diferencias fruto de la preeminencia del factor económico, pero también evidencia la importancia del factor social (familiar y escolar) en estas particulares desigualdades. Aun notando que el nivel socioeconómico tiende a enmascarar varios de los determinantes diferenciales, es clave el hecho de que capital económico y capital cultural están, en la sociedad actual, bastante relacionados (Yang, Rosén y Gustafsson, 2011). En efecto, el capital cultural tiene un efecto positivo en los niveles de comprensión, mientras que el aspecto económico aislado no tiene la misma influencia. No obstante, la distinción entre ambos aspectos llega a desdibujarse bastante, más aún cuando pagar más por educación representa también vincularse con quienes han tenido mayores oportunidades educativas, lo que incluso puede llevar a los padres a cambiar de escuela a sus hijos cuando notan que las redes sociales del establecimiento se deterioran (Yang-Hansen, 2009; Yang, Rosén y Gustafsson, 2011).

Siendo la comprensión lectora una habilidad tan importante para el desarrollo del aprendizaje y de la inserción profesional, así como en la obtención de oportunidades futuras de diversa índole, y habiendo enunciado en este punto que hay factores sociales y económicos que influyen en las habilidades cognitivas vinculadas, en este estudio se le toma como la variable principal de la mediación de aprendizajes y representa el tema alrededor del cual se abordan los posibles determinantes tanto extra como intrapsíquicos.

### **CAPÍTULO 3: Un modelo teórico para la comprensión lectora**

Con la pretensión de desarrollar un estudio de las diferencias entre escuelas públicas y privadas y, en última instancia, evidenciar a un nivel detallado cuáles son los atributos cognitivos que contribuyen a los esperados desempeños diferenciales, se hace necesario adoptar un modelo teórico sobre comprensión lectora que constituya la base con la cual establecer qué habilidades específicas determinan un mayor o menor grado de comprensión lectora y cómo se pueden establecer diferencias a partir de ellas.

Los modelos a aplicar son variados y no hay una prescripción específica, siendo del arbitrio del juez o experto, el que decide qué atributos cognitivos participan en la ejecución de una tarea específica de comprensión lectora.

A lo anterior se suma el hecho de que la comprensión de la lectura no es explicada por una única teoría, sino que existen múltiples opciones que enfatizan diversos mecanismos que intentan explicar la génesis, desarrollo y consolidación de la aptitud comprensiva. No obstante, para poder aplicar los modelos de diagnóstico cognitivo es necesario plantear un modelo teórico que cumpla un conjunto de requisitos: a) Los jueces o expertos hayan hecho uso del mismo con el fin de establecer las habilidades en las que se basan las tareas de comprensión de la prueba aplicada en la investigación y b) sea un modelo cognitivo respaldado por la evidencia empírica y un referente en lingüística y psicología (Karreman, 2004) que intente brindar una explicación completa, desde el reconocimiento de palabras hasta la representación del significado de un texto. El énfasis de la teoría estará en comprender, justamente, dicho significado.

#### **Tipos de modelos en comprensión de lectura**

Era de interés proponer una teoría a partir de la cual se pudiese trabajar con el instrumento de recolección de información que se seleccionó para determinar el nivel comprensivo lector, pues con visiones teóricas distintas las tareas de evaluación también varían porque deben concordar con la perspectiva sostenida. Cuando ello ocurre, las habilidades a la base de la resolución de dicha tarea también cambian y en consecuencia, el éxito o no de los estudiantes dependerá de los atributos o habilidades que estén contenidos en la tarea que se les proponga en un instrumento (Tatsuoka, 2009).

Canet-Juric, Andrés y Ané (2005), distinguen tres tipos de modelos:

- Los modelos de procesamiento ascendentes (bottom-up): En que se le concede especial importancia a los procesos perceptivos en desmedro de los procesos cognitivos de orden superior (Alonso y Mateos, 1985). Por ejemplo, en el caso del modelo de Just y Carpenter se plantean cinco etapas: Percepción, codificación-acceso lexical, asignación de casos, integración intra-clausal y cierre de frase. Un modelo como este considera la lectura como un proceso en secuencia jerarquizado, que comienza al identificar las letras, ampliando su identificación a unidades lingüísticas más grandes como palabras y eventualmente frases. Dicho de otro modo, la información se procesa de abajo hacia arriba, comenzando por lo más simple (las letras) hasta llegar a algo tan complejo como la globalidad del texto (Alonso y Mateos, 1985). En modelos así, resulta evidente la importancia de la decodificación (en consecuencia, el aprendizaje de los estudiantes se centraba en una adecuada decodificación). La comprensión de lectura se entiende un derivado de la correcta interpretación de los caracteres de un texto.
- Los modelos de procesamiento descendentes (top-down): En los que, a diferencia del caso anterior, se le da importancia a los procesos cognitivos superiores, que se asume controlan y dirigen la comprensión de lectura. El modelo de lectura dirigida conceptualmente de Goodman (Brunning, Schraw, Norby y Ronning, 2007) es un buen ejemplo de esta aproximación. La lectura, está ahora guiada por las hipótesis que el lector propone sobre el significado de los textos y para ello, este hace uso anticipado de sus conocimientos sintácticos y semánticos, antes que de los aspectos gráficos de dichos textos (Alonso y Mateos, 1985). El proceso de lectura tiene comienzo en quien lee, quien durante dicho proceso elabora supuestos o hipótesis sobre el texto que está frente a él. En cierto modo, ahora quien lee no es un actor pasivo o mero receptor de información, sino que su conocimiento previo y otras habilidades interpretativas, así como otras estrategias cognitivas y metacognitivas se ponen en funcionamiento para dar cuenta del procesamiento de nivel inferior (Davoudi y Hashemi, 2015). De este modo, el procesamiento en los niveles

inferiores, está dirigido por el procesamiento inferencial de nivel superior. Esta aproximación sigue siendo jerárquica, pero de carácter descendente. El acento ya no está en el texto o el material de lectura, sino en el lector, quien da significado a lo leído en función de sus esquemas previos de conocimiento, experiencias y modelos mentales. En términos de aprendizaje, la comprensión lectora no es ya un subproducto, sino un proceso privilegiado en que el lector juega un papel fundamental. La información ahora se procesa de arriba hacia abajo, iniciando con lo más complejo (palabras, frases) y de ahí hacia los elementos que las componen (letras, sonidos). No obstante lo anterior, la visión no estuvo libre de críticas a pesar de la preeminencia del procesamiento cognitivo superior, pues estudios mediante movimientos oculares revelaron que los lectores observan cada palabra mientras leen, lo que es un indicador de que no puede haber solo un procesamiento top-down.

- **Modelo interactivo:** Es un intento de unificación de los dos modelos anteriores. Se propone que el procesamiento ascendente y descendente ocurren de modo paralelo. La comprensión lectora está dirigida de modo simultáneo por los datos explícitos presentes en el texto y por el conocimiento preexistente del propio lector. El enfrentamiento de un texto lleva a quien lee a generar expectativas en distintos niveles, tanto al nivel inferior (reconocimiento de palabras, decodificación) como a nivel superior (sintáctico, semántico). En el nivel inferior, la información que se procesa se convierte en un input para el siguiente nivel y se propaga a niveles cada vez más elevados. Algo similar ocurre en el nivel superior donde emergen hipótesis que requieren verificación mediante indicadores del nivel inferior. La interacción de procesos ascendentes y descendentes lleva a la comprensión. Como consecuencia, el lector elabora un significado a partir de la interacción con el texto mediante el aprendizaje de ideas relevantes del mismo y las relaciona con las ideas que ya posee. En este sentido, comprender implica poseer un esquema o estructura que le permita asociar la información que se presenta en el texto con lo que el lector ya conoce. Un buen lector será aquel que sea consciente de estos esquemas de conocimiento, pues de este modo puede adoptar estrategias que le permitan organizar y estructurar la

información con el objeto de desarrollar una representación coherente, ordenada y jerárquica (Vidal-Abarca y Gilabert, 1990). En el modelo interactivo, esta noción de esquema es relevante. Se corresponde con una estructura cognitiva que representa los aspectos de mayor relevancia de un objeto, hecho o concepto. Se constituyen en estereotipos que se activan cuando se procesa información, lo que denota que las personas tienen estructuras cognitivas abstractas que son usadas para el procesamiento y comprensión de la información. En este modelo la lectura se considera como un proceso constructivo donde el significado de lo leído depende del lector y no solo del material de lectura, pues la persona aporta sus propios esquemas para favorecer la comprensión (Alonso y Mateos, 1985).

- Modelos proposicionales: Desde esta perspectiva se asume que la proposición es la unidad básica para representar adecuadamente el significado de un texto. Su génesis proviene desde al menos los años 70 y emerge como crítica y reflexión desde las teorías relacionadas con la representación de los textos, y el recuerdo de lo más importante dentro de un texto. La idea proviene del hecho que los textos denotan un cierto tipo de discurso y que dicho tipo impone una forma y nivel de comprensión distinto. En efecto, entre los tipos de discurso (descripción, narración y exposición-argumentación) se sabe que los textos de carácter narrativo son más fáciles de entender que los expositivos por diversas razones, entre ellas porque los narrativos son ficticios y no debemos preguntarnos constantemente por su valor de verdad, los narrativos refieren a un contexto específico, mientras en los expositivos dicho contexto o no está presente o solo lo está de modo ambiguo, en los narrativos hay una secuencia de eventos y se hacen más inferencias, en los expositivos los medios para dar cuenta de las relaciones de coherencia tienen un papel mucho más importante que en los narrativos (Alonso y Mateos, 1985). Dado que los textos de enseñanza suelen ser de carácter expositivo, en estos la comprensión reside en otros atributos distintos a los que se suelen tomar en consideración en la comprensión de la narración. En específico, en la narración la construcción del significado parece descansar en el contenido, mientras que en los textos expositivos el significado reside en la estructura proposicional y superficial

del texto. Es por ello que la mayor parte de la investigación en los textos expositivos ha estado tradicionalmente fundada en el estudio de su estructura, su relación con la estructura teórica resultante y en el recuerdo de los lectores. Para Kintsch y Van Dijk (1978) el lector es un componente fundamental pues son ellos quienes representan el significado de los textos a través de la construcción de la denominada “microestructura” o conjunto de proposiciones formadas por un predicado y uno o más argumentos vinculados entre sí. Un texto tiene una base proposicional que se presenta como una lista jerárquica de proposiciones. La base o microestructura corresponde a la lista de proposiciones ordenadas siguiendo un orden jerárquico y de nivel que intenta representar el sentido de un texto. Aquellas proposiciones que ocupan un nivel más alto de importancia, son más fácilmente recordadas. El recuerdo de los sujetos denota que el significado del texto ha sido conectado con el mundo, con la situación que este describe. Esta habilidad de recuerdo de las proposiciones e ideas más importantes no es una adquisición que pueda darse por sentado, pues se requiere de un proceso de desarrollo y aprendizaje de parte de los sujetos. A partir de la microestructura, se construye la denominada “macroestructura”, que se define como la representación semántica del significado global del texto (García, Cordero, Luque y Santamaría, 1996). Un buen lector es capaz de desarrollar la macroestructura del texto leído utilizando como base su conocimiento sobre cómo están organizados los textos en función de los signos o elementos que los caracterizan (títulos, marcadores retóricos, resúmenes). En este sentido, entiende la formación de la macroestructura como un proceso en que el lector aplica su saber para identificar, reconocer y seleccionar las ideas más importantes de un texto, utilizando las indicaciones y las señalizaciones que están incorporadas en este (García, Cordero, Luque y Santamaría, 1996). Sobre la macroestructura, se encuentra la superestructura, que se corresponde con diferentes estructuras globales de orden convencional y que son comunes a algunos textos. La perspectiva proposicional considera la lectura como un proceso de carácter interactivo entre el lector y el autor del texto, donde el primero construye el significado, pero siempre apoyándose en las pistas de organización que ofrece

el texto, junto con la información proveniente de sus propias experiencias y conocimientos previos.

Aunque existen otras clasificaciones sobre los modelos teóricos, algunos de ellos con énfasis histórico y temático (Rayner y Reichle, 2010; Tracey y Morrow, 2017), la esencia de esta referencia es brindar un panorama general de las aproximaciones en el ámbito de la comprensión lectora para posteriormente presentar con algo más de detalle, la perspectiva que fue considerada en este estudio y que se dispone en el siguiente apartado.

### **Introducción al modelo comprensivo lector de Kintsch y Van Dijk**

Se ha adelantado que el trabajo de Kintsch y Van Dijk se encuentra dentro de los denominados modelos proposicionales que son los que actualmente concitan mayor interés, por ser de base cognitiva y contar con un adecuado sustento empírico (Herrada-Valverde y Herrada, 2017; Karreman, 2004). Esta teoría describe el proceso completo de lectura, desde el reconocimiento de palabras hasta la construcción de una representación del significado de un texto.

Las bases de la aproximación cognitiva de Kintsch se encuentran en los primeros trabajos llevados a cabo por él en los años 70, momento en los que el autor propone la teoría de procesamiento de lenguaje a partir de la descripción semántica de los textos. En este marco, las proposiciones conforman la base de la representación del significado de los textos. Las proposiciones están constituidas por un conjunto de conceptos o microproposiciones, las que se reconocen según su función como argumento o bien como predicado, de tal manera que dichos conceptos se definen a través de la proposición de la misma. La base proposicional también debe ser entendida como una lista jerárquica de proposiciones, lo que se conoce como “base de texto”. Desde el punto de vista de esta perspectiva, el foco de una teoría psicolingüística como esta está en intentar explicar si la estructura semántico-proposicional de un texto influye en la comprensión y en el recuerdo del mismo. El recuerdo de estos conceptos constituiría información que el lector ya tiene almacenada en su memoria de largo plazo y refleja su saber y sus experiencias previas. Se asume que habrá un mejor recuerdo con las proposiciones que dentro del texto tengan un lugar más alto en la jerarquía semántica. Propositiones que corresponden a las

denominadas macroproposiciones (Alonso y Mateos, 1985; Canet-Juric, Andrés y Ané, 2005).

El modelo original de Kintsch evolucionó para incorporar nuevos aspectos superando la propuesta teórica inicial y pasó a conocerse como modelo de construcción-integración (Herrada-Valverde y Herrada, 2017).

En el modelo de construcción-integración, se define la comprensión lectora como un proceso mental en donde el lector construye dos tipos de representación: a) La mencionada base de texto o base textual y b) el denominado modelo de situación o representación situacional (Herrada-Valverde y Herrada, 2017, Tijero, 2009).

Por un lado, para poder acceder al significado del texto, se necesita acceder al nivel semántico superficial o microestructura, lo que permite aproximarse a la coherencia local del texto o relaciones lineales entre las proposiciones (lo que también se entiende como captar las relaciones entre dichas ideas o proposiciones) y por otro, introducirse en el nivel semántico profundo o macroestructura, lo que permite acercarse a la coherencia global del texto o relaciones jerárquicas entre las ideas principales (Alonso y Mateos, 1985; Canet-Juric, Andrés y Ané, 2005; Herrada-Valverde y Herrada, 2017).

Por otro lado, la construcción de la representación situacional, o como también se le conoce, el “modelo situacional” o de situación, consiste en simular mentalmente la situación en la que se enmarca lo proposicional, con el fin de recrear el mundo en el que esto está inserto. El modelo de situación define hechos que son descritos en las relaciones semánticas que se aprecian a nivel local y global (son aspectos considerados intencionales y plausibles).

Aunque “base de texto” y “modelo de situación” contribuyen para permitir la comprensión de textos, la diferenciación que se hace entre ambas pretende ser una forma didáctica de separar una forma de comprensión más cercana a lo literal (lo semántico-proposicional) y otra más próxima a lo inferencial (Herrada-Valverde y Herrada, 2017). En la práctica, el lector no obtiene dos representaciones separadas, sino que construirá solo una a partir de las dos anteriores. La base de texto, de hecho, puede presentarse incoherente o sin sentido hasta que se inserta dentro de un modelo de situación pertinente.

En consecuencia, el modelo de construcción-integración tiene dos grandes niveles (Herrada-Valverde y Herrada, 2017):

- La representación textual: En la que i) la proposición es la estructura básica del significado, ii) la microestructura es el nivel semántico superficial que permite determinar la coherencia local del texto y iii) la macroestructura es el nivel semántico profundo que permite establecer la coherencia global de un texto.

En la representación textual necesitamos identificar el significado de cada palabra, teniendo en cuenta el contexto verbal en las que se enmarcan. Este conjunto, que se denomina proposición, contiene dos o más conceptos relacionados y constituye la unidad mínima de significado dentro de un texto. Es posible atribuirle valor de verdad, en particular dentro del tipo de texto expositivo, como suelen ser los textos usados para aprender. Las proposiciones constan de un predicado (un verbo, un adverbio o un sintagma nominal) y uno o más argumentos (Kintsch y Van Dijk, 1978). Hay que tener presente que la construcción de una proposición no es libre. Como estas poseen un predicado, este determina el tipo y número de argumentos que lo pueden acompañar. Por ejemplo, el verbo “dar” requiere tres argumentos para conformar una proposición: El agente que da, el objeto que se da y el objetivo de quien da. Las proposiciones pueden constituir proposiciones más complejas si se incrustan unas dentro de otras. Cuando esto ocurre, emerge un núcleo proposicional (Fuentes, 2014) que integra argumentos pero prescinde de los predicados de proposiciones de menor nivel (Herrada-Valverde y Herrada, 2017). Las proposiciones complejas son estructuralmente más simples, pero recogen el mismo significado agregando mayor fluidez al discurso, en virtud de dicha simplificación.

De modo interesante, el campo de la comprensión ha tomado las proposiciones simples (también llamadas atómicas) y las proposiciones complejas, en una doble vertiente: Como unidades textuales y como unidades psicológicas que un lector construye en función de su lectura (Kintsch y Van Dijk, 1978).

La microestructura textual, representación o “código de superficie” (Riffo, 2016), representa el nivel semántico superficial del cual es posible extraer proposiciones, sean estas atómicas o complejas, y supone establecer la coherencia que hay entre dichas proposiciones a nivel local, es decir, el tipo de relaciones entre dichas proposiciones. Para ello es necesario no obstante, reconocer e inferir los dos tipos de relaciones que se dan entre las proposiciones. De acuerdo a lo expuesto por Herrada-Valverde y Herrada (2017) estas son: Las relaciones temáticas y las relaciones estructurales.

Las relaciones temáticas son vínculos de referencia entre argumentos que implican solapamiento argumental de proposiciones. Es en virtud de estas relaciones que el lector puede establecer la coherencia temática de un texto y ser capaz de reconocer el hilo conductor, así como los cambios que puedan producirse (Kintsch y Van Dijk, 1978).

Las relaciones estructurales facilitan la progresión temática aunque no agregan nada nuevo al contenido, algunas pueden aludir a solución (unas proposiciones representan la solución a un problema o necesidad), a resultado (unas proposiciones representan el resultado que provocan otras), a causa, a motivación, a propósito y a justificación. Es posible también que algunas de las relaciones estructurales sean de contraste (unas proposiciones contrastan con lo que expresan otras), o de listado (las proposiciones representan elementos de una lista) (Herrada-Valverde y Herrada, 2017).

La macroestructura textual, que también es parte de la representación textual permite reconocer cuales de las proposiciones son relevantes y establecer coherencia a nivel global entre ellas. Representa un nivel semántico profundo que permite al lector identificar y reorganizar jerárquicamente las principales ideas presentes en la microestructura (Herrada-Valverde y Herrada, 2017). Se debe recordar que el lector construye a partir de la microestructura la representación semántica de su significado global. Lograr formar la macroestructura implica el logro de la coherencia global al conectar entre sí las ideas en función del significado de conjunto del texto (García, Cordero, Luque y Santamaría, 1996).

Esta macroestructura está formada por macroproposiciones que representan el tema general de un texto. El sujeto llega a ella a través de un proceso inferencial mediante la aplicación de ciertas estrategias y la información gatillada en el lector por el propio texto (García, Cordero, Luque y Santamaría, 1996). En el proceso, resulta de crucial necesidad la forma organizativa de un texto, pues este brinda una serie de claves acerca de las proposiciones de la microestructura que pueden ayudar a formar macroproposiciones de forma organizada y jerárquica. Esta organización textual también se le conoce como superestructura, que facilita mucho la comprensión (Van Dijk y Kintsch, 1983). La superestructura permite al lector disponer de ciertos criterios que, al tener en cuenta las claves propias de cada texto, permiten categorizarlos con base en una forma organizativa específica, pudiendo ser capaz de identificar textos con énfasis en: Un problema y una solución, la causalidad, lo comparativo, lo enumerativo, lo descriptivo y lo argumentativo. De este modo, la superestructura del texto permite al lector extraer la macroestructura de la microestructura (Herrada-Valverde y Herrada, 2017).

- La representación situacional: O modelo de situación, que considera necesariamente el recuerdo y los conocimientos previos del lector. Implica integrar la base textual dentro de un contexto que le dé sentido (Tijero, 2009). En efecto, si el lector desea comprender un texto, debe encontrar un referente claro que le permita comprender la situación a la que se refiere el texto. De no lograrlo, la comprensión no será posible y además, el recuerdo asociado será pobre. La referencia de lo situacional posibilita que el lector encuentre una referencia que le permita encuadrar el contenido de un texto en una situación concreta. De este modo, el lector situará el contenido textual dentro de una dimensión temporal, causal y espacial (Herrada-Valverde y Herrada, 2017; Tijero, 2009). Así, ciertos referentes de un texto, pueden llevar al lector a situarlo, por ejemplo, en una compra de pizza, en la antigua Roma, en un espacio específico relativo a estos ejemplos y con base en una cierta relación causal de hechos.

Hay dos elementos más dentro de este modelo, que aquí no se consideran parte de estos niveles (aunque algunos lo mencionen de ese modo), pero que son aspectos fundamentales de la teoría: El uso que tiene la memoria de trabajo del lector (que permite que ambos tipos de representación interactúen) y los procesos inferenciales.

- La participación de la memoria de trabajo: El modelo de construcción-integración entiende la comprensión lectora como un proceso híbrido en donde la memoria de trabajo recibe como input o entrada un conjunto de símbolos lingüísticos relacionados gramaticalmente entre sí (código de superficie) y produce como output o salida una red de símbolos semánticos (base de texto) que se basa en las relaciones del input con los conocimientos previos del lector, afinando la base de texto que se ha obtenido mediante la generación de un modelo mental (modelo de situación) que se integra dentro del conjunto de conocimientos del lector pasando a formar parte de la memoria de largo plazo. Se debe tener presente que la implicación de la memoria de trabajo u operativa implica que durante la lectura se procesa información por partes o ciclos y no todo de una vez. Se retiene durante un tiempo limitado (alrededor de 20 segundos por vez) y cada ciclo de tiempo permite retener unas cinco o nueve proposiciones. La representación textual y situacional se construyen de forma simultánea dentro de cada ciclo y de modo progresivo a través de diferentes ciclos (González, 2016; Herrada-Valverde y Herrada, 2017).
- La relevancia de las inferencias en el modelo de construcción-integración: Se le asigna importancia a la información que proporciona el texto en sí para la construcción de inferencias asociadas a la coherencia local, de modo que dicho proceso sería esencialmente ascendente. No obstante, este modelo contempla la posibilidad de añadir inferencias en la representación textual, aunque su capacidad explicativa no se detiene allí, ya que también pueden haber inferencias basadas en el conocimiento del lector. Kintsch establece dos criterios básicos: Por un lado la recuperación de información inmediatamente disponible en memoria y por otro, la generación de información a partir de lo

propuesto en el texto, ambas inferencias contribuyen a la generación de un modelo de situación (González, 2016).

Aunque hay diversos matices presentes en el modelo de construcción-integración que no pueden ser tratados aquí debido a su complejidad intrínseca, los elementos expuestos dan cuenta de los aspectos centrales y cruciales de esta aproximación.

La comprensión lectora es una habilidad que se desarrolla progresivamente no sin esfuerzo y cuyo avance depende, como se ha visto, de la coordinación y organización de un sinnúmero de procesos psicolingüísticos que, de no tener la oportunidad de llevarse a cabo como se espera, puede mostrar insospechados resultados donde el peor escenario es no lograr el nivel de comprensión requerido para enfrentar textos de dificultad académica creciente.

En este nivel sólo se puede conjeturar, pero es esperable que un conjunto de factores relevantes, previo a la implicación cognitiva que emerge de la complejidad expuesta, reside en la oportunidad de compartir mundos en donde el vocabulario, la riqueza de las experiencias y el conocimiento personal, tengan la posibilidad de desarrollarse. Ya sea producto de la construcción intercultural, zona de desarrollo próximo o como se le llame, el proceso comprensivo expuesto aquí se nutre de este amplio contexto donde aprender se empapa del entorno. Y el entorno, para los estudiantes, está representado mayormente por sus familias y por su escuela y más allá, por el contexto social en que ambos entornos se integran.

## **CAPÍTULO 4: Modelos de diagnóstico cognitivo en comprensión lectora**

Las aproximaciones cognitivas, como se ha visto, no son nuevas en lecto-escritura. No obstante, esta dimensión es asumida dentro de un marco teórico complejo e interactivo de relaciones entre capacidades que lleva a la construcción de una representación del significado general de un relato (Gutiérrez, García, Elosúa, Luque y Gárate, 2002; Lorenzo, 2001a; Lorenzo, 2001b). Su complejidad y alcance muchas veces lleva a múltiples opciones sobre el tipo de evaluación que es posible realizar con el fin de dar cuenta apropiadamente del o los dominios que la teoría intenta explicar (Ibáñez, 2012; Pérez, 2005).

Así, la evaluación de la comprensión ha adoptado muchas modalidades, algunas centradas en un enfoque de medida cuantitativo, otras en uno cualitativo, otras se han enfocado en el tipo de ítems o preguntas propuestas, en las características de los textos sobre los que se consulta o, soslayando otras cuestiones, en que la prueba aplicada tenga la validez propia de cualquier instrumento de los contruidos para tal propósito (Ibáñez, 2012). En efecto, la comprensión lectora, como ocurre con muchas otras variables de carácter cognitivo, se revela en virtud del tipo de tarea a la que son sometidos los alumnos. Dependiendo de la prueba, test o evaluación que se aplique, se obtendrán resultados específicos para la configuración de tarea de dichas pruebas (Martínez-Arias, Hernández y Hernández, 2006). Ibáñez (2012) y Pérez (2005) exponen extensivos listados de formas de evaluación que repercuten en los resultados de comprensión de textos que se obtendrán con su utilización. Esto es aún más importante porque, como se ha visto, no hay una sola aproximación teórica al fenómeno de comprensión lectora, abundando modelos más o menos populares, tales como el de lectura dirigida por datos de Gough, el modelo de lectura dirigida conceptualmente de Goodman o el modelo de construcción-integración de Kintsch (Brunning, Schraw, Norby y Ronning, 2007; Herrada-Valverde y Herrada, 2017). Con visiones teóricas distintas, las tareas de evaluación también varían porque deben concordar con la perspectiva sostenida. Cuando ello ocurre, los factores a la base de la resolución de dicha tarea también cambian. Dependiendo de qué atributos o factores un estudiante domine se obtendrá o no éxito en su resolución (Tatsuoka, 2009).

Bajo este panorama, la compleja teorización en torno a la comprensión se nutre de investigación empírica en que el enfoque de evaluación y en particular el modelo teórico que debe responder a un modelo de medición, condicionan la forma en que se da cuenta de los desempeños de los estudiantes, pues están determinados por el constructo o constructos específicos que se desea medir y la teoría que se teje en torno a ellos para argumentarlos.

Esta forma de acceder a las operaciones cognitivas implicadas en la resolución de una tarea definida en un instrumento, puede ser explicada desde la “perspectiva del razonamiento evidencial” (Martínez-Arias et al., 2006; Yan, Mislevy y Almond, 2003), que se refiere a cómo extraer inferencias sobre lo que los estudiantes conocen, pueden hacer o comprenden. La perspectiva desarrolla un conjunto de modelos entrelazados entre sí, desde el llamado modelo del estudiante, donde a nivel teórico se establece el constructo o constructos que se desea evaluar, pasando por vincular dicho constructo o constructos con un modelo de medida o modelo psicométrico, donde se establece la relación entre un rasgo latente y variables observables, cumpliendo con principios psicométricos que aseguran la calidad de esas medidas (fiabilidad, validez, comparabilidad, equidad), a las que posteriormente se les asocia un modelo de puntuación, que indica el procedimiento que permite asignar una valoración a las respuestas dadas por el examinado (Mislevy, Wilson, Ercikan y Chudowsky, 2003). Este enfoque expuesto por Martínez-Arias et al. (2006) obedece a la perspectiva propuesta dentro del marco de Diseño Centrado en Evidencia (Zieky, 2014) que se detalla más adelante.

La evaluación se transforma entonces en “evidencia”, que no solo informa del desempeño de un estudiante, sino también contribuye a fortalecer o al menos apoyar, la definición que se ha hecho de un constructo específico. Esta evidencia representa datos que sostienen una hipótesis teórica acerca de cómo los estudiantes adquieren conocimiento y habilidades (Pellegrino, Chudowsky y Glaser, 2001), lo que también representa un aporte a la validez de las interpretaciones que se hacen de las puntuaciones (Abad, Olea, Ponsoda y García, 2010).

En concreto, aproximarse a la comprensión de lectura y su medida, como puede ocurrir al intentar dar cuenta de cualquier otro rasgo que suponga la adquisición de una habilidad o aprendizaje, requiere la propuesta no solo de un modelo de medición, sino

también de un modelo de cognición, que permita describir como las personas representan el conocimiento (Kane y Bejar, 2014; Pellegrino, Chudowsky y Glaser, 2001). Por supuesto, una perspectiva como la de Kintsch (Brunning, Schraw, Norby y Ronning, 2007; Kintsch, 1998; van Dijk y Kintsch, 1983), provee un punto de partida fundamental cuando se trata de diseñar dicho modelo de cognición para la evaluación de la capacidad comprensiva en lectura, pues permite especificar con detalle los procesos cognitivos que juegan un papel en la resolución exitosa de esta habilidad, lo que bajo el marco del razonamiento evidencial y en particular, bajo un modelo de medición adecuado, permite atender los principios de un modelo psicométrico, en particular fiabilidad, validez y equidad (Martínez-Arias et al., 2006; Pellegrino, Chudowsky y Glaser, 2001).

Sin embargo, además de tener un modelo de cognición basado en evidencia empírica acerca de cómo se lleva a cabo el aprendizaje de un dominio, y brindar una descripción del pensamiento que se pone en juego frente a la resolución de una tarea (del “modelo de puntuación” al “modelo de tarea”, que pone a prueba los atributos que permiten resolverla) (Embretson y Gorin, 2001; Martínez-Arias et al., 2006), también se debe: a) proveer una identificación clara y diferenciada de los altos y bajos desempeños, b) representar la variedad de maneras en que típicamente un examinado puede entender dicha tarea, c) seleccionar una parte o subconjunto del dominio que se desea evaluar, pues es necesario que el evaluador sepa que aspectos necesitan ser mejorados y retroalimentados para poder pasar a dominios más avanzados y por último, d) un modelo de cognición debe prestarse para ser combinado dentro de un conjunto mayor de información, para que pueda ser usado para propósitos distintos a la evaluación original (Pellegrino, Chudowsky y Glaser, 2001).

La noción de un modelo de cognición así comprendido representa una metodología que requiere un mayor desarrollo que los tradicionales modelos, teniendo fuertes implicaciones en la calidad psicométrica de las evaluaciones. Las interpretaciones que pueden hacerse de la medida del desempeño de un estudiante implicarán descripciones más complejas del logro de los estudiantes, especificando las habilidades o atributos que estos dominan o no, lo que puede ir incluso más allá, al incorporar el avance del aprendizaje mediante la identificación de niveles de complejidad entre dichas habilidades (Pellegrino, Chudowsky y Glaser, 2001). Esta aproximación, por supuesto,

sobrepasa lo que se puede conseguir a través de mediciones realizadas con tests unidimensionales (Kane y Bejar, 2014)

La perspectiva tradicional del desarrollo de tests, aun cuando se base en un marco conceptual sólido en torno al modelo cognitivo que sostiene el desempeño esperado en un instrumento, suele estar pensado de modo más global. Los ítems miden el desempeño explicado por un único constructo de base, esto es, una medida unidimensional (Abad, Olea, Ponsoda y García, 2010, Santisteban y Alvarado, 2001). Incluso modelos sofisticados, como los relacionados con la Teoría de Respuesta al Ítem (TRI), aunque son capaces de brindar información invariante sobre los parámetros del examinado y los ítems del test, siguen vinculados a índices globales sobre el rasgo de los examinados o de su nivel de competencia (Embretson y Yang, 2012; Yan, Mislevy y Almond, 2003). En el caso particular de la evaluación tradicional en educación, ya sea que esté enmarcada en la TRI o bien en la Teoría Clásica de los Tests (TCT), la medida del estudiante se establece por una puntuación que lo posiciona en un continuo de desempeño que si bien sirve para muchos propósitos, el de diagnóstico detallado sobre las capacidades de dicho estudiante no se encuentra entre ellas. Al decir de Ravand (2015), esto es porque las medidas tradicionales están pensadas como “evaluaciones del aprendizaje antes que evaluaciones para el aprendizaje”<sup>2</sup> (p. 1) si bien son de utilidad, pues las puntuaciones así obtenidas se pueden aplicar a diversos propósitos, como la clasificación entre quienes son o no seleccionados, la promoción o no de curso, o selección de candidatos para un programa, asignación de becas, entre otras opciones, lo cierto es que no satisfacen los requerimientos de información diagnóstica detallada.

Es por ello que se requiere de una aproximación que vaya más allá de la evaluación global. Otra aproximación brindará, eventualmente, un mejor diagnóstico de la situación de los estudiantes que estén en situación vulnerable o que requieran algún apoyo extra, en función de las condiciones de desigualdad que les toque experimentar. Si, como se sostiene en este trabajo, las diferencias entre estudiantes de escuelas públicas y privadas van más allá de lo socioeconómico, repercutiendo en sus desempeños académicos, será de mayor relevancia tener información precisa sobre sus habilidades

---

<sup>2</sup> “*assessment for learning rather than assessment of learning*” (Ravand, 2015, p. 1).

con el propósito de proveer una mejor orientación sobre los cambios a implementar para nivelar o al menos suplir en algún grado las desventajas.

Los modelos de diagnóstico cognitivo (CDM) representan una posible respuesta a esta pretensión. Tienen la capacidad de proveer información detallada sobre las habilidades de los examinados, brindando además la oportunidad de ofrecer retroalimentación sobre las razones del éxito o fracaso frente a un test dado (Li, Hunter y Lei, 2015; Ravand, 2015).

Un modelo de cognición con mejores propiedades de diagnóstico, representa una mejor ayuda para indicar lo que debe aprenderse en la escuela e implementar mejoras concretas que repercutan efectivamente en el desempeño de los estudiantes. Entendido de este modo, el desarrollador de la evaluación puede comenzar desde cero un análisis cognitivo sobre un dominio particular. No obstante, lo habitual es basarse en principios de cognición y aprendizaje que ya hayan sido establecidos, pues siempre se sostendrá que la construcción de una evaluación se genera desde lo que los estudiantes comprenden y saben de un determinado tema (Pellegrino, Chudowsky y Glaser, 2001).

Gorin (2007), así como Rupp, Templin, y Henson (2010) describen dos aproximaciones que profundizan en los principios de estos modelos: a) Desde una perspectiva educativa, se les enmarca en lo que se ha llamado Diseño Centrado en Evidencia (o ECD de su nombre en inglés, Evidence Centered Design) (de la Torre y Michen, 2014; Kane y Bejar, 2014; Rupp, Templin y Henson, 2010; Zieky, 2014). Se la considera una aproximación sistemática y lógica a la creación de tests; b) Sistema de Diseño Cognitivo (CDS, de su nombre en inglés, Cognitive Design System) cuya principal representante es Susan Embretson (Yang y Embretson, 2007). Curiosamente, aunque ambas perspectivas proponen cosas similares, siguen caminos distintos y rara vez hay citas cruzadas entre dichas líneas de trabajo, salvo algunas menciones de modelos específicos dentro del marco comprensivo de alguna de ambas visiones (DiBello, Roussos y Stout, 2007). Se detallan brevemente a continuación.

#### **a) Sistema de diseño cognitivo:**

En el caso del diseño de evaluaciones diagnósticas, el enfoque estuvo puesto originalmente en los procesos de respuesta con objeto de medir habilidades cognitivas básicas como rotación espacial u otras de la misma índole. Los constructos siguen siendo

la pieza fundamental de estas evaluaciones y por tanto, el soporte teórico que define, describe y explica dichos constructos es la base de estas propuestas (Rupp, Templin, y Henson, 2010).

De acuerdo a Embretson y Gorin (2001), el procedimiento de diseño debe iniciarse con el marco conceptual. Se ha mencionado la importancia que tiene el constructo con el cual se elaboran los tests. Se debe sin embargo, hacer una distinción, pues dos son los aspectos de este principio: a) Una adecuada representación del constructo, ocupándose de los procesos, estrategias y las estructuras de conocimiento que están involucradas en la resolución de un ítem. La investigación que se hace en torno al constructo alimenta esta representación. En función del constructo, se proponen los aspectos que conforman el estímulo (tarea o ítem), con el fin de variar las demandas cognitivas que dicho estímulo requiere; b) Por su parte, en este marco también debe incluirse el alcance nomotético, o las relaciones de las puntuaciones de los tests con otras medidas. Está vinculado con las correlaciones de las diferencias individuales entre variables medidas<sup>3</sup>.

Una vez aclarado el relevante papel que siempre impone un adecuado marco conceptual, se llega a la propuesta del marco de procedimiento, teniendo presente que si la intención es generar ítems o tareas para medir un constructo específico, entonces se requiere una explicación válida sobre cómo las características de esos ítems influyen en los procesos que están comprometidos en el desempeño frente a dicho ítem. Dado que el desempeño es una tarea compleja, a menudo se requiere más de una estrategia para llegar a una solución. Si los examinados pueden variar las estrategias que aplican para la resolución de la tarea, entonces las tareas complejas, que demandan desempeños complejos, son tareas de orden multidimensional (Embretson y Gorin, 2001). Como señala Embretson y Yang (2006): “Tareas de medición complejas, tales como la habilidad y el desempeño en un test, son normalmente multidimensionales en su naturaleza, más

---

<sup>3</sup> Se hace una distinción entre alcance nomotético y la denominación de red nomotética de Cronbach y Meehl (1955), porque el alcance nomológico se centra en la significancia y no en el significado y además, un fuerte sistema de hipótesis que se haya generado desde la investigación de la representación del constructo, debería guiar la investigación sobre el alcance nomotético (Embretson y Gorin, 2001, p. 350).

que unidimensionales. Las visiones contemporáneas sobre teoría cognitiva, incluso lo refieren a las habilidades más simples”<sup>4</sup> (p. 336).

De este marco de procedimiento se desprenden las siete fases centrales del diseño cognitivo. Las etapas son reproducidas en la obra de Rupp, Templin y Henson (2010):

- Especificación de los fines de la medición: La aproximación de diseño cognitivo requiere que se especifiquen dos tipos diferentes de fines: a) Para la representación del constructo y b) los relativos al alcance nomotético, anteriormente mencionado.
- Identificación de aspectos relevantes en el dominio de tarea: La identificación de aspectos específicos en el diseño de tareas es más sistemático y preciso en la labor de diseño cognitivo comparado con las aproximaciones tradicionales del desarrollo de tests. Se examinan las características de los ítems con el fin de manipular la representación del constructo y de este modo afectar los procesos cognitivos, las estrategias y las estructuras de conocimiento.
- Desarrollo de un modelo cognitivo del desempeño en la tarea: En esta etapa, primero es necesario identificar los procesos cognitivos, estrategias y estructuras de conocimiento y organizar todo en un modelo unificado. Una revisión de la literatura es fundamental pues no suele estar organizada en función de tipos de tareas. Segundo, las características del estímulo (tareas) que influyen en los procesos deben ser operacionalizadas. Para construir un modelo cognitivo para un tipo particular de ítem, estas características deben ser cuantificadas ya sea sobre ítems ya existentes o bien sobre aquellos que se hayan creado. Si el fin es generar ítems, entonces las características deben ser manipulables y puntuables. Tercero, el impacto de las características cognitivas sobre las propiedades psicométricas de un instrumento deben ser estudiadas empíricamente sobre ítems ya existentes. El impacto relativo de dichas características sobre la dificultad y la discriminación de un ítem

---

<sup>4</sup> “Complex measurement tasks, such as ability and achievement test items, are usually multidimensional, rather than unidimensional, in nature, Contemporary views on cognitive theory regard even the most simple ability items” (Embretson y Yang, 2006, p. 336).

permitirán evaluar el potencial de varios modelos cognitivos para la generación de ítems.

- Generación de ítems de acuerdo al modelo cognitivo: En esta etapa, las estructuras de los ítems así como las reglas de sustitución son desarrolladas para operacionalizar las características del estímulo en los ítems. Si las etapas precedentes han sido exitosas, entonces las variaciones en las estructuras de los ítems representan variaciones en los procesos. Ahora es cuando las características son seleccionadas para satisfacer las estructuras de los ítems y las reglas de sustitución. Entonces es cuando los ítems pueden ser incorporados en un test para su prueba empírica.
- Evaluación empírica del modelo cognitivo a través de las tareas administradas: Es la prueba empírica del modelo que subyace a la generación de los ítems. El éxito de esta etapa es fundamental para sustentar la representación del constructo y evaluar la modalidad de generación de ítems. Se evalúa tanto el modelo cognitivo como el modelo psicométrico. El primero mediante la predicción de desempeño en el ítem, considerando como variables dependientes los tiempos medios de respuesta y la dificultad, mientras que las variables independientes son la estructura de los ítems y las características del estímulo del ítem que operacionaliza los procesos cognitivos. El impacto relativo de las características considera el impacto de los procesos, estrategias y estructuras de conocimiento que ellas representan. El modelo psicométrico, por su parte, se evalúa mediante el ajuste de los datos obtenidos de la respuesta a los ítems. Si el modelo no ajusta, entonces es posible que se requiera modificaciones al modelo cognitivo. Se pueden hacer ajustes mediante la manipulación experimental de las características de los ítems y de este modo extraer conclusiones más sólidas sobre los efectos de los procesos cognitivos en los resultados de desempeño.
- Construcción de un banco de ítems en función de su complejidad cognitiva: Si el sistema de generación de ítems es efectivo en la predicción de sus propiedades, los ítems pueden ser almacenados en función de su complejidad cognitiva. Si los modelos matemáticos proveen una predicción

suficientemente buena de la dificultad del ítem, entonces estos pueden ser descompuestos en las fuentes que contribuyen a la dificultad. En consecuencia, los ítems pueden ser diseñados en función de su complejidad cognitiva y su dificultad.

- Validación del modelo mediante la revisión del alcance nomotético: Los ítems que se hayan creado, deben ser evaluados teniendo presente su alcance nomotético. Se realizan predicciones específicas a partir de la representación del constructo respecto de las correlaciones de las puntuaciones del test con conocimiento similar en tests de referencia o criterio.

Tal como señala Rupp, Templin y Henson (2010), los siete pasos resaltan la importancia de incorporar la determinación teórica de los procesos de respuesta del modelo cognitivo dentro del desarrollo de tarea. La integración viene dada por los modelos psicométricos que vinculan las características de la tarea (la dificultad y la discriminación del ítem) con variables propias del diseño de dicha tarea, lo que influye en las demandas del procesamiento cognitivo de esta.

El modelamiento cognitivo representa el paso desde el modelo teórico acerca de la forma de responder a una tarea, hasta la propuesta de tareas con características específicas (Martínez-Arias et al., 2006)

#### **b) Diseño centrado en evidencia:**

Más cercano al área educativa es la perspectiva del diseño centrado en evidencia. Rupp, Templin, y Henson (2010) señalan que fue desarrollada para ayudar a darle estructura al pensamiento de los diseñadores de evaluaciones diagnósticas, con mayor independencia de los constructos que sustentan la medida. La esencia del ECD (de su nombre en inglés, Evidence Centered Design) lleva a los diseñadores de evaluaciones a estructurar claramente la narrativa diagnóstica que les gustaría desarrollar sobre los examinados. La idea esencial del ECD es que las tareas se deben diseñar de modo que evidencien comportamientos que sean lo más informativos posibles sobre la estructura de la habilidad latente de los examinados.

Zieky (2014) señala que el ECD trata la evaluación como un proceso de razonamiento que se inicia con evidencia limitada de lo que los estudiantes hacen en una

situación de evaluación, para llegar a establecer lo que ellos saben y son capaces de hacer en la vida real. El ECD se define como “marco de principios para el diseño, producción y prestación de evaluaciones educacionales”<sup>5</sup> (Zieky, 2014, p. 79). El ECD asegura que la manera en que la evidencia se recopila e interpreta se apoya en el conocimiento subyacente y en los propósitos que orientan la evaluación. No son procedimientos rígidos, más bien es un conjunto de prácticas que ayudan a clarificar las inferencias que se hacen a partir de las puntuaciones, y de este modo determinar la mejor manera de proveer evidencia para apoyar dichas inferencias dentro de las restricciones que un programa de evaluación puede tener. Hay que tener presente que la esencia de un ECD es, precisamente, disponer este conjunto de prácticas de una forma tal que se dé cuenta de progreso en los aprendizajes.

Los pasos o etapas de esta perspectiva están organizados en función de “capas” e incluyen el diseño, desarrollo y uso de los tests (Kane y Bejar, 2014; Zieky, 2014). Estas capas se describen a continuación:

- **Análisis del dominio de conocimiento, habilidades o atributos de interés:** En esta etapa se revisa investigación relacionada con la forma en que los estudiantes aprenden, incluyendo una mirada profunda sobre la clase de situaciones que podrían servir para producir evidencia sobre el aprendizaje de los estudiantes y específicamente el nivel de desempeño que estos estudiantes denotan. Los niveles de progreso en el aprendizaje pueden proveer metas respecto de las cuales el estudiante puede ser evaluado. La determinación de logro va más allá que la mera clasificación jerárquica del desempeño, pues en este caso, el progreso se vincula con desempeños bien específicos que pueden usarse como guías en el desarrollo de evaluaciones que reflejen adecuadamente dicho progreso.
- **Modelado de dominios:** Donde se formulan los argumentos que sostienen la evaluación. Un resultado de este modelado es que las inferencias han de estar justificadas con la interpretación adecuada de las puntuaciones o bien el uso de información de apoyo. Estas inferencias, son afirmaciones que los examinadores desean hacer respecto del desempeño de los examinados. Otro

---

<sup>5</sup> “*principled framework for designing, producing, and delivering educational assessments*” (Zieky, 2014, p. 79)

resultado del modelado es uno o más patrones de desempeño como una vía para comenzar especificando componentes tangibles de dicho patrón. Estos patrones, que bien pueden considerarse patrones de aprendizaje debieran incluir distinciones entre niveles de progreso del aprendizaje. Además, los patrones debieran incorporar evidencia del desempeño que sería característico de estudiantes que estén a diferentes niveles en dicho desempeño. En esta etapa también deberían establecerse aquellos esquemas de las situaciones que podrían servir para gatillar la clase de evidencia que se ha mencionado.

- Marco conceptual de evaluación: Que en esencia describe los momentos presentados y que han sido expuestos por Martínez-Arias et al. (2006). Contiene en particular el modelo del estudiante, el modelo de evidencia (que abarca el modelo de medida y el modelo de puntuación), el modelo de tarea y el modelo de montaje o ensamblaje. Se entiende: a) El modelo del estudiante como las habilidades que, se hipotetiza, son de carácter latente y que se asumen como parámetros en un modelo psicométrico, por ello los parámetros son denotados por  $\theta_i = (\theta_{i1}, \dots, \theta_{ik})$ , para un examinado  $i$ , caracterizando los conocimientos, habilidades y/o atributos que son objeto de inferencia. Las variables en el modelo del estudiante son denotadas como  $\theta$  (Yan, Mislevy y Almond, 2003); b) los modelos de tarea son aquellos en que se detalla la tarea y sus características, así como las instrucciones para llevarlas a cabo; puede ser entendida como las actividades a proponer en un test. Los resultados de la tarea son expresados como  $Y_j = (Y_{j1}, \dots, Y_{js})$ , para la tarea  $j$  (Yan, Mislevy y Almond, 2003). El conjunto completo de resultados de tarea son denotados por  $Y$ ; c) el modelo de evidencia describe tanto las puntuaciones asociadas con ciertos desempeños de los estudiantes y las reglas para aplicarlas, así como el modelo psicométrico propiamente tal que hace las veces de nexo entre el modelo del estudiante y el modelo de tarea. El modelo de evidencia está basado en comportamiento observable o productos observables de la conducta, consecuencia de las respuestas a un determinado modelo de tarea. Las reglas producen los valores de las variables observables, que son denotadas por  $X_j = (X_{j1}, \dots, X_{js})$ , para la tarea  $j$ . Un modelo de evidencia también describe –

haciendo justicia a su condición de nexos- la estructura por la cual  $X_j$  depende de  $\theta$ . El conjunto completo de respuestas en todos los examinados y para todas las tareas se expresa como  $X$  (Yan, Mislevy y Almond, 2003). En el modelo de evidencia, la tarea del desarrollador de un test es describir en detalle los comportamientos que dan cuenta de los conocimientos, habilidades o atributos que son el foco de la medición; d) por último, el modelo de ensamblaje, que describe como el test se debe apreciar en su totalidad. Un modelo de ensamblaje contiene la información necesaria para construir formas paralelas de un test. En este modelo es donde se indican los atributos estadísticos que se espera tenga el test (dificultad, discriminación, fiabilidad, factor de velocidad). El modelo de evidencia permite además realizar un vínculo entre las inferencias y las tareas, si este indica los conocimientos, habilidades y atributos y las inferencias para las cuales se ha propuesto cada una de las tareas.

- Implementación de la evaluación: Estrechamente vinculado con las actividades de construcción de ítems y su incorporación en un test. Un concepto fundamental es la Estructura de Tareas, que permite generar una familia de tareas más o menos del mismo tipo, por lo que su ventaja de uso es permitir la construcción de tareas que pueden asumirse como paralelas. Por ejemplo: “¿Cuál es la [mediana, mediana, moda] de la siguiente distribución: [25-30 números de dos dígitos en orden aleatorio]?”<sup>6</sup> (Zieky, 2014, p. 83), donde lo que está entre paréntesis cuadrados es variable y puede ser modificado según interés. Además de los elementos variables, una estructura de tareas puede contener: i) una afirmación de los conocimientos, habilidades o atributos que pueden ser medidos por la tarea, ii) el sentido que se debe aplicar a las tareas generadas, iii) especificaciones para cualquier material que sea provisto con la tarea, iv) para tareas de elección múltiple, reglas para generar los distractores y v) rúbricas de puntuación para tareas de respuesta construida. Cabe señalar que si bien el flujo lógico parece ser crear una

---

<sup>6</sup> “What is the [mean, median, mode] of the following distribution: [25-30 two-digit numbers in random order]?” (Zieky, 2014, p. 83)

estructura de tareas y usarla para crear otras diferentes, no se requiere un estricto apego a estos pasos.

- Presentación de la evaluación: Consta de cuatro procesos: a) Selección de actividades, en que se seleccionan las tareas que serán presentadas al examinado. Tres son los mecanismos para presentar la tarea: i) Previamente seleccionadas, ii) al momento de evaluación, pero las tareas previas no condicionan las siguientes y iii) adaptativamente, donde las tareas posteriores están condicionadas por el desempeño de las tareas previas, y que es propio de los procedimientos de tests adaptativos informatizados (Barrada, 2012; Thompson y Weiss, 2011). El fin es seleccionar los ítems más apropiados para el nivel de habilidad del examinado, que en estos últimos dos casos se seleccionan mientras el test está siendo administrado; b) Presentación: Primero se presenta la tarea al evaluado y segundo, se registra su respuesta. Normalmente se suele hacer en formato impreso, aunque también existe la opción de hacerlo mediante sistemas informáticos. Las tareas donde se pone a prueba el desempeño, estas pueden ser de corta o larga duración; c) Procesamiento de las respuestas: En la ECD no se requiere ningún procedimiento particular de puntuación de respuestas a las tareas, lo que sí se requiere es que dichas puntuaciones estén basadas en una cadena de conexiones lógicas explícitas, esto es, pensar las puntuaciones de modo que sigan la misma cadena de razonamiento que guiaron la construcción de los test, pero en la dirección opuesta, es decir, mientras que la construcción del test se mueve desde las inferencias al modelo de evidencia y a las tareas, la intención aquí es que vaya desde el nivel de tarea, pasando por el modelo de evidencia, hasta obtener una evaluación o juicio sobre el nivel que tiene el examinado con respecto al conocimiento, habilidad y/o atributo que este denota; d) Resumen de puntuaciones: Rara vez se realizan juicios sobre una inferencia propia del modelo de estudiante a partir de una única tarea, de modo que los cuatro procesos previos requieren de la aplicación de algún tipo de procesamiento cuantitativo. Los sistemas de puntuación pueden variar ampliamente, pudiendo tener simplemente la contabilización de respuestas correctas, o bien se pueden tener múltiples habilidades y por tanto se pueden

calcular varias puntuaciones parciales. Un modelo de evidencia tendrá múltiples conexiones entre tareas y los diferentes conocimientos, habilidades y/o atributos y por tanto se requerirán modelos de puntuación mucho más sofisticados como los modelos de diagnóstico cognitivo (de la Torre y Minchen, 2014). Sin perjuicio del modelo de puntuación, uno de los aspectos más importantes de un ECD, es el vínculo que se puede establecer con categorías cualitativas y una representación enriquecida del desempeño que vaya más allá de la cuantificación.

Estas cuatro “capas” representan el núcleo de trabajo de los ECD. Por supuesto, hay un gran conjunto de matices que se deberían considerar en el proceso. No obstante, los pasos o capas señaladas dan cuenta de los mecanismos generales de construcción (Cane y Bejar, 2014; Deane y Song, 2014; Zieky, 2014)

De acuerdo a Rupp, Templin y Henson (2010) tanto los sistemas de diagnóstico cognitivo como la del diseño centrado en la evidencia representan contribuciones al proceso de validación de un test, pues constituyen un marco de trabajo dentro del cual se puede estructurar la relación entre teoría y proceso de respuesta, permitiendo la inclusión de la psicología cognitiva en la evaluación, así como elementos propios de los modelos estadísticos que se pueden usar para analizar los datos.

En consecuencia, los modelos de diagnóstico cognitivo juegan un papel de relevancia brindando apoyo a los argumentos de validez de los tests y representan una aproximación más precisa y exhaustiva que otras de cara a la evaluación de carácter diagnóstico, lo que es especialmente beneficioso si se requiere brindar apoyo específico en poblaciones que denotan situación de vulnerabilidad y de las que cabe suponer dificultades específicas, como se mencionará posteriormente.

### **Breve génesis de los modelos de diagnóstico cognitivo**

Los marcos de trabajo anteriormente expuestos representan un punto de partida para esta clase de modelos (Gorin, 2007; Rupp, Templin y Henson, 2010). No obstante, no hay un acuerdo sustancial al respecto y es necesario señalar, sin pretensión de exhaustividad, que otros autores sitúan el inicio de estos modelos mucho antes. En efecto, Baghaei y Kubinger (2015) indican, en el contexto de la validez de los tests y la

relevancia que tiene en ello conocer los procesos y operaciones cognitivas, que uno de los primeros modelos para estimar la dificultad de las operaciones cognitivas y su efecto sobre el desempeño fue especificado por Fisher denominándolo Modelo Logístico Lineal (o LTMM, de su nombre en inglés. No confundir con el modelo de regresión logístico lineal) (Fischer, 1973). En este caso, el modelo de Fischer es una extensión del modelo Rasch, que descompone los parámetros del ítem en una combinación lineal de varios parámetros básicos que han sido definidos a priori. El modelo también ha sido extendido a ítems politómicos con categorías de respuesta ordenada (Baghaei y Kubinger, 2015). Curiosamente, el mismo Fischer menciona que un paso previo y fundamental en la construcción de la historia de los modelos cognitivos lo dio Suppes, quien pretendió explicar la dificultad de problemas aritméticos elementales como la función de varios factores típicos, como por ejemplo, el tamaño de los números, más grandes y más pequeños, la forma de la ecuación que tiene que ser resuelta o el número de operaciones aritméticas implicadas (Fischer, 1973).

Según Romero (2010), los modelos de diagnóstico cognitivo tienen su raíz en el trabajo de Lazarsfeld y Henry (1977), relativo al análisis de la estructura latente, mediante el que se buscaba establecer la relación entre elementos observables y las clases latentes que permiten clasificar dichos elementos.

Por su parte, para von Davier (2005) los modelos de diagnóstico cognitivo propiamente tales encuentran su origen en los trabajos de Tatsuoka y Tatsuoka (1983). En ellos se asocian modelos basados en la TRI y en la clasificación de patrones de respuesta aberrante. La misma autora desarrolló estudios proponiendo una aproximación probabilística a la clasificación y diagnóstico de reglas erróneas de operación que provienen de una mala concepción en un determinado dominio aritmético. La variabilidad en los errores de respuesta es, efectivamente, modelada mediante la TRI (Tatsuoka, 1985).

Además de las contribuciones anteriores, cuya aparición en términos cronológicos refleja el temprano interés por desarrollar una aproximación operacional cognitiva en evaluación, están los hitos teóricos generados por autores de la talla de Messik (1989) y su trabajo sobre la validez de constructo, que indirectamente legitima la presencia de otras evidencias sobre la interpretación de los resultados de un test; o la reconocible mención

de Nichols (1994) a la necesidad de la sociedad por contar con medidas que vayan más allá de lo que él llamó “la aproximación de medidas psicométrico-educacionales”<sup>7</sup> (p. 577) o EPM (de su nombre en inglés), pues estas medidas tradicionales están más bien vinculadas a contenidos antes que a mecanismos cognitivos. Por último, están los argumentos de Snow y Lohman (1989), quienes aludieron a la relación esperable entre psicología cognitiva y la medición educacional.

En cuanto a los ámbitos de aplicación de los modelos, estas han abarcado áreas tan diversas como el diagnóstico clínico, la evaluación basada en estándares, con directa relación a la rendición de cuentas y quizás con mayor énfasis, en la evaluación diagnóstica en educación (Rupp, Templin y Henson, 2010). En estos dos últimos ámbitos resaltan los trabajos en dos dominios que juegan un papel de relevancia en el conocimiento y desarrollo de habilidades de otras áreas: Lenguaje y Matemática.

En los inicios del modelamiento cognitivo primó sobre todo el estudio de tareas de índole matemática (Klein, Birenbaum, Standiford y Tatsuoka, 1981; Tatsuoka, 1985). El dominio de un ítem requería poner en juego un conjunto de aptitudes para que el examinado pudiese lograr una respuesta correcta. Aunque se ha seguido investigando en esta área (Dogan y Tatsuoka, 2008; Tatsuoka, Corter, y Tatsuoka, 2004), se han incorporado nuevos elementos buscando explicar de manera más profunda el desempeño del estudiante y vincular con mayor solidez la evaluación cognitiva con lo que el profesor hace en clases, pues dicho desempeño se ve influido también por quién enseña y cómo le es enseñado o transmitido el conocimiento (Leighton, Gokiart, Cor y Heffernan, 2010). Estas líneas de trabajo representan los aportes futuros en este campo, pues se sabe que el conocimiento deficiente del profesor en los contenidos, en el componente pedagógico y en el conocimiento y transmisión pedagógica del contenido, tiene importantes efectos sobre los déficits comprensivos que presentan posteriormente los estudiantes (Goldrine, Estrella, Olfos, Cáceres, Galdames, Hernández, y Medina, 2015; Leighton, Gokiart, Cor y Heffernan, 2010).

El área de Lenguaje fue la segunda mayor área a abordar dentro del ámbito educativo desde la perspectiva de los modelos cognitivos (Buck, Tatsuoka y Kostin, 1997; Buck, Van Essen, Tatsuoka, Kostin, Lutz y Phelps, 1998; Embretson y Wetzel, 1987) y

---

<sup>7</sup> “*educational psychometric measurement (EPM) approach*” (Nichols, 1994, p. 577).

cobró importancia por las mismas razones que el dominio matemático: Su determinante contribución a la comprensión de otras disciplinas. En las siguientes secciones se discutirá con mayor detalle las implicaciones del lenguaje y en particular de la comprensión de lectura en el desempeño de los estudiantes.

### **Etapas de desarrollo del diagnóstico cognitivo**

Anteriormente se han descrito las etapas de desarrollo de marcos de trabajo más amplios en torno a la concepción de evaluación cognitiva, desde la psicología cognitiva y desde los modelos de evaluación (Gorin, 2007). En este apartado se destacan los pasos que se deben seguir para llevar a cabo un proceso de diagnóstico cognitivo en evaluación. La secuencia se le debe a Nichols (1994) quien la elaboró a partir de los trabajos de Gitomer en el Educational Testing Service.

- Paso 1. Construcción de teoría sustantiva: La base sustantiva se centra en el desarrollo de una teoría o modelo que describe el conocimiento y las habilidades que se piensa pueden estar implicadas en el desempeño, así como las características de la tarea o ítem que se conjetura interactúan con esos conocimientos o habilidades.
- Paso 2. Selección del diseño: Aquí el desarrollador del test selecciona los diseños de medición y observación. Esta selección se guía por la base sustantiva que ha sido construida previamente. Posteriormente, el desarrollador construye ítems o tareas que serán respondidas por los examinados de modos predecibles en función de los conocimientos y habilidades específicas sobre las que dichos ítems han sido elaborados, así como en función de otros atributos que también sean importantes en la teoría.
- Paso 3. Administración del test: Incluye cada aspecto del contexto en el cual los examinados lo responden, esto es, el formato de los ítems, la naturaleza de la respuesta, la tecnología usada para presentar los materiales del test y la forma en que se registran las respuestas y, por supuesto, el ambiente en la sesión de evaluación. Debería haber investigación asociada que permita informar sobre la forma en que el ambiente puede influir en la sesión de evaluación.

- Paso 4. Puntuación de las respuestas: El fin de este paso es asignar valores a los patrones de respuesta dados por los examinados de modo que se pueda vincular esos patrones a constructos teóricos. Un procedimiento de puntuación es parte de la operacionalización del diseño de evaluación.
- Paso 5. Revisión del diseño: Es el proceso por el cual se obtiene apoyo para un modelo o una teoría. Al igual que con cualquier teoría científica, la teoría usada en el desarrollo de un instrumento nunca está completamente probada; la evidencia se acumula gradualmente. En este paso, los resultados de la administración son usados para revisar la base sustantiva sobre la cual se funda la construcción de la evaluación.

Posteriormente, esta organización por pasos ha sido re-elaborada con mayor especificidad por DiBello, Roussos y Stout (2007), que es la que se expone a continuación:

- Descripción del propósito de la evaluación: Esta tiene importantes efectos para la caracterización que posteriormente debe hacerse del espacio del(os) atributo(s) latente(s). El propósito de una evaluación tendrá un relevante impacto si el conocimiento, habilidad o atributo es modelado con una o más variables, y si estas variables serán de carácter cualitativo, para propósitos de clasificación, o bien cuantitativo, para generar escalamiento. Es importante tener presente que las inferencias sobre las aptitudes pueden hacerse a partir de las tareas, solo si se presta la atención apropiada a la información que estas proveen sobre las aptitudes. En consecuencia, al determinar el propósito, también se está definiendo las tareas que deben ser seleccionadas. Uno de los aspectos importantes en cuanto al propósito es que se pueden definir dos tipos de enfoques diagnósticos de aptitudes: i) Análisis de datos de una evaluación ya existente, con la esperanza de extraer las aptitudes a la base de la misma intentando de este modo proveer información más completa que la que puede brindar un enfoque meramente unidimensional; ii) diseño de un test desde el comienzo con un enfoque diagnóstico basado en aptitudes. Muchas de las aplicaciones, sin embargo, son del primer tipo, algunas veces llamado “retroajuste”, en donde el diagnóstico de atributos se lleva a cabo mediante

análisis post hoc con propósitos de demostración y como un intento de extraer más y mejor información de una evaluación que fue diseñada con menores pretensiones.

- Descripción de un modelo para las aptitudes latentes de interés diagnóstico (en el espacio de aptitudes): La pregunta de lo que debe ser evaluado invariablemente lleva a los desempeños y los tipos de inferencia deseadas. En esta segunda etapa, se requiere una detallada representación del espacio de conocimientos, habilidades y atributos que serán medidos. Este espacio de aptitudes se desarrolla a la luz de la literatura de relevancia para el área, en medición, psicología educacional, ciencias. De acuerdo a DiBello, Roussos y Stout (2007), esta literatura no siempre lleva a representaciones adecuadas de dicho espacio, por lo que a veces se acude a representaciones múltiples antes de hacer una selección definitiva. Contingentemente, cabe indicar que la comprensión lectora invita a proponer un espacio de aptitudes desde muchas perspectivas teóricas. No hay una receta específica, pues el argumento dominante no es si las aptitudes o atributos elegidos para conformar el espacio son o no los correctos para una determinada tarea, sino si son adecuados para el propósito para el que el test ha sido creado. Quizás la única restricción real es la relativa al número de aptitudes o atributos a definir en función de su tratamiento estadístico, que impone cierto límite.
- Desarrollo y análisis de las tareas de evaluación: Al proponer tareas se debe tener presente el tipo y la cantidad de evidencia que es necesaria para sostener las inferencias que se hagan de los atributos de los examinados. Los desarrolladores de tests deben considerar una amplia variedad de posibles tareas, eligiendo aquellas que mejor se corresponden con el propósito de la evaluación. En el contexto de los CDM, no es necesario que los desarrolladores eviten tareas que requieran combinación de aptitudes o atributos por cada una de ellas. En los casos donde el test ya existe, el asunto llega a ser cuánta información puede proveer sobre un conjunto de aptitudes o atributos en particular. Si el instrumento se crea ex profeso, entonces no solamente se requiere saber las aptitudes o atributos involucrados en cada

tarea, sino también su dificultad y si hay alguna interacción entre ellas. Para indagar en torno a estos aspectos, se ha propuesto el uso de entrevistas cognitivas (Almond et al., 2009), el análisis de las características textuales de las tareas y estudios de análisis semántico latente, por ejemplo (Svetina, Gorin y Tatsuoka, 2011). Un desafío que emergerá en el proceso es el número de aptitudes o atributos que estarán representados en las tareas. Se tiene evidencia que hay un máximo tratable que no debiera sobrepasar los 10 (de la Torre y Michen, 2014). No obstante, siempre puede haber más de aptitudes que las que son estadísticamente aceptables. En dichos casos se debe llegar a un compromiso entre las aptitudes o atributos propuestos inicialmente y las que pueden ser exitosamente procesadas mediante un modelo estadístico.

Por supuesto, se debe tener en mente que puede haber atributos correlacionados con otros o bien atributos que interactúan. Uno de los aspectos más importantes en la interacción intra tarea son las referidas a: i) La interacción conjuntiva sobre una tarea, que significa que para tener un desempeño exitoso en dicha tarea se requiere cumplir con todas las aptitudes o atributos especificados para la misma; ii) la interacción compensatoria, que significa que un nivel de competencia suficiente en una aptitud o habilidad puede compensar un nivel de competencia más baja en otra aptitud o habilidad, permitiendo obtener un resultado exitoso en la tarea, a pesar de este dominio parcial de atributos. El extremo de una situación compensatoria sería el caso disyuntivo, donde hay varias aptitudes o habilidades que podrían llevar a una resolución correcta de la tarea, bastando que se domine solo una de ellas para obtener un resultado exitoso.

La elección de un modo de interacción entre atributos depende claramente del tipo de diagnóstico, los propósitos de la evaluación y de cómo estos hayan sido definidos. No es, sin embargo, el único aspecto que se debe tener presente, pues también pueden presentarse relaciones entre aptitudes o atributos que tiendan a co-ocurrir en las tareas, o bien jerarquías entre aptitudes en función de su dificultad, o el dominio de una aptitud, siempre que otras se hayan dominado previamente, o incluso aptitudes o atributos

correlacionados positivamente cuando se presume que dicha situación es útil en la elección de un modelo.

Por último, otro aspecto de esta fase es el desarrollo de reglas de codificación específicas en esta relación aptitud-tarea, que son usadas para la asignación de aptitudes a las tareas de modo que codificadores independientes puedan hacer esta atribución de modo confiable. Unas reglas de codificación adecuadas permiten realizar un buen proceso de asignación, que se ve reflejado en una matriz, referida como matriz Q, que pone en relación tareas y atributos, como será explicado más adelante.

- Especificación de un modelo psicométrico vinculando desempeño a las aptitudes latentes: Se asume en la especificación de DiBello, Roussos y Stout (2007) que los modelos psicométricos aquí mencionados corresponde a aquellos que explícitamente contienen múltiples variables que se corresponden con aptitudes o atributos a ser diagnosticados. Los modelos psicométricos a considerar pueden ser clasificados de diversas maneras, una de ellas es de acuerdo a si las aptitudes que deben ser modeladas como discretas o continuas, otra es si las aptitudes para una tarea específica son conjuntivas o bien compensatorias. Este modelo psicométrico adopta la estructura de una función matemática que especifica, como en los modelos TRI, la probabilidad de respuesta a un ítem en términos de las aptitudes o atributos del examinado y las características del ítem (sus parámetros). En estos modelos se pueden asumir ciertos supuestos que llevan a una simplificación del conjunto de aptitudes y/o de los parámetros de los ítems. Si bien esta forma de proceder puede ayudar a un mejor tratamiento estadístico de los datos mediante el modelo, también es cierto que se corre el riesgo de generar ciertos sesgos.
- Calibración y evaluación del modelo: El modelo que se selecciona debe ser calibrado (es decir, estimación tanto de los parámetros de los examinados como de los ítems). Con la calibración efectuada, se obtienen clasificaciones de los examinados en el conjunto de aptitudes. La calibración puede enfrentar problemas, pues pueden no haber suficientes ítems por aptitud o bien

demasiadas aptitudes como para obtener una calibración precisa. En consecuencia, la calibración debe llevarse a cabo con modelos simples, con mayor número de ítems y menor número de aptitudes o atributos. Otros aspectos a tener presente en esta etapa son el método de estimación y estudio de la fiabilidad, validez interna y externa del modelo.

- Desarrollo de sistemas de informes de resultados de evaluación para examinados, profesores y otros: Es la etapa que brinda algunas sugerencias sobre la forma en que se debe informar a diversos actores de interés de los resultados obtenidos por los examinados, considerando que estos deben ser fáciles de entender e interpretar, ajustándose al uso que se le dará a dicha información. Deben ir acompañados, en lo posible, de sugerencias de acción instruccional o de actividades de aprendizaje para las distintas clasificaciones de desempeño.

Los dos sistemas de desarrollo de un diagnóstico cognitivo reflejan los aspectos teóricos y empíricos que se ponen en juego para producir un patrón de aptitudes o atributos que brinde información más específica sobre los desempeños de los examinados. De esta exposición se desprenden dos elementos claves, que es necesario discutir en detalle: a) La representación de la relación entre tareas y aptitudes mediante la denominada matriz Q y b) la forma en que es necesario validar el modelo psicométrico seleccionado, para sostener la complejidad del mismo, mayor que la presente en los más tradicionales modelos unidimensionales.

## **Construcción de la matriz Q**

### ***Descripción de la matriz Q.***

Ya sea en el paso dos de la selección de un diseño (Nichols, 1994) o bien en la etapa descrita por DiBello, Roussos y Stout (2007) que implica el desarrollo o análisis de las tareas de evaluación, un aspecto fundamental de cualquier aproximación mediante modelos cognitivos es la construcción de la denominada matriz Q, también llamada matriz de incidencia (Buck, Tatsuoka y Kostin, 1997; Buck, Van Essen, Tatsuoka, Kostin, Lutz y Phelps, 1998; Chen y de la Torre, 2014; Chiu, 2013; García, Olea y de la Torre, 2014; Lee, Hunter y Lei, 2015; Lee y Sawaki, 2009a; Svetina, Gorin y Tatsuoka, 2011; Tatsuoka, 2009).

Los CDM, en particular en el ámbito de la medición educacional, consideran un dominio de evaluación que puede ser descompuesto en múltiples habilidades específicas. Estas, que a partir de este momento serán llamadas “atributos”, representan aptitudes que un examinado debe poseer para evidenciar manejo en dicho dominio. No obstante, es posible que el individuo solo domine algunos de dichos atributos, en cuyo caso se puede estar frente a diversas configuraciones de atributos en los que examinado sea o no competente. A estas configuraciones se les llama “patrón de atributos” y representan la particular conformación de fortalezas y debilidades por parte de un individuo en el dominio evaluado (Chiu, 2013).

La matriz  $Q$  es una estructura de operacionalización que pone en relación las tareas y los atributos. En el contexto psicométrico, las tareas son propuestas en ítems. En efecto, la matriz  $Q$  describe la relación entre los ítems y los atributos que serán medidos. Es, de acuerdo a Tatsuoka (2009), una matriz booleana. Sea  $q_{jk}$ , el elemento-atributo que se corresponde con la fila  $j$  y la columna  $k$  de una matriz  $Q$   $J \times K$ , donde  $J$  y  $K$  representan el número de ítems y atributos respectivamente (Chen y de la Torre, 2014). El elemento  $q_{jk}$  podrá tomar uno de dos valores (Henson y Douglas, 2005):

$$q_{jk} = \begin{cases} 1 & \text{si el ítem } j \text{ requiere el atributo } k \\ 0 & \text{en otro caso} \end{cases}$$

Es decir, si el  $k$ -ésimo atributo es requerido para responder el ítem  $j$ , entonces  $q_{jk} = 1$ , de otro modo  $q_{jk} = 0$  (García, Olea y de la Torre, 2014). La  $j$ -ésima fila de una matriz  $Q$  recibe el nombre de  $j$ -ésimo vector- $q$ , lo que brinda la especificación del atributo para el ítem  $j$ , o el patrón de atributos para ese ítem en particular. Por su parte,  $K_j^* = \sum_{k=1}^K q_{jk}$  representa el número de atributos requeridos para el ítem  $j$ . Aunque el número de atributos puede ser grande, se asume que  $K_j^*$  solo incluye los primeros atributos de relevancia para el ítem  $j$  (la restricción estadística indica entre cinco y 10 de ellos) (Chen y de la Torre, 2014; de la Torre, 2011; von Davier, 2005).

Ahora bien, los atributos que se requieren para el ítem  $j$  pueden ser expresados o denotados por el vector reducido  $\alpha_{lj}^* = (\alpha_{l1}, \dots, \alpha_{lK_j^*})'$ , donde  $l = 1, \dots, 2^{K_j^*}$ , que representa el número total de grupos o patrones de atributos latentes requeridos. Al adoptar el concepto de atributos requeridos, se puede simplificar el modelo porque el

número de patrones de atributos que son considerados para el ítem  $j$  se reduce de  $2^K$  a  $2^{K^*}$  (Chen y de la Torre, 2014). También puede comprenderse como un vector  $\alpha$  multidimensional de atributos binarios (Liu, 2015), pues se asume que estos atributos pueden adoptar dos posibles estados: a) dominar el atributo o b) no dominarlo. En efecto, la combinación de habilidades que un examinado denota se define por un patrón  $\alpha$  de atributos, y que adopta la representación de un vector de largo  $K$  latente, en el que un valor 0 representa no dominio de ese atributo  $k$  y un valor 1 representa dominio (de la Torre y Minchen, 2014). En consecuencia, en los CDM, a cada examinado se le asigna un vector  $\alpha$  o patrón de atributos binarios (domina-no domina). Por ejemplo, suponiendo que se tienen tres atributos cognitivos en los que se debe ser competente o  $K = 3$ , entonces al evaluado se le asigna un patrón o vector  $\alpha = (1, 0, 1)$ , que significa que este domina o es competente en el primer y tercer atributo, no así en el segundo. El ejemplo corresponde a uno de muchos posibles patrones de atributos, pues dado que los  $K$  atributos son binarios, entonces puede haber  $2^K$  patrones de atributos y por tanto, con tres atributos cognitivos, se podrían obtener  $2^3 = 8$  posibles patrones de atributos que también pueden ser expresados como:

$$\alpha_{ij}^* = (\{000\}, \{100\}, \{010\}, \{001\}, \{110\}, \{101\}, \{011\}, \{111\});$$

Y que pueden ser referidas como clases latentes, pues son comprendidas como categorías no observables (Huebner, 2010). En la tabla 1 se dispone un ejemplo de matriz Q que responde a lo mencionado:

Tabla 1. Ejemplo de matriz Q con 4 ítems y 3 atributos cognitivos.

Ítem	Atributo $\alpha_1$	Atributo $\alpha_2$	Atributo $\alpha_3$
1	1	1	0
2	0	1	0
3	1	1	1
4	1	0	0

En el ejemplo se podrá apreciar que el ítem 1 requiere, para su correcta resolución, que el examinado domine el atributo 1 y el 2, no siendo necesario dominar el atributo 3, pues no está presente en la tarea propuesta por el ítem 1, en consecuencia, el patrón de atributos para dicho ítem es igual al vector  $\alpha = (1, 1, 0)$ . Otros ítems requieren otros

patrones en función de los dominios en los que sea necesario ser competente para resolver correctamente la tarea propuesta.

Es importante notar que el patrón de atributos y el patrón de respuestas no son lo mismo, pero están relacionados. Así, siguiendo con el ejemplo de la tabla 1, para un examinado cuyo patrón de atributos esté representado por el vector  $\alpha = (0, 0, 0)$ , su correspondiente patrón de respuestas esperado para 4 ítems sería  $Q = (0, 0, 0, 0)$ , pues no hay ningún ítem que no requiera de, al menos, el dominio de un atributo de los que se han asignado. Otras posibles relaciones son posibles derivándolas de la matriz  $Q$  (Svetina, Gorin y Tatsuoka, 2011). No debe perderse de vista que, mientras el patrón de atributos es latente, el patrón de respuestas sí es observable.

A efectos de evaluación, la meta de un modelo diagnóstico es estimar  $\alpha$  del modo más preciso posible, o dicho de otro modo, clasificar a los estudiantes en uno de las  $2^K$  combinaciones potenciales de atributos, cada una representando una clase latente única (patrón único de dominio-no dominio). Las inferencias que se hagan sobre los estudiantes solo son válidas en la medida que la lista de atributos esté completa y los atributos estén correctamente asociados con los ítems y que, por supuesto, los ítems sean relevantes para los atributos que se pretende medir (DiBello, Roussos y Stout, 2007).

Una vez que los datos de la evaluación diagnóstica han sido analizados, será útil una revisión de la consistencia interna de la matriz  $Q$  que ha sido construida, para determinar si hay desajustes entre las dificultades que han sido observadas o lógicamente inferidas de las tareas (proporción de examinados que responden exitosamente a cada tarea) y la dificultad de los atributos que han sido asignados a las tareas o ítems. Para el caso dicotómico de dominio/no dominio de atributos, esto se refiere a la proporción estimada de la población que domina un atributo en función de la dificultad de dicho atributo. Por ejemplo, en una matriz  $Q$  no se debería asignar un atributo que supuestamente es difícil de dominar a una tarea que en realidad es fácil (DiBello, Roussos y Stout, 2007).

### ***Especificación de la matriz $Q$***

Tema aparte es el modo como se especifica la matriz de  $J$  ítems x  $K$  atributos. No hay un procedimiento automático para construirla, pues en primer lugar las tareas o ítems

que se propongan depende de los creadores del instrumento, que tienen libertad para desarrollar aquellas tareas que mejor se ajusten al modelo teórico que está a la base del desarrollo del test. Por otro lado, dado que la teoría cambia, se requiere de un elaborado juicio crítico que sea capaz de establecer: a) Qué modelo teórico responde a las tareas propuestas, b) qué atributos de dicho modelo son asignables a las distintas tareas y c) si dichas tareas responden o no a alguno de los atributos definidos dentro del modelo.

Al construir la matriz Q se debe tener presente una primera distinción, que condicionará los procedimientos de especificación. En el primer caso, es posible construir desde el inicio un instrumento centrado en los CDM. En consecuencia, su proceso de desarrollo estará mejor ajustado a la base teórica que lo sustenta y los atributos cognitivos que rigen el desempeño de los ítems. Un segundo caso es tomar un test ya construido, que no fue desarrollado teniendo en mente los CDM e intentar extraer desde él, haciendo uso de múltiples métodos, aquellos atributos que mejor explican el desempeño. La prevalencia de este proceder mediante “retroajuste” es mayor (de la Torre y Minchen, 2014). Se sabe, sin embargo, que es un método que proporcionará menos información que en el primer caso (Chen y de la Torre, 2014). Muchas veces sin embargo, no existe otra opción y aunque su uso se desaconseja, en ciertas ocasiones se le considera aceptable, como cuando se usa una estructura jerárquica de atributos y estos son asociados a distintos momentos del continuum de desempeño (Leighton, Gierl y Hunka, 2004).

Hecha la distinción, hay varios ejemplos que se pueden consultar para conocer el procedimiento de especificación concreto que se ha usado para llegar a una matriz Q que satisfaga la estructura esperada (García, Olea y de la Torre, 2014; Gierl, Leighton y Hunka, 2000; Svetina, Gorin y Tatsuoka, 2011; Tatsuoka, 2009).

En términos generales, el procedimiento implica en primer lugar asumir una teoría para el dominio general a evaluar. Por supuesto, es necesario que dicha teoría sea de carácter cognitivo o que en ella predomine esta perspectiva. En segundo lugar, evaluar aquellos aspectos cognitivos son los requeridos para cada ítem, describiendo el proceso a través del cual estos aspectos están ligados al desempeño en el ítem. Tercero, participación de jueces que puedan ayudar a desarrollar y validar la matriz Q respecto a los constructos que se supone la prueba mide (Luna, 2017).

Diferentes autores han llevado a cabo procesos específicos para construir la matriz Q y validarla a nivel de contenido. Una metodología, particularmente importante en este caso por centrarse en el dominio de la comprensión lectora, es la presentada por Buck, Tatsuoka y Kostin (1997), siendo además, una de las primeras referencias sobre el proceso de especificación aquí tratado. En primer lugar, hacen una lista de las características de los ítems que podrían afectar su desempeño, muchas de estas fueron obtenidas observando el desempeño directo de los examinados. En segundo lugar, codifican los ítems en función de esos atributos y los analizan. Después se realiza un análisis estadístico a la luz de los fundamentos teóricos y se estudia la posible interacción entre atributos. La codificación posterior fue realizada, como suele ser casi por defecto, por jueces. En este caso sin embargo, se calcularon índices de concordancia para asegurar la convergencia de opinión de los expertos participantes.

Lee y Sawaki (2009b) mencionan algunos procedimientos que enriquecen y mejoran las posibilidades de construir una matriz Q libre en lo posible de incertezas. En la definición de atributos el investigador puede guiarse por varias fuentes, como la tabla de especificaciones de los tests, las teorías sobre el dominio, análisis del contenido de los ítems o protocolos de revisión en voz alta del proceso de respuesta a los ítems, entre otros. Para estos autores, es fundamental que la definición de los atributos esté adecuadamente sostenida en investigación empírica sobre el proceso cognitivo de pensamiento que subyace a las habilidades y conocimientos que se someten a evaluación. De acuerdo a esta visión, los CDM requieren ser modelos sustentados empíricamente sobre el desempeño en la tarea para cada uno de los ítems del test. No obstante lo anterior, cuando se procede mediante “retroajuste” es posible no contar con dicho modelo de desempeño de tarea, en cuyo caso los autores sugieren acudir a “lluvia de ideas” sobre los posibles atributos que pudiesen estar tras la especificación del test y establecer qué ítems se corresponden con dichos atributos.

Es distinta la situación cuando se procede de modo exploratorio, pues en estos casos, se deben definir atributos que permitan extraer información diagnóstica útil sobre el conocimiento y las habilidades de los examinados. Aunque los autores reconocen que hay varios métodos para el desarrollo de la matriz Q, el procedimiento tradicional suele ser el análisis de contenido que hacen los jueces en virtud de su experiencia. En ocasiones, este se combina con análisis de datos, propios de modelos psicométricos de diagnóstico

cognitivo o CDPM (de su nombre en inglés: Cognitively Diagnostic Psychometric Models). El análisis de contenido implica inspección de cada ítem y múltiples jueces que los categorizan, procedimiento ya conocido. La información empírica basada en CDPM es otro tema. El propósito de este análisis es identificar una potencial especificación inadecuada de la matriz Q, debido a variaciones en las categorizaciones de jueces o bien al desajuste entre el juicio experto y el desempeño de los examinados de los ítems. Los CDPM producen información sobre el funcionamiento individual de los ítems. Datos como la dificultad y discriminación provista por este procedimiento representan información invaluable para revisar la matriz Q y reevaluar la categorización de los ítems en función de los atributos identificados. El procedimiento mediante CDPM es repetido hasta obtener resultados satisfactorios.

A pesar de la exhaustividad metodológica propuesta por Lee y Sawaki, en un trabajo empírico de los mismos autores, destinado a elaborar una prueba de escucha y lectura en inglés, utilizaron esencialmente expertos en el contenido de la prueba y en el desarrollo de matrices Q. Las habilidades presentes en los ítems fueron obtenidas del desarrollo de constructos identificados en las especificaciones del tests (Lee y Sawaki, 2009a).

Jang (2009), siguiendo un procedimiento por fases identificó nueve atributos primarios de lectura, analizando protocolos verbales de pensamiento en voz alta y ejecutando análisis estadísticos sobre los ítems y el test completo. El proceso fue personalmente realizado por el investigador quien posteriormente pidió a cinco jueces para que evaluaran los atributos de lectura que había identificado, realizando una selección y jerarquización de los atributos más importantes para responder correctamente cada uno de los ítems. El autor también dejó abierta la posibilidad de que se sugirieran nuevas habilidades o atributos.

La perspectiva de trabajo de Svetina, Gorin y Tatsuoka (2011) se centra en la génesis de la matriz Q como un proceso exploratorio e iterativo. Al inicio, los investigadores o expertos en el área a medir recopilan una lista de atributos que se cree describen la interacción del estudiante con el ítem y la forma en que este lo resuelve. Como se ha comentado, estos atributos pueden ser obtenidos desde la teoría al uso o bien mediante entrevistas cognitivas, o lisa y llanamente desde la experiencia. Después de esta

etapa informativa, los autores indican haber obtenido un listado de 22 atributos relevantes para la evaluación. Estos atributos fueron categorizados en términos de variables cognitivas relevantes para los modelos de comprensión actuales (Kintsch, 1998; van Dijk y Kintsch, 1983). Posteriormente, se codifica o asignan los atributos a los ítems, proceso que se lleva a cabo de dos maneras: a) investigadores que establecieron si era necesario dominar un atributo en particular para responder exitosamente un ítem y b) uso de los llamados procesadores automáticos de lenguaje natural, tales como el análisis semántico latente (LSA de su nombre en inglés Latent Semantic Analysis), que es una técnica analítica para extraer y representar la similitud en el significado de palabras y párrafos mediante el análisis de grandes corpus textuales, generándose un índice del grado de similitud semántica entre las proposiciones de diferentes segmentos de un texto. El proceso fue revisado por los autores del estudio para llegar a un consenso entre la codificación humana del ítem y la codificación automática.

En otro estudio (Park y Cho, 2011), se menciona que las habilidades o atributos son listados en función de las características de los ítems. Iniciando con un conjunto de atributos mayor al necesario. Posteriormente, al construir la matriz Q, se consideró: a) un criterio de identificación, pues con muchos atributos, la estimación de parámetros de los ítems no pueden ser identificados, así que por fuerza fue necesario reducir el número de atributos hasta cinco más apropiados. El segundo criterio fue b) la interpretación, pues si las habilidades son muy abstractas, el conocimiento que se puede obtener de ellas podría ser difícil de interpretar y en particular serían poco útiles para sugerir mejoras concretas una vez hecho el diagnóstico. En este estudio, dedicado a evaluación de la gramática en inglés, se pensó en mantener aquellos atributos que fuesen lo suficientemente específicos para ser interpretables y a la vez, en número adecuado.

Ravand, Barati y Widhiarso (2013) que pretendieron determinar la capacidad diagnóstica de un test de comprensión lectora en estudiantes universitarios, trabajaron sobre un instrumento que originalmente no fue pensado desde la perspectiva diagnóstica, procediendo a su análisis para montar el modelo de diagnóstico desde el “retroajuste”, extrayendo información desde el test ya creado. Dado que un modelo cognitivo detallado no estaba disponible, se usó la técnica mencionada por Lee y Sawaki (2009b), usando lluvia de ideas y posteriormente comparando juicios de cinco expertos.

Otra modalidad de especificación de la matriz Q es la ofrecida por García, Olea y de la Torre (2014), quienes optaron por consultar a cuatro expertos, a partir de un conjunto de siete atributos definidos como competencias a la base del instrumento estudiado (test de juicio situacional), sobre las competencias más adecuadas para obtener las respuestas correctas definidas para cada ítem. El proceso se llevó a cabo mediante el método Delphi, conducido en tres rondas. En la ronda inicial, cada experto trabajó individualmente. En la segunda ronda, a cada experto se le entregó el resultado anónimo del trabajo hecho por los otros jueces y se les comentó que si lo deseaban, podrían cambiar sus especificaciones iniciales. En la tercera ronda los cuatro expertos se encontraron personalmente para discutir el proceso. No fue necesario construir una única matriz Q, generando las que fueran viables y estudiando, mediante el criterio de información bayesiano (BIC de su nombre en inglés, Bayesian Information Criteria), la matriz Q de mejor ajuste. Se seleccionó la matriz con el BIC más pequeño.

Aunque en los estudios anteriores hay procedimientos más o menos específicos, lo cierto es que usualmente los trabajos que hacen alusión a modelos de diagnóstico cognitivo –en particular en comprensión de lectura- coinciden en los mecanismos de especificación de una matriz Q, siendo obligada la revisión por expertos de los atributos que están a la base del desempeño en los ítems y algún método para establecer la concordancia de estos con lo teorizado como proceso cognitivo (Chen y de la Torre, 2014; Hemmati, Baghaei y Bemani, 2016; Kim, 2015; Ravand, 2015; Sawaki, Kim y Gentile, 2009).

### **Tipologías de modelos de diagnóstico cognitivo utilizados en comprensión de lectura**

A esta altura conviene hacerse cargo de brindar una definición que permita diferenciar estos modelos de diagnóstico cognitivo de los marcos de trabajo que han sido mencionados anteriormente. En primer lugar, Zieky (2014) se hace cargo de definir modelo: “Un modelo es un término usado para referirse a una representación simplificada, comprensible y aplicable de una realidad mucho más compleja. Por ejemplo, un mapa de rutas podría ser llamado un ‘modelo de terreno’. Es altamente útil porque está muy simplificado con respecto a la realidad [...]. Un modelo permite enfocarse en los aspectos

importantes de un segmento de la realidad para un propósito particular”<sup>8</sup> (Zieky, 2014, p. 83).

Gorin (2007) por su parte, aporta la definición de diagnóstico en el término. Comprende el diagnóstico como uno de los usos más complejos de los indicados en los Standards for Educational and Psychological Testing (AERA, 2014). Para la autora, el diagnóstico se refiere “a un proceso que incluye la recolección e integración de los resultados de un test con información actual y previa sobre una persona, junto con relevantes condiciones contextuales, para identificar características de funcionamiento psicológico saludable así como de trastornos psicológicos. Los trastornos podrían evidenciarse en la información obtenida durante la medición de atributos sociales, físicos, neuropsicológicos, de personalidad, de comportamiento, adaptativos, emocionales o cognitivos, que son propios del individuo”<sup>9</sup> (p. 173) Gorin, no obstante, especifica que en el ámbito de la educación, el diagnóstico se ha enfocado en la jerarquización, clasificación, posicionamiento o resultados de evaluación y en consecuencia, la mayoría de los tests pertenecen a alguno de estos propósitos.

En función de estas definiciones, los modelos de diagnóstico cognitivo constituyen una representación simplificada de la influencia de ciertos factores de carácter eminentemente cognitivo con el objeto de clasificar y/o jerarquizar a los individuos que son evaluados mediante un test desarrollado con ese propósito. En el mismo sentido, Leighton y Gierl (2007) aunque asumen los términos, consideran necesario modificar levemente la definición al tratarse de medición educacional. La definición cognitivo-psicológica debe incorporar la noción de resolución de problemas, conocimiento y habilidades. Así, sugieren que el término de modelo cognitivo en educación se refiere a: “una descripción simplificada de la resolución de problemas humanos en tareas educacionales estandarizadas, las cuales ayudan a caracterizar el conocimiento y las

---

<sup>8</sup> “A model is a term used to refer to a simplified, understandable, usable representation of a far more complex reality. For example, a road map could be called a “terrain model.” It is highly useful because it is greatly simplified compared to reality [...]. A model allows focus on the important aspects of a segment of reality for a particular purpose.” (Zieky, 2014, p. 83)

<sup>9</sup> “Diagnosis refers to a process that includes the collection and integration of test results with prior and current information about a person, together with relevant contextual conditions, to identify characteristics of healthy psychological functioning as well as psychological disorders. Disorders may manifest themselves in information obtained during the testing of an individual’s cognitive, emotional, adaptive, behavioral, personality, neuropsychological, physical, or social attributes.” (Gorin, 2007, p. 173)

habilidades de los estudiantes en los distintos niveles de aprendizaje que han adquirido y facilitar la explicación y predicción del desempeño de los estudiantes”<sup>10</sup> (p. 6).

La confluencia de significados lleva al planteamiento, necesario por lo demás, de una definición específica de los modelos de diagnóstico cognitivo dentro de la tradición que se rescata en este trabajo. Rupp y Templin (2008) propusieron, basándose en diversos estudios, la siguiente definición:

Los modelos de clasificación diagnóstica (DCM) son modelos de variable latente, multidimensionales confirmatorios, probabilísticos, con una estructura de saturación compleja o simple. Son apropiados para modelar variables de respuesta categórica y contienen variables predictoras categóricas no observables (latentes). Las variables predictoras son combinadas en formas compensatorias y no compensatorias para generar clases latentes. Los DCM permiten múltiples interpretaciones referidas a criterio y se asocian con retroalimentación para propósitos diagnósticos, lo que es típicamente provisto con un nivel de detalle muy específico. Esta retroalimentación puede estar, aunque no necesariamente, basada en una teoría sobre procesamiento de respuesta basada en psicología cognitiva aplicada. Algunos DCM son capaces de manejar diseños de muestreo complejos para ítems y examinados, así como la heterogeneidad debida a la estrategia de uso<sup>11</sup> (p. 226).

Los aspectos de esta definición serán discutidos a continuación, no obstante, es a través de la aportación de Rupp y Templin (2008) así como la de Leighton y Gierl (2007) que surge la necesidad de prestar atención a aquellos modelos que verdaderamente representan un apoyo en el diagnóstico de inferencias sobre los procesos de pensamiento y retroalimentación hacia los estudiantes.

---

<sup>10</sup> “simplified description of human problem solving on standardized educational tasks, which helps to characterize the knowledge and skills students at different levels of learning have acquired and to facilitate the explanation and prediction of students’ performance.” (Leighton y Gierl, 2007, p. 6).

<sup>11</sup> “Diagnostic classification models (DCM) are probabilistic, confirmatory multidimensional latent-variable models with a simple or complex loading structure. They are suitable for modelling observable categorical response variables and contain unobservable (i.e., latent) categorical predictor variables. The predictor variables are combined in compensatory and noncompensatory ways to generate latent classes. DCM enable multiple criterion-referenced interpretations and associated feedback for diagnostic purposes, which is typically provided at a relatively fine-grain size. This feedback can be, but does not have to be, based on a theory of response processing grounded in applied cognitive psychology. Some DCM are further able to handle complex sampling designs for items and respondents, as well as heterogeneity due to strategy use” (Rupp y Templin, 2008, p. 226)

El problema con los modelos es que en la literatura el número es creciente y permanentemente se están llevando a cabo nuevos desarrollos (Luna, 2016). De acuerdo a Lee y Sawaki (2009), se han descrito hasta 62 modelos distintos. Según Rupp, Templin y Henson (2010), quienes adoptan una posición conservadora y no se comprometen brindando una taxonomía exhaustiva de estos modelos, limitan su listado a todos aquellos modelos que pretenden clasificar a los examinados en un número específico de clases latentes. Aunque mencionan 18 de estos modelos, probablemente es más importante denotar que estos han sido elaborados teniendo presente algunos aspectos tipológicos referidos al componente estructural del modelo, esto es, la forma en que se combinan los atributos cognitivos en clases latentes, y al componente de medida que recae principalmente en el tipo de ítems, en tanto variables observables.

### ***Criteria de clasificación general de los modelos de diagnóstico cognitivo***

Al hablar de una tipología inicial de modelos, es necesario comenzar distinguiéndolos en función de algunos criterios generales:

#### **a) Los modelos de diagnóstico cognitivo en función del tipo de relación entre atributos:**

En los modelos tradicionales basados en Teoría de Respuesta al Ítem y análisis factorial, las variables latentes se combinan mediante sumas. Lo que implica que las contribuciones de cada rasgo latente a las respuestas son aditivas. En modelos especializados de Teoría de Respuesta al Ítem, las variables o atributos latentes pueden ser combinados de otras formas, por ejemplo, mediante el producto entre estas, lo que quiere decir que las contribuciones ahora son de carácter multiplicativo (Rupp, Templin y Henson, 2010). Esta distinción puede llevarse a la diferenciación estructural de los modelos.

En efecto, bajo esta premisa surgen, por una parte, los denominados modelos de variable latente compensatoria, que ya fueron adelantados en un apartado previo, y que son más comúnmente reconocidos como modelos compensatorios. Como se mencionó, en este tipo de modelos un valor bajo en una variable o atributo latente, puede ser compensada por un alto valor en otra variable latente, de modo que no es necesario contar con un dominio en todos los rasgos para enfrentar una tarea, pues un dominio en alguno de los rasgos puede compensar las carencias en otros. Cuando un modelo psicométrico

usa sumas o diferencias, se está trabajando con atributos en forma compensatoria (Li, Hunter y Lei, 2015; Rupp, Templin y Henson, 2010).

Por su parte, en los modelos de variable latente no compensatorios, o más simplemente conocidos como modelos no compensatorios, desempeños bajos en un atributo no pueden ser compensados por un valor alto en otro atributo. Esta es la situación cuando el modelo usa la multiplicación (Li, Hunter y Lei, 2015; Rupp, Templin y Henson, 2010).

Sin embargo, no siempre producto y suma están asociados indefectiblemente con uno u otro modelo, aunque en la práctica se puede concebir pues ayuda a conceptualizar claramente la diferencia. Así, el producto o multiplicación también puede estar presente en modelos compensatorios. Ellos surgen cuando pensamos en los modelos según otra terminología que distingue entre modelos conjuntivos y disyuntivos y que responden aproximadamente a lo no compensatorio y compensatorio, respectivamente. Los primeros son aquellos en que para responder correctamente un ítem, todos los atributos para ese ítem deben haber sido dominados y ejecutados. De este modo, una estructura conjuntiva es no compensatoria, pues otros atributos no compensarán los atributos en donde no se tenga dominio. Los modelos disyuntivos en cambio, son aquellos en que si se tiene dominio sobre el atributo y se ejecutan exitosamente solo unos cuantos de ellos, será suficiente para responder correctamente el ítem. Es decir que una estructura disyuntiva de los atributos es también compensatoria. La mayoría de los CDM son modelos conjuntivos/no compensatorios (Lee y Sawaki, 2009).

Lo conjuntivo y disyuntivo son entendidas como reglas de condensación y aunque hay otras, estas son las más comunes. Para entender las implicaciones de estas reglas de condensación se puede pensar en términos determinísticos, en donde un examinado que domina un atributo  $\alpha$  recibe el valor de 1, mientras que en aquellos que no domina recibe el valor 0. Si se asume que para obtener un resultado correcto en un ítem se requiere que domine todos los atributos relacionados con su éxito, entonces  $X_i = \alpha_1 * \alpha_2 * \alpha_3, \dots, \alpha_N$ , lleva a establecer que un valor 0 (no dominio) en cualquiera de los atributos, llevará también a un valor cero el resultado final (desde una perspectiva algo más realista, llevaría a una menor probabilidad de responder correctamente) (Rupp, Templin y Henson, 2010).

En el caso de una regla de condensación disyuntiva, todos los atributos latentes  $\alpha$  que están involucrados en la determinación de un resultado exitoso  $X$  para el  $i$ -ésimo sujeto también están multiplicados entre sí, pero de un modo más flexible. Así podemos observar que  $X_i = 1 - [(1 - \alpha_1) * (1 - \alpha_2) * (1 - \alpha_3) * \dots * (1 - \alpha_N)]$ , que desde una perspectiva determinista implica que al menos uno de los atributos latentes debe ser igual a 1 para obtener una respuesta correcta  $X$  por parte del  $i$ -ésimo sujeto. Si varios, pero no todos, de los atributos tienen valor igual a 1, el resultado aún será exitoso, pero si todos tienen valor 0, entonces el resultado final será también igual a cero. Nótese que en este caso, está también presente lo multiplicativo, pero no exclusivamente (Rupp, Templin y Henson, 2010).

**b) Los modelos de diagnóstico cognitivo en función de la escala de la variable respuesta:**

Es posible trabajar con dos de los formatos más comunes en instrumentos de carácter evaluativo. Por un lado, se tienen los ítems de carácter dicotómico, que son aquellos que tienen dos opciones de respuesta. Es común, sin embargo, que los ítems de los tests tengan más de dos opciones de respuesta, pero en virtud de la definición de respuesta correcta e incorrecta, es posible transformar los ítems de opción múltiple considerando la respuesta correcta con valor 1, mientras que todas las respuestas alternativas incorrectas se reducen a valor 0. Por su parte, los ítems politómicos, que poseen más de dos alternativas de respuesta también tienen modelos dedicados considerando dicha variabilidad, sin reducir el resultado a dicotomías (Martínez-Arias et al., 2006; Moreno, Martínez y Muñiz, 2006).

Hay que destacar, sin embargo, que en realidad todos los DCMs son apropiados para ítems dicotómicos (Rupp, Templin y Henson, 2010). No obstante, modelos específicos para ítems politómicos han estado también desarrollándose, pues se han obtenido evidencias de que son más adecuados para propósitos de diagnóstico (Ma y de la Torre, 2016).

**c) Los modelos de diagnóstico cognitivo en función de la escala de los atributos latentes:**

También es posible organizar los CDMs de acuerdo a la forma en que estos atributos son definidos. La forma más simple corresponde a dominio/no dominio, por lo que los atributos latentes pueden ser concebidos en su naturaleza dicotómica. En estos

casos, cada respuesta al ítem dependerá de un conjunto de atributos, pero en donde el examinado debe tener o no dominio de dicho atributo. Es una aproximación de todo o nada a nivel de atributo, aun cuando en la combinación de estos se asuma una estructura compensatoria o no compensatoria (Lee y Sawaki, 2009).

Los CDMs que incorporan atributos dicotómicos son de menor complejidad que aquellos que asumen una estructura politómica. En este último caso se consideran atributos específicos pero en los que se obtienen niveles de dominio, de modo que pueden construirse ítems específicos para medir niveles específicos de un atributo (Chen y de la Torre, 2013). Los CDMs politómicos van más allá que lo descrito por la multidimensionalidad, pues esta característica se refiere a que puede haber más de un atributo que condicione el desempeño en los ítems. En este caso, se agrega un nivel mayor de dificultad pues, además de múltiples atributos o variables latentes, cada una de ellas puede tener más de dos niveles. Sin embargo, en número de parámetros en estos casos tiende a incrementarse y se requieren estrategias de análisis específicas para afrontar la complejidad derivada de modelos como estos. Sin embargo, en varias aplicaciones, se sigue definiendo el modelo en función de la matriz  $Q$ , solo que esta ahora es de naturaleza politómica (Chen y de la Torre, 2013; Sun, Xin, Zhang y de la Torre, 2013; von Davier, 2005; von Davier, DiBello y Yamamoto, 2006).

### ***Tipos de modelos de diagnóstico cognitivo aplicados en comprensión de lectura***

Dada la gran cantidad de modelos disponibles y descritos en la literatura (Lee y Sawaki, 2009; Ravand, 2016; Rupp y Templin, 2008; Rupp, Templin y Henson, 2010) en este trabajo la presentación de modelos específicos se limita solo a aquellos que hayan sido utilizados con cierta eficacia en el diagnóstico de la comprensión lectora en estudiantes de nivel elemental.

Li, Hunter y Lei (2015) realizaron un estudio que describió aquellos modelos que habían sido utilizados en el contexto de la comprensión de lectura exclusivamente. De acuerdo a estos autores, una decisión de importancia corresponde a la elección de un modelo compensatorio o no compensatorio, también llamado conjuntivo. Estos últimos han sido mucho más comunes en comprensión de lectura (Buck, Tatsuoka y Kostin, 1997; Buck, Van Essen, Tatsuoka, Kostin, Lutz y Phelps, 1998; Hemmati, Baghaei y Bemani,

2016; Jang, 2009; Ravand, Barati y Widhiarso, 2013; Ravand, 2016; von Davier, 2005). Algunos conocidos modelos de tipo no compensatorio son el Rule Space Model (Svetina, Gorin y Tatsuoka, 2011; Tatsuoka, 2009), el Attributes Hierarchy Model o AHM (Wang y Gierl, 2011), el modelo Deterministic-Input, Noisy-And-Gate, DINA, el modelo Reparameterized Unified Model (RUM) o su sutil derivación, el Fusion Model (Jang, 2009) y el modelo Reduced Reparameterized Unified Model, o RRUM (Jang, 2009; Jang, Dunlop, Wagner, Kim y Gu, 2013; Li, Hunter y Lei, 2015), también reconocido como Noncompesatory Reparameterized Unified Model, NC-RUM (Ravand y Robitzsch, 2015). Por su parte, un conocido modelo compensatorio es el Deterministic-Input, Noisy-OR-Gate, DINO (Liao, Kuo y Deenang, 2015; Park y Cho, 2017), mientras que los modelos generales o saturados están representados por el denominado Generalized Deterministic-Input, Noisy-And-Gate Model, G-DINA (Chen y Chen, 2016; Hemmati, Baghaei y Bemani, 2016; Ma y Meng, 2014; Ravand, 2016), que agrupa los modelos DINA, DINO y RRUM y el General Diagnostic Model o GDM (von Davier, 2005),

No está muy claro que los modelos deben ser no compensatorios, como los que se han mencionado, o bien compensatorios, como proponen otros autores tal como Stanovich o Bernhardt (Li, Hunter y Lei, 2015). Usó-Juan (2006), por ejemplo, en un trabajo orientado a determinar la naturaleza compensatoria de la comprensión de lectura, determinó que la carencia de conocimiento sobre el tema tratado, podría verse compensado por el dominio de la lengua inglesa.

El asunto es que la elección de una u otra estructura de relación entre atributos depende del modelo teórico sobre comprensión al uso y en particular, del estado de avance en el área. De acuerdo a un modelo de lectura, llamado Visión Simple de Lectura, propuesto por Gough (Dreyer y Katz, 1992) y que rescata un esquema sencillo de comprensión, la lectura es fruto de la comprensión y la decodificación. Lo primero refiere a la interpretación de discurso desde la información léxica y lo segundo corresponde a la representación de la información impresa, reconociendo y pronunciando las palabras. Ambos componentes son igualmente relevantes, apoyando la visión no compensatoria de la lectura. En consecuencia, de acuerdo a Li, Hunter y Lei (2015) “hay evidencia teórica

mixta respecto a si la comprensión lectora es compensatoria o no compensatoria, y no es claro qué CDMs deberían ser usados para las pruebas de comprensión lectora”<sup>12</sup> (p. 3).

En trabajos más recientes, donde se han probado modelos compensatorios y no compensatorios, el resultado no es mucho mejor. Se ha señalado que los modelos no compensatorios tendieron a sobreestimar la probabilidad de dominio de los atributos, mientras que los compensatorios los subestimaron. En buena medida, el mejor o peor ajuste dependió de los atributos que se pusieron en juego en los modelos (Park y Cho, 2017).

Dada esta situación, parece lógico presentar los modelos más generales, también llamados modelos saturados así como también los modelos más específicos. A los primeros pertenecen el modelo GDM (von Davier, 2005), los modelos log-lineales (Rupp, Templin y Henson, 2010) y el modelo G-DINA (Chen y Chen, 2016; Ravand, 2016) y a los segundos modelos derivados como DINA (Ravand, Barati y Widhiarso, 2013), DINO (Li, Hunter y Lei, 2015) y RRUM (Jang, 2009; Jang, Dunlop, Wagner, Kim y Gu, 2013; Li, Hunter y Lei, 2015). Además de esta condición, se restringe la presentación a aquellos modelos que han sido aplicados en la evaluación de la comprensión lectora, a saber DINA, DINO, RRUM, G-DINA y GDM (Hemmati, Baghaei y Bemani, 2016; Ravand, Barati y Widhiarso, 2013; Ravand, 2016). En consecuencia, de acuerdo a estos criterios, los modelos a discutir se presentan a continuación.

**a) El modelo Deterministic-Input, Noisy-And-Gate DINA:**

Este modelo, traducido por Luna (2017) como modelo de insumos deterministas y salidas aleatorias, es de carácter conjuntivo o no compensatorio. Sea  $X_{ij}$  la respuesta de un sujeto  $i$  a un ítem  $j$ , donde  $i = 1, 2, \dots, I$ , y donde  $j = 1, 2, \dots, J$ , y sea  $\alpha_i = \{\alpha_{ik}\}$  el vector de atributos del examinado. En tanto,  $k = 1, 2, \dots, K$ , donde un valor igual a 1 en el  $k$ -ésimo elemento del vector denota el dominio sobre el atributo o habilidad, y el valor 0 ausencia o no dominio de dicho atributo. Como otros CDMs, este modelo requiere la construcción de una matriz  $Q$  como la que ha sido presentada, con una estructura  $J \times K$  o matriz de ítems por atributos formada exclusivamente por ceros y unos, donde el elemento que se encuentra en el cruce de la  $j$ -ésima fila con la  $k$ -ésima columna de la matriz es  $q_{jk}$ ,

---

<sup>12</sup> “there is mixed theoretical evidence regarding whether reading comprehension is compensatory or non-compensatory, and it is not clear what CDMs should be used for reading comprehension tests.” (Li, Hunter y Lei, 2015, p. 3).

que indica si el atributo  $k$  es requerido para responder correctamente el ítem  $j$ . En el modelo DINA, el vector de atributos del examinado y la matriz  $Q$  produce un vector de respuesta latente que se denota como  $\eta_i = \{\eta_{ij}\}$ , donde

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}. \quad (1)$$

La ecuación 1, referida a la respuesta latente, asume un valor igual a 1 si el examinado  $i$  posee la totalidad de los atributos requeridos para responder el ítem  $j$  y asume el valor 0 si el evaluado carece de al menos uno de los atributos. En efecto, en el modelo, el término “y” (AND) de su nombre, refiere al proceso conjuntivo en la determinación de  $\eta_{ij}$  que implica el dominio de todos los atributos requeridos por un ítem en particular, de allí su carácter no compensatorio. El dominio de otras habilidades no reemplazará las carencias de dominio en los atributos que son claves. Este proceder permite conformar dos grupos distintos por ítem, aquellos que dominan todos los atributos para responder correctamente ese ítem y aquellos que no dominan al menos uno de los atributos y por tanto se espera una respuesta incorrecta en el ítem (de la Torre, 2009; Romero, 2010). Ambos grupos se conectan con dos parámetros, que forman parte del componente “noise” de la designación del modelo: a) desliz o distracción, en que el examinado del grupo que posee todos los atributos para un ítem puede distraerse y fallar y b) conjetura, donde el grupo de los examinados que carecen de al menos uno de los atributos, puede suponer o adivinar la respuesta correcta, lo que conlleva a probabilidades de respuesta distintas a cero. En este modelo, estos parámetros del ítem  $j$  pueden expresarse como  $s_j = P(X_{ij} = 0 \mid \eta_{ij} = 1)$  (parámetro desliz o distracción) y como  $g_j = P(X_{ij} = 1 \mid \eta_{ij} = 0)$  (parámetro conjetura) (de la Torre, 2009). De manera que la probabilidad de que un examinado  $i$  con el vector de atributos  $\alpha_i$  responda correctamente el ítem  $j$  está dada por

$$P_j(\alpha_i) = P(X_{ij} = 1 \mid \alpha_i) = g^{1-\eta_{ij}}(1 - s_i)^{\eta_{ij}} \quad (2)$$

El modelo DINA tiene entonces entradas determinísticas y salidas probabilísticas, donde el componente determinístico reside en la variable latente que asume el valor 1

cuando todos los atributos requeridos son dominados por el evaluado y el valor 0 cuando no es así. Con cada ítem se estiman los parámetros  $s_j$  y  $g_j$  y para cada examinado se estiman  $k$  variables latentes, que conforman el vector de dominio de atributos de los examinados  $\alpha_{ik}$  (Romero, 2010).

El modelo se suele estimar mediante máxima verosimilitud conjunta (JML de su nombre en inglés) dado que DINA es una distribución condicional conjunta de  $X_{ij}$  en función de un vector de atributos  $\alpha_i$ . Sin embargo, existen aproximaciones mediante estimaciones de máxima verosimilitud marginalizada, que puede ser estimada mediante el algoritmo EM. También se ha informado de su estimación mediante el método Monte Carlo Cadenas de Markov (MCMC) y se ha aportado evidencia a favor de estimaciones fiables, pero en particular en los modelos DINA de orden superior (HO-DINA) (de la Torre, 2009).

**b) El modelo Deterministic-Input, Noisy-OR-Gate DINO:**

Representa un modelo compensatorio, análogo al modelo DINA y también incluye los parámetros de desliz/distracción y conjetura, que son modelados a nivel de ítem. En orden a obtener la expresión que representa al modelo DINO, se requieren tres elementos. El primer elemento es la variable latente de respuesta  $\eta_{ij}$ , aquellos individuos que dominan o satisfacen todos los atributos entonces  $\eta_{ij} = 1$ , teniendo una probabilidad de repuesta positiva igual a  $1 - s_j$ , mientras que aquellos que carecen de dominio sobre alguno de los atributos –al menos uno- esto es,  $\eta_{ij} = 0$  tienen una probabilidad de respuesta igual a  $g_j$ . En cuanto al dominio, el modelo DINO está definido de modo similar a DINA, la diferencia reside en que los grupos o clases de equivalencia probabilísticos determinados por  $\eta_{ij}$  se basan ahora en la regla disyuntiva, de modo que  $\eta_{ij}$  se redefine usando la notación  $\omega_{ij}$ , donde:

$$\omega_{ij} = 1 - \prod_{k=1}^K (1 - \alpha_{ik})^{q_{jk}} \quad (3)$$

Donde  $\omega_{ij}$  divide a los examinados en dos grupos, un grupo donde los sujetos satisfacen al menos uno de los atributos requeridos presentes en la matriz Q ( $\omega_{ij} = 1$ ) y un grupo de individuos que no satisfacen ninguno de los atributos de los presentes en

dicha matriz ( $\omega_{ij} = 0$ ). De acuerdo a la definición de  $\omega_{ij}$ , el modelo DINO define la probabilidad de obtener una respuesta correcta de un modo semejante al modelo DINA:

$$P(X_{ij} = 1|\omega_{ij}) = g_j^{1-\omega_{ij}}(1 - s_i)^{\omega_{ij}} \quad (4)$$

Donde  $1 - s_j > g_j$ . Como en el modelo DINA,  $s_j$  y  $g_j$  representan el parámetro deslíz y el parámetro conjetura, respectivamente.

**c) El modelo Reduced Reparameterized Unified Model (RRUM):**

El modelo RUM (Jang, 2009; Jang, Dunlop, Wagner, Kim y Gu, 2013; Kim, 2011; Li, Hunter y Lei, 2015; Templin, Henson, Templin y Roussos, 2008) pertenece a la familia de los modelos de Teoría de Respuesta al Ítem multidimensionales. Se aplica con el fin de determinar los niveles de dominio de un examinado sobre habilidades o atributos específicos, vinculando probabilidades de respuesta al ítem a dicho conjunto de atributos. Como en otros modelos esta asociación entre ítems y atributos se establece en función de la matriz Q. En la representación de esta matriz, los ítems  $j = 1, \dots, J$  se definen asociados con los atributos o habilidades  $k = 1, \dots, K$ , con  $q_{jk} = 1$  indicando que se requiere dominar el atributo  $k$  para responder el ítem  $j$ , mientras que  $q_{jk} = 0$  indica que no se requiere dominar el atributo  $k$  para responder el ítem  $j$ . De modo que los parámetros relativos a la habilidad del examinado  $i$  son modelados como  $\alpha_{ik} = 1$  si el evaluado  $i$  domina el atributo  $k$ , y  $\alpha_{ik} = 0$  en otro caso (Templin, Henson, Templin y Roussos, 2008).

Este modelo asume una relación conjuntiva entre los atributos que se requieren para resolver un ítem, es decir, para responder correctamente el examinado debe dominar todos los atributos cognitivos. En el modelo RRUM la probabilidad de responder correctamente un ítem se expresa como:

$$P(X_{ij} = 1|\alpha_i) = \pi_j^* \prod_{k=1}^K r_{jk}^{*(1-\alpha_{ik})q_{jk}} \quad (5)$$

Donde  $P(X_{ij} = 1|\alpha_i)$  es la probabilidad del sujeto  $i$  de responder correctamente al ítem  $j$  dado el dominio de su patrón de atributos  $\alpha_i$ . Por su parte,  $\pi_j^*$  es la probabilidad de aplicar correctamente todos los atributos especificados sobre el ítem  $j$  de acuerdo a la

identificación ítem-atributo establecida por medio de la matriz  $Q$  ( $q_{j\cdot} = \sum_{k=1}^K q_{jk}$ ). El parámetro  $r_{jk}^*$  puede interpretarse como la penalización impuesta sobre  $\pi_j^*$  cuando no se domina un atributo  $k$ , con  $r^*$  definido en el intervalo  $[0,1]$ . De modo que el valor de la función de respuesta al ítem dada por el modelo RUM es un producto multiplicativo de  $\pi^*$  y, a cada uno de los atributos especificados en la matriz  $Q$  que el examinado no domina, le acompaña el parámetro  $r^*$ . RUM provee la habilidad para discriminar entre patrones de atributos en niveles que no son posibles con modelos como DINA. La parametrización de RUM, especialmente los parámetros  $r^*$ , permiten que se definan probabilidades adicionales de respuesta al ítem (considere, por ejemplo, que DINA solo provee probabilidades de respuesta al ítem solo para el grupo que domina todos los atributos definidas en la matriz  $Q$  y para el grupo que no domina al menos uno de dichos atributos) (Templin, Henson, Templin y Roussos, 2008).

**d) El modelo Generalized Deterministic Inputs, Noisy and Gate (G-DINA):**

Se trata de una generalización del modelo DINA, pero con supuestos más flexibles. G-DINA se considera un modelo saturado y aunque es menos restrictivo, esta clase de modelos son más complejos y requieren muestras más grandes si se desea obtener estimaciones precisas de los parámetros (Rojas, de la Torre y Olea, 2012).

El modelo DINA es quizás uno de los más simples modelos propuestos en CDM cuando los ítems se puntúan de modo dicotómico. Como se ha presentado, DINA solo requiere dos parámetros por cada ítem sin perjuicio del número de atributos requeridos para el ítem. Su aplicación es pertinente cuando la tarea requiere un conjunto de atributos igualmente importantes, de modo que carecer de dominio en un atributo en particular equivale a no tener dominio en ninguno de los atributos requeridos (de la Torre, 2011).

En el modelo DINA, en un test compuesto de  $J$  ítems que miden  $K$  atributos, se dividen los  $2^K$  vectores de atributos latentes o clases, en dos grupos. En función de los requerimientos del dominio de atributos para un ítem  $j$ , se establece de modo determinista la clasificación de cada uno de los  $2^K$  vectores de atributos en uno de estos dos grupos,  $\eta_{j1}$  que corresponde a la agrupación de los vectores de atributos que cumplen con todos los atributos requeridos para el ítem  $j$ , y  $\eta_{j0}$ , que es el grupo que reúne los vectores de atributos que carecen de al menos uno de los atributos requeridos para responder correctamente el mismo ítem (de la Torre, 2011).

Con el modelo DINA, los vectores de atributo que han sido reunidos en el mismo grupo se asume que tienen la misma probabilidad de responder correctamente el ítem. Pero este supuesto no necesariamente se mantiene para el grupo  $\eta_{j0}$  pues debido a que los vectores de atributos que lo conforman pueden llegar a tener grados variables de deficiencia con respecto a los atributos requeridos para dar una respuesta correcta, sus probabilidades de éxito podrían no ser idénticas. La generalización del modelo DINA, o G-DINA relaja el supuesto de igual probabilidad de éxito para todos los vectores de atributos en  $\eta_{j0}$  (de la Torre, 2011).

Para cada uno de los ítems, el modelo G-DINA divide las clases latentes en  $2^{K_j^*}$  grupos latentes, donde  $K_j^* = \sum_{k=1}^K q_{jk}$  representa el número de atributos requeridos para el ítem  $j$ . Se define  $\alpha_{lj}^*$  como el vector de atributos reducido cuyos elementos son los atributos requeridos para el ítem donde, como se ha mencionado anteriormente  $l = 1, \dots, 2^{K_j^*}$ , que representa el número total de grupos latentes o patrones de atributos latentes requeridos<sup>13</sup>. En el modelo G-DINA, la probabilidad de que un examinado perteneciente a la clase latente  $l$ , con un patrón de atributos  $\alpha_{lj}^*$  responda correctamente un ítem  $j$  se denota como  $P(X_j = 1 | \alpha_{lj}^*) = P(\alpha_{lj}^*)$  (García, Olea y de la Torre, 2014). La formulación original del modelo G-DINA basada en  $P(\alpha_{lj}^*)$  puede ser descompuesta en la suma de los efectos debido a la presencia de atributos específicos y sus interacciones, así:

$$P(\alpha_{lj}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'} \alpha_{lk} \alpha_{lk'} \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk} \quad (6)$$

Donde  $\delta_{j0}$  es el intercepto para el ítem  $j$ ,  $\delta_{jk}$  es el efecto principal debido a  $\alpha_k$ ,  $\delta_{jkk'}$  es el efecto de interacción debido a  $\alpha_k$  y  $\alpha_{k'}$  y  $\delta_{j12\dots K_j^*}$  es el efecto de interacción debido a  $\alpha_1 \dots \alpha_{K_j^*}$ . De los parámetros expuestos, se puede señalar que  $\delta_0$  es la probabilidad basal o probabilidad de respuesta correcta cuando ninguno de los atributos está presente,  $\delta_k$  es el cambio en la probabilidad de una respuesta correcta como resultado de dominar un único atributo (por ejemplo  $\alpha_k$ ),  $\delta_{kk'}$  es un efecto de interacción de primer

<sup>13</sup> La explicación detallada de estas expresiones pueden revisarse en el apartado previo de construcción de la matriz Q.

orden y representa el cambio en la probabilidad de una respuesta correcta debido al dominio tanto en  $\alpha_k$  y  $\alpha_{k'}$  que está más allá del impacto aditivo de dominio de ambos atributos, y  $\delta_{12\dots K_j^*}$  representa el cambio en la probabilidad de una respuesta correcta debido al dominio de todos los atributos requeridos más allá del impacto aditivo de los efectos principales y de interacción. El intercepto siempre es no negativo, los efectos principales son usualmente no negativos, pero los efectos de interacción pueden tomar cualquier valor. Cabe notar que los efectos principales son no negativos si  $P(\mathbf{0}_{K_j^*}) \leq P(\alpha_{ij}^*)$  para  $\sum_{k=1}^{K_j^*} \alpha_{ijk}^* = 1$ , donde  $\mathbf{0}_{K_j^*}$  es el vector nulo de largo  $K_j^*$ . Esto implica que tener dominio sobre cualquiera de los atributos cognitivos requeridos para un determinado ítem representa un incremento en la probabilidad individual de obtener un resultado exitoso en el ítem (de la Torre, 2011).

**e) El modelo Generalized Deterministic-Input, Noisy-And-Gate Model GDM:**

El modelo Generalized Deterministic-Input, Noisy-And-Gate Model es una propuesta de von Davier (2005) que puede ser aplicado a ítems de carácter politómico al igual que a atributos que tienen más de dos niveles de desempeño. Además, son modelos que permiten el manejo de respuesta omitida, sin necesidad de hacer ajustes a los datos. En realidad es una aproximación generalista y más amplia que permite especificaciones o versiones de otros modelos.

Estos modelos pueden ser de naturaleza compensatoria y no compensatoria. Dentro de los modelos GDM puede haber varias versiones de los mismos, desde estimaciones a partir de modelos Rasch, modelos de crédito parcial generalizado, modelos logísticos de dos parámetros (2PL) o incluso el modelo FACETS de Muraki (von Davier, 2005).

Como los anteriores modelos, el modelo GDM inicia a partir de un conjunto de atributos o habilidades propias de los examinados que permite generar una matriz Q como la que se ha explicado en apartados previos. Sin embargo, el modelo (o más bien, modelos) GDM permite versiones más generales de la matriz Q, representando una aproximación de carácter más amplio en la especificación de la interacción entre el patrón de atributos y dicha matriz.

Dado que es posible obtener varias versiones del GDM, a continuación se presenta una forma general de expresión de esta clase de modelos para posteriormente mostrar y comentar el modelo en términos logísticos.

La formalización de von Davier (2005) inicia describiendo el modelo, en forma loglineal, como discreto, multidimensional, con variable latente  $\theta$ , donde  $\theta = (\alpha_1, \dots, \alpha_K)$ , con  $K$  niveles discretos de atributos. En el caso más simple y más habitual, los atributos son dicotómicos, tomando por tanto dos valores, de modo tal que  $\alpha_K \in \{0, 1\}$ . En un caso así, los atributos son interpretados bajo la lógica del dominio-no dominio o uno y cero en el atributo  $K$ , respectivamente. Siendo  $\theta = (\alpha_1, \dots, \alpha_K)$  un patrón de atributos  $K$ -dimensional consistente de niveles de atributos politómicos  $\alpha_k$  donde  $k = 1, \dots, K$ . Se puede entonces definir el modelo como:

$$\log \left[ \frac{P(X = x | \beta_i, q_i, \gamma_i, \alpha)}{P(X = 0 | \beta_i, q_i, \gamma_i, \alpha)} \right] = \beta_{xi} + \gamma_{xi}^T \cdot h(q_i, \alpha) \quad (7)$$

Con elementos  $q_i$  de la matriz  $Q$  donde  $q_i = (q_{i1}, \dots, q_{iK})$  y  $q_{ik} \in \{0, 1, 2, \dots\}$  para  $k = 1, \dots, K$ . Además, hay parámetros de dificultad  $\beta_{ix}$  con valoración real y un parámetro pendiente  $k$ -dimensional  $\gamma_{xi} = (\gamma_{xi1}, \dots, \gamma_{xiK})$  para cada categoría  $x$  de respuesta distinta de cero, donde  $x \in \{1, 2, \dots, m_i\}$ . El modelo descompone la probabilidad condicional de una respuesta  $x$  en el ítem  $i$  en dos sumandos, la dificultad general  $\beta_{xi}$ , y una combinación lineal del nivel de atributo por los términos de la matriz  $Q$   $h(q_i, \alpha) = (h_1(q_i, \alpha), \dots, h_k(q_i, \alpha))$ . Dado un elemento, distinto a cero, de la matriz  $Q$ , la pendiente  $\gamma_{ix}$ . en la expresión lineal anterior determina cuanto los componentes de un atributo particular en  $\alpha = (\alpha_1, \dots, \alpha_K)$  contribuye a las probabilidades de respuesta en un ítem  $i$ .

Como se sabe, elementos de la matriz  $Q$ ,  $q_{ik}$ , relacionan el ítem  $i$  al atributo  $k$  y determinan si dicho atributo se requiere para ese ítem  $i$ . Si un atributo  $k$  se requiere, entonces  $q_{ik} > 0$ . Si en cambio,  $k$  no es requerido, entonces  $q_{ik} = 0$ . En consecuencia, esto implica que si el atributo  $k$  se define como no requerido para el ítem  $i$  en la matriz  $Q$ , esto es  $q_{ik}=0$ , entonces el nivel de atributo  $\alpha_k$  no contribuye en lo absoluto a las probabilidades de respuesta para ese ítem (von Davier, 2005).

De lo expuesto, se puede señalar que  $h_i(q_{i.}, \alpha) \mapsto \mathbb{R}$  son los bloques fundamentales en la construcción de los GDM. La función  $h_i$  asigna los niveles del atributo  $\alpha_k$  y los elementos  $q_{ik}$  de la matriz  $Q$  a los números reales (von Davier, DiBello y Yamamoto, 2006). La asignación de  $h$  proyecta los niveles de atributo  $\alpha_k \alpha = (\alpha_1, \dots, \alpha_K)$  usando los elementos  $q_{i.}$ . En la mayoría de los casos, la misma proyección será utilizada para todos los ítems (de ahí que realmente no sea necesario incluir el subíndice  $i$ ). Las  $h$  son asignaciones que especifican cómo los elementos de la matriz  $Q$  determinan el impacto del patrón de atributos sobre las probabilidades de respuesta condicional  $P(X = x | \beta_i, q_i, \gamma_i, \alpha)$  (von Davier, 2005).

La formulación log-lineal de von Davier (2005) puede ser expresada en términos logísticos. El modelo anterior es equivalente a:

$$P(X = x | \beta_i, q_i, \gamma_i, \alpha) = \frac{\exp[\beta_{xi} + \gamma_{xi.}^T h(q_i, \alpha)]}{1 + \sum_{y=1}^{m_i} \exp[\beta_{yi} + \gamma_{yi.}^T h(q_i, \alpha)]} \quad (8)$$

Donde el modelo incorpora un patrón de atributos  $k$ -dimensional  $\alpha = (\alpha_1, \dots, \alpha_K)$  y algunas necesarias restricciones sobre  $\sum_k \gamma_{xik}$  y sobre  $\sum \beta_{xi}$  para poder identificar el modelo. Usando esta reformulación y además especificando la asignación de  $h$ , se muestra que los GDMs pueden contener modelos de TRI habituales así como otros modelos de carácter cognitivo compensatorio, entre ellos el Fusion Model (Jang, 2009), razón por la que este último no se dispone aquí en detalle. Todos ellos constituyen casos especiales de GDM. En la misma lógica, ahora los parámetros  $\beta_{xi}$  y  $\gamma_{xik}$  puede ser interpretados como los parámetros dificultad y discriminación, respectivamente (von Davier, 2005; von Davier, DiBello y Yamamoto, 2006). Por último, señalar que los modelos GDM utilizan para su estimación una aproximación basada en máxima verosimilitud, misma que está implementada en algunos de los escasos programas informáticos para su estimación (von Davier, 2005).

Estos modelos cognitivos, presentados con algún detalle, representan los más comúnmente aplicados en el ámbito de la comprensión lectora y permiten avanzar en el conocimiento detallado de este dominio, más allá que la sola puntuación que permita establecer una jerarquía de desempeño general.

En lo que sigue, además de determinar, en primera instancia, la presencia de diferencias entre los establecimientos públicos y privados en el dominio de comprensión lectora, también se estudia el patrón de atributos cognitivos que se ve comprometido en función de estas diferencias. La pretensión última lleva al vínculo de variables sociales con variables de índole cognitivo, con un detalle mayor, en términos diagnósticos, al que suele representar la medida de desempeño tradicional.

## **PREGUNTA DE INVESTIGACIÓN**

En función de los antecedentes teóricos expuestos en la introducción, se configuran las pretensiones de este estudio que en términos simples, pueden ser comprendidos como la determinación de diferencias en el grado de comprensión lectora de estudiantes de centros públicos y privados y el papel que juega el patrón de atributos cognitivos en el dominio de la comprensión de lectura y en la generación de diferencias en el desempeño lector entre los estudiantes de ambas dependencias.

En consecuencia, la pregunta de investigación que guía este estudio es: ¿Cuáles son las diferencias en el dominio de la comprensión de lectura entre los centros educativos privados y públicos de la zona central de Chile y cuáles son los atributos cognitivos que mejor explican dicha diferencia?

## **OBJETIVOS DE INVESTIGACIÓN**

Con base en estas preguntas de investigación, los principales objetivos son:

1. Evaluar la influencia de variables contextuales y económicas sobre el desempeño académico en comprensión de lectura entre estudiantes de escuelas públicas y privadas.
2. Establecer, a través de la aplicación de un modelo de diagnóstico cognitivo, el patrón de atributos cognitivos que explican la ejecución de tareas de comprensión lectora en una evaluación efectuada sobre estudiantes de 8 a 10 años de edad (correspondientes a segundo y tercer ciclo de escuela primaria) que pertenecen a la educación pública y privada.
3. Determinar las diferencias en el patrón de atributos cognitivos propios de la comprensión lectora en centro de educación pública y privada.

## **PRESENTACIÓN DE LOS ESTUDIOS**

Los tres objetivos mencionados cubren las actividades de investigación principales que desarrolla este estudio. Mediante el primer objetivo se pretende justificar la importancia de considerar las diferencias sociales y económicas entre escuelas públicas y privadas y su efecto en una de las habilidades a la base de cualquier aprendizaje, como ocurre con la comprensión de lectura.

El segundo objetivo intenta determinar los atributos cognitivos que explican el desempeño en comprensión lectora, mediante el estudio por retroajuste de una de las herramientas psicométricas más populares a nivel nacional. Se rescata un patrón de habilidades, todos aquellos atributos cognitivos específicos que predicen un buen desempeño en dicha variable. A través de la aplicación del modelo cognitivo de mejor ajuste a los ítems de una prueba de comprensión, se obtienen patrones de atributos o también llamados estados de conocimiento que explican el desempeño del estudiante en la evaluación. Una vez extraídos los atributos y clasificados los estudiantes conforme a su propio patrón de dominio de atributos o “estado de conocimiento”, se realiza un diagnóstico sobre las falencias o fortalezas en el desempeño de los estudiantes, aportando además a la validez de constructo del instrumento.

El tercer objetivo se alcanza con el estudio de las diferencias entre centros públicos y privados al nivel de los atributos cognitivos comprometidos en cada caso, no sin antes efectuar una evaluación del funcionamiento diferencial del ítem, considerando como grupo focal, los centros públicos y como grupo de referencia los centros privados. Una vez determinada la presencia/ausencia de DIF, se procedió a determinar la presencia o no de las diferencias entre atributos cognitivos individuales entre los centros comparados.

A continuación se presentan los estudios elaborados para dar cuenta de cada una de estas finalidades.

## **Estudio 1: El efecto de factores contextuales y económicos sobre el grado de comprensión en lectura**

### **Introducción**

El éxito académico del alumnado ha acaparado una parte importante de la atención tanto de instituciones internacionales como nacionales como lo atestiguan diversos estudios que desarrollan estudios comparativos. Es el caso de las investigaciones elaboradas por Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la UNESCO (Flotts, Manzi, Jiménez, Abarzúa, Cayuman y García, 2015), el International Association for the Evaluation of Educational Achievement y su estudio Trends in International Mathematics and Science Study TIMSS (Agencia de Calidad de la Educación, 2011) o el muy conocido Programme for International Student Assessment (PISA) de la Organisation for Economic Cooperation and Development (OECD, 2014).

En Chile se ha hecho lo propio desde los primitivos intentos de evaluación en 1982 con el antiguo Programa de Evaluación del Rendimiento (PER), que evaluaba a niños de cuarto y octavo básico, con la pretensión de trasladar la iniciativa de mejoramiento a los docentes y apoderados (Bravo, 2011). El sistema evolucionó hasta el actual Sistema de Calidad de la Educación (SIMCE), que cada año mide a estudiantes de los mismos niveles, además de segundo medio, para determinar si han alcanzado los objetivos fundamentales y los contenidos mínimos obligatorios del currículo nacional. En este sentido, la prueba SIMCE es actualmente un referente obligado de rendimiento (Agencia Calidad de la Educación, 2015).

Estas evaluaciones implementadas a nivel nacional tienen mucho sentido pues mucho del desempeño futuro, ya sea logro profesional o simplemente un mejor nivel de vida se fundamenta en haber tenido las adecuadas oportunidades escolares y una educación de calidad (Bellei, 2007; Comisión Económica para América Latina, 2008; Ministerio de Educación, 2003; Valenzuela, Bellei y de los Ríos, 2013; Villalobos y Valenzuela, 2012). Aspectos claves en la educación de calidad son las competencias vinculadas a lectura y escritura, matemáticas y ciencias (OECD, 2014). En particular, destaca la comprensión de lectura pues su papel no es solo la decodificación de información y la interpretación de información escrita, sino también y fundamentalmente

el desarrollo de conocimiento y del potencial personal que favorece una mejor inserción en la sociedad (Lorenzo, 2016; OECD, 2009; OECD, 2014; Pérez, 2009).

De ahí la importancia de la calidad de la educación y los resultados, pues se imponen como una necesidad para la igualdad de oportunidades (Torche, 2005). De larga data es el debate respecto de la educación y la (des)igualdad de oportunidades, y de cómo afecta la dificultad de acceso y una mala calidad de educación a las posibilidades futuras del alumnado y en definitiva de la sociedad (Donoso-Díaz, 2014).

A nivel nacional, el logro académico ha llegado a convertirse en sinónimo de la calidad de la educación y el Estado se ha empeñado en obtener evidencias de dicha calidad en función de resultado (Agencia Calidad de la Educación, 2015) a diferencia de las evaluaciones de proceso (Morgan, 2016). En lo concreto los organismos encargados de dicha evaluación asumen que la calidad está vinculada a la obtención de ciertos estándares.

Varios estudios realizados a nivel nacional han abordado la problemática de la calidad en educación o en la eficacia de las escuelas, a menudo desde una perspectiva económica. En su mayoría, esta clase de estudio ha buscado establecer los factores que pueden explicar las diferencias en el desempeño de los estudiantes o, de modo más general, en la eficacia del centro escolar. Es habitual comparar la educación privada y la pública, incluyendo el papel que puede jugar el aporte de un “voucher” (cheque escolar) a estas finalidades. Los estudios difieren en los énfasis dados al control de ciertas variables, como considerar el nivel socioeconómico de las familias, infraestructura de las escuelas, edad, educación de los padres, selección en la admisión de estudiantes, entre otros. No obstante, ya sea como consecuencia o no de estas decisiones metodológicas, hay poco consenso en los resultados, muchas de las variables incluidas en los estudios son estadísticamente significativas pero su relevancia en la predicción no es igualmente consistente (Correa, Parro y Reyes, 2014; Elacqua, 2012; McEwan y Carnoy, 2000; Somers, McEwan y Willms, 2004; Mizala y Torche, 2012; Osses, Bellei y Valenzuela, 2015; Thieme, Tortosa-Ausina, Prior y Gempp, 2012). Mayor acuerdo parece existir dentro de las líneas de investigación compartida por otros autores, quienes tienden a reafirmar que las diferencias en el desempeño de los estudiantes en áreas particulares (en especial matemática y lenguaje) son producto de las diferencias sociales y culturales

producto de la marcada segregación del sistema educativo chileno, lo que tiende a acentuar y perpetuar la polarización social de las familias, que permanecen ancladas en escuelas públicas o privadas según sea su condición (Muñoz-Chereau y Thomas, 2015; Torche, 2005; Villalobos y Valenzuela, 2012).

Buena parte de este numeroso conjunto de estudios hace uso, para desarrollar su análisis empírico, de las bases de datos que el SIMCE pone a disposición de la comunidad científica (Cádiz, 2006; Elacqua, 2012; Donoso y Hawes, 2002; Merino y Maldonado, 2014; Mizala y Torche, 2012; Muñoz-Chereau y Thomas, 2015; Orellana y Merino, 2013).

Son particularmente útiles pues además de los resultados en las pruebas con las que se evalúa el aprendizaje de los estudiantes, se han incorporado encuestas que deben ser respondidas por el profesor, los padres y el mismo estudiante. La información que recopilan recibe el nombre genérico de variables contextuales, pues en dichos instrumentos se solicitan datos sobre ingresos económicos y educación de los padres, acceso a algunos bienes (como internet) y clima escolar (Agencia de Calidad de la Educación, 2013). Estas variables contextuales son frecuentemente utilizadas como factores a considerar al momento de intentar determinar diferencias en el desempeño de los estudiantes o bien de la escuela. En el ya clásico estudio de Donoso y Hawes (2002) se establece que los grupos socioeconómicos son determinantes en el desempeño de lenguaje.

El efecto determinante de la condición social no ha cambiado mucho desde entonces. En estudios previos, como el de McEwan (2003) y Somers, McEwan y Willms (2004) también se enfoca, entre otras habilidades, en el lenguaje. Del conjunto de las variables independientes, aquellas relativas a: nivel socioeconómico de la familia, condición de ruralidad, dependencia del establecimiento (privado o público), años de estudio de los padres, implicación de los padres en la educación de los hijos, número de libros en el hogar, indicadores sobre atributos de los pares, género del estudiante y otras, fueron consideradas de relevancia para concluir que las escuelas privadas denotan mayor logro académico que las públicas.

Para Orellana y Merino (2013) hay cierto consenso en los predictores utilizados por los científicos en este tema. Los autores destacan especialmente la condición de

ruralidad, el acceso a la energía, el acceso al computador, acceso a internet, crecimiento comunal, analfabetismo y evaluación docente. No obstante, salvo el acceso a computador y la evaluación docente, las restantes variables contribuyen poco a explicar la varianza de la variable dependiente, que en este caso corresponde al promedio de las evaluaciones de matemática, lenguaje y comprensión del medio.

En el estudio de Correa, Parro y Reyes (2014) se consideraron las mismas variables dependientes de los estudios previos (lenguaje y matemáticas), estableciendo que los factores como el ingreso familiar, la educación del padre y la de la madre están significativamente relacionados con estas evaluaciones. De modo similar, el trabajo de Muñoz-Chereau y Thomas (2015) se destacaron factores como el género, la educación de la madre, el nivel educacional de la madre, número de libros en el hogar, y número de personas compartiendo la casa.

Mizala y Torche (2015) realizaron un estudio orientado a la estratificación del logro educacional que incorporó nuevas variables independientes medidas en SIMCE. De particular interés es la inclusión de las expectativas parentales, los años de experiencia del docente, el haber tenido educación preescolar y haber repetido de curso, todas estadísticamente significativas en la relación con la medición de lenguaje. Las expectativas parentales, a pesar de su intangibilidad, representa una fuerte influencia en el rendimiento académico (Davis-Kean, 2005; Pérez, Betancort y Cabrera, 2011).

La importancia de incluir la comprensión lectora no solo se entiende por su relevante papel en el desarrollo de conocimiento y potencial personal de los individuos, sino también porque es una evaluación clave para el curriculum nacional vigente y responden a la medición de habilidades que se consideran fundamentales en términos de aprendizaje (Agencia de Calidad de la Educación, 2015). La relevancia de la comprensión de lectura ha sido reconocida como un factor que impacta fuertemente en el logro académico en general (Fondo de Naciones Unidas para la Infancia, 2004) y ha sido tema recurrente de investigación, como lo demuestran numerosos estudios que establecen correlatos con el factor familiar, con la calidad de la lectura oral o directamente con factores cognitivos (Gil, 2009; Marchant, Lucchini y Cuadrado, 2007; Nuñez y Donoso, 2000; Svetina, Gorin y Tatsuoka, 2011), con modelos de comprensión (Parodi, 2011) e incluso de intervención (Ripoll y Aguado, 2014). Sin dejar de mencionar estudios que

insisten en vincular el factor socio-económico con el desempeño en el lenguaje, en aspectos específicos como las tareas de repetición de sentencias y comprensión de vocabulario (Balladares, Marshall y Griffiths, 2016),

En el contexto de lo expuesto, el propósito de este estudio es destacar, de las propuestas en estudios previos, aquellas variables que más influyen en el rendimiento lector, incluyendo no solo los factores socioeconómicos, sino también factores familiares, pues la evidencia empírica respalda su inclusión en la explicación de las diferencias en el grado de comprensión de lectura observado entre escuelas públicas y privadas.

## **Metodología**

### ***Muestra***

Para este estudio se utilizan datos secundarios, recogidos de la aplicación nacional de la prueba SIMCE (Sistema de Medición de la Calidad de la Educación). Es una prueba que se aplica cada año a estudiantes de 9-10 años (4° EGB), de 13-14 años (2° de ESO) y estudiantes de 15-16 años (4° de ESO) para evaluar la calidad de la enseñanza en función del producto o resultados.

La base original contiene los resultados de todos los estudiantes de 4° EGB evaluados durante el año 2014, que corresponden a 7667 escuelas y centros de estudio, con un total de 215.939 observaciones.

Siguiendo un diseño estratificado proporcional, se tomó en consideración a) la dependencia, que implica si la escuela recibe o no fondos del Estado, desde público (financiado íntegramente con aportes del estado), subvencionado (copago de padres más ayudas estatales) y privado (solo financiación de los padres) y b) el grupo socioeconómico, establecido estadísticamente a partir de varios criterios por la Agencia de Calidad de la Educación (grupo Bajo, Medio-Bajo, Medio, Medio-Alto y Alto). La Tabla 2 disponible la organización de los estudiantes en función de estas variables de estratificación:

Tabla 2. Distribución de frecuencia conjunta de estudiantes por variables de estratificación.

Grupo Económico	Dependencia			Total
	Municipal	Subvencionado	Privado	
Bajo	51			51
Medio-bajo	1772	1599		3371
Medio	1991	4294		6285
Medio-alto	172	2907		3079
Alto		398	875	1273
Total	3986	9198	875	14059
Fuente: Elaboración propia				

La muestra definitiva fue de 183 centros que reunían a 14.059 estudiantes y 463 cursos o clases. Se consideró que las escuelas y los centros no tuviesen menos de 50 estudiantes y a su vez, que se pudiese seleccionar más de 50 centros, pues hay algunas limitaciones reconocidas en el análisis con muestras pequeñas a nivel de centros (Maas y Hox, 2005).

### ***Diseño***

El diseño es ex post facto de carácter retrospectivo, donde los efectos de todas las variables ya han ocurrido, pero donde es posible reconocer factores y variables criterio (Montero y León, 2014). Se consideran como factores aquellas variables incluidas en los cuestionarios conjuntamente aplicados con las pruebas SIMCE. Solo se incluye un factor aleatorio (escuela). De un amplio conjunto de posibles variables, por motivos teóricos fueron seleccionadas originalmente las siguientes, que se asumen en el posterior modelo como variables de efectos fijos de nivel individual y grupal (escuela):

Tabla 3. Variables de nivel individual y grupal incluidas inicialmente en el modelo.

<b>Nivel individual</b>	Expectativas de logro académico
	Nivel de educación padre (en años)
	Nivel de educación madre (en años)
	Nivel de ingresos (ln)
<b>Nivel escuela</b>	Dependencia
	(municipal o pública, subvencionada y privada)
	Condición de ruralidad (0 = rural)
	Índice de vulnerabilidad por escuela (proporción)
	Grupo socio-económico
	(bajo, medio-bajo, medio, medio-alto y alto)
Fuente: Elaboración propia	

No obstante, después de los análisis preliminares, se determinó que a nivel grupal la dependencia, el índice de vulnerabilidad por escuela y el grupo socio-económico representaban casi la misma información y en consecuencia se seleccionó esta última por ser más comprensible fuera del contexto nacional.

Como variable de resultado del diseño, se ha definido el resultado que los estudiantes obtienen en la prueba de comprensión de lectura.

### ***Instrumentos***

Los datos se obtuvieron a partir de la aplicación de los instrumentos desarrollados por la Agencia Calidad de la Educación (2015) que utiliza una escala Likert, con cuatro opciones de respuesta y algunas preguntas que solicitaron datos específicos como edad, nivel de escolaridad. Se utilizaron los datos Cuestionario Psicosocial a Padres, respondido por los padres y/o apoderados de cada estudiante evaluado. Contiene información sobre relación de parentesco con el estudiante, edad del padre o madre, nivel educacional y de ingresos, etnia, expectativas de los padres sobre el estudiante, percepción de la actitud de los profesores o los pares hacia el estudiante, características y calidad organizacional del establecimiento (apoyo, comunicaciones, compromiso, exigencia, entre otros). De los cuestionarios, que agrupan estas variables contextuales, se han definido los factores del estudio.

Por su parte, la variable criterio correspondió a los resultados obtenidos en la prueba de comprensión de lectura, que ha sido desarrollada para garantizar la cobertura de aprendizajes del currículo nacional. Es una prueba centrada en habilidades y se la

considera la piedra angular para el aprendizaje de otras disciplinas, de allí su importancia (Agencia Calidad de la Educación, 2015; Paris y Stahl, 2005). Considera tres habilidades principales: a) Localizar: tareas de lectura donde el estudiante debe operar con los elementos explícitos del documento, implica discriminación y extracción de información específica; b) Relacionar e interpretar, el estudiante opera con elementos implícitos del texto, a los que se accede a partir de vínculos de elementos explícitos, lo que implica inferencia, lenguaje figurado, relaciones de significado, comprensión general entre otros aspectos; c) Reflexionar: confrontar diversos aspectos de los textos en forma y contenido, con el conocimiento de mundo y las lecturas previas que los estudiantes tienen.

### *Análisis*

El análisis de datos se realizó haciendo uso de la hoja electrónica de cálculo Microsoft Excel 2013 y el entorno de programación R (R Core Team, 2017). El análisis se inició con la descripción de las variables del estudio, para, además de caracterizarlas, evaluar el cumplimiento de los supuestos estadísticos y su ulterior inclusión en el modelo estadístico seleccionado. A partir de estos resultados, se efectuó un análisis de depuración de datos. Las observaciones atípicas fueron evaluadas mediante el paquete de R “mvoutlier” (Filzmoser, Garrett y Reimann, 2005) y eventualmente eliminadas.

Con los datos depurados, se decidió la imputación de las observaciones perdidas, cuando ello fue posible. A tal efecto se utilizó el procedimiento de imputación múltiple implementado en el paquete de R “Amelia”, basado en el algoritmo EMB (Honaker, King y Blackwell, 2001).

Posteriormente se obtuvieron los estadísticos descriptivos definitivos, se analizó la covariación entre variables y se decidió por la inclusión de aquellas que contribuyeron más a la solución en el modelo estadístico seleccionado

Para dar cuenta de los objetivos de investigación se seleccionó un modelo multinivel que incluyó una variable resultado, cuatro variables de efectos fijos de nivel individual, una variable de efecto fijo de nivel grupal y solo un factor aleatorio representado por las escuelas. El modelo puede ser expresado como:

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + b_{i1} z_{i1} + \epsilon_{ij}$$

El modelo original se interpreta como el valor de comprensión de lectura que es predicho por los factores: nivel educacional del padre y la madre, nivel de ingreso, expectativas de los padres en un mejor futuro social y económico, además de un factor de efecto fijo grupal (asignado a las escuelas), el grupo socioeconómico. Como factor aleatorio se consideró a las escuelas. El modelo se ajustó en R con el paquete “lme4” (Bates, Machler, Bolker y Walker, 2015).

## Resultados

En la Tabla 4, se disponen los resultados del análisis descriptivo de todas las variables que fueron definitivamente incluidas en el modelo.

Tabla 4. Estadísticos descriptivos de los factores individuales y la variable resultado organizados por la variable grupal grupo socio-económico.

	Bajo (n=51)		Medio-bajo (n=3371)		Medio (n=6285)		Medio-alto (n=3079)		Alto (n=1263)	
	Media	DT	Media	DT	Media	DT	Media	DT	Media	DT
Nivel educación padre (años)	10.8	3.52	11.8	3.56	13.5	3.52	16.5	3.8	20.1	3.06
Nivel educación madre (años)	9.2	4.52	11.8	3.32	13.5	3.18	16.3	3.58	19.2	3.34
Nivel de ingresos (ln)	11.9	0.84	12.4	0.64	12.8	0.63	13.4	0.69	14.3	0.49
Expectativa de los padres	4.3	1.31	4.5	1.15	4.9	0.88	5.3	0.64	5.7	0.51
Desempeño prueba lectura	237.6	49.1	255.7	47.89	269.9	48.28	281.0	46.36	300.2	43.29

Fuente: Elaboración propia

Se aprecia que la mayor parte de las diferencias en el nivel de educación del padre, la madre y las expectativas de los padres implican tamaños de efecto medianos (d de Cohen para diferencias de un 50% de la DT) o grandes (para diferencias de un 80% de la DT o superiores). En el nivel de ingresos (se informan los logaritmos naturales, en función de una práctica de análisis común en estudios socio-económicos) también se evidencia su relevancia medida en tamaño del efecto, especialmente si comparamos grupos extremos, en las diferencias en el grado de comprensión de lectura. Los cambios en los factores y en la variable resultado son concomitantes y reflejan posibles efectos sobre esta última.

En la tabla 5 se disponen los modelos con la información de ajuste. Al inicio se presenta el modelo nulo. Los resultados obtenidos son de interés para contrastar el resultado de los subsiguientes ajustes. En particular, es importante porque con las varianzas de los efectos aleatorios se puede calcular el coeficiente de correlación intraclase (CCI), como una forma de evaluar si el efecto aleatorio de las escuelas está presente en los datos, o bien qué proporción de la varianza total es explicada por las

diferencias entre centros. En este caso,  $352.8/(352.8+2066.7) = 0.15$  que significa que aproximadamente un 15% de la variabilidad presente en la prueba de comprensión lectora puede ser explicado por las diferencias entre escuelas.

El siguiente modelo incluyó el factor grupo socio-económico. Respecto de los efectos aleatorios, representado por las escuelas, el porcentaje de varianza se reduce. En efecto,  $177.9/2244.5 = 0.08$ , esto es, el desempeño en comprensión lectora es explicado en un 8% por las diferencias entre las escuelas. La importancia de este cambio reside en que es la variable grupo socio-económico la que está explicando buena parte de la varianza. Por su parte, los efectos fijos denotan coeficientes que reflejan de qué modo el grupo socioeconómico influye sobre el desempeño en comprensión lectora (la variable “dummy” que representa la categoría “Alto” coincide con el valor del intercepto y por ello no está presente).

La incorporación de nuevos factores permitió el ajuste del modelo definitivo, como la opción de mejor ajuste que incluyó las variables informadas en la parte final de la Tabla 5. En este último modelo ajustado, respecto de los efectos aleatorios (escuela), se aprecia nuevamente la reducción de varianza explicada, pues  $152.8/2147.5 = 0.07$  o 7% del desempeño en comprensión de lectura que es explicado por las diferencias entre escuelas. La reducción en la varianza producto de la incorporación de las nuevas variables individuales fue menor. No obstante, de los efectos fijos, se aprecia que uno de los mayores coeficientes (no estandarizado, comprendido como unidad de cambio en Y) corresponde a expectativa de los padres.

Tabla 5. Estadísticos descriptivos de los factores individuales y la variable resultado organizados por variable grupo socio-económico.

Efectos aleatorios			Efectos fijos					
Escuelas	Varianza	Desv. Est.		Coefficiente	Error estándar	gl	t	p-valor
<b>Modelo nulo</b>								
Intercepto	352.8	18.78	Intercepto	270.573	1.445	180.51	187.30	0.000
Residual	2066.7	45.46						
Total	2419.5							
<b>Modelo con variable de grupo</b>								
Intercepto	177.9	13.34	Intercepto	301.524	3.587	176.01	84.06	0.000
Residual	2066.6	45.46	GSE Bajo	-63.899	15.208	197.27	-4.20	<b>0.000</b>
Total	2244.5		GSE Medio-bajo	-46.131	4.131	177.66	-11.17	<b>0.000</b>
			GSE Medio	-32.745	3.945	175.80	-8.30	<b>0.000</b>
			GSE Medio-alto	-20.151	4.248	176.62	-4.74	<b>0.000</b>
<b>Modelo con variables individuales y de grupo</b>								
Intercepto	152.8	12.36	Intercepto	216.805	9.599	6457.00	22.59	0.000
Residual	1994.7	44.66	Nivel educación padre	0.405	0.122	13954.00	3.33	<b>0.001</b>
	2147.5		Nivel educación madre	0.588	0.135	14000.00	4.37	<b>0.000</b>
			Nivel de ingresos	1.328	0.678	14059.00	1.96	0.050
			Expectativa padres	8.104	0.460	14000.00	17.63	<b>0.000</b>
			GSE Bajo	-39.241	14.357	211.00	-2.73	<b>0.007</b>
			GSE Medio-bajo	-25.656	4.103	232.00	-6.25	<b>0.000</b>
			GSE Medio	-17.710	3.850	214.00	-4.60	<b>0.000</b>
			GSE Medio-alto	-12.517	4.022	191.00	-3.11	<b>0.002</b>

Fuente: Elaboración propia

Destaca del análisis anterior que las variables individuales denotan efectos estadísticamente significativos. En particular, las expectativas de los padres y el nivel educacional. No obstante, cabe señalar dos hechos respecto de este resultado: a) Se observa que la reducción de la varianza explicada debido a la inclusión del conjunto de variables individuales apenas la reduce en un 1%, con lo que el efecto real de estas es escaso, b) en este sentido, la significación de estas variables individuales obedece más bien a la magnitud del tamaño muestral, que vuelve más sensibles los contrastes, antes que a efectos sustantivos. Esta situación justifica aún más incluir las escuelas como factor aleatorio en el estudio y resalta la relevancia del grupo socio-económico en el desempeño en comprensión de lectura.

La variable grupo socio-económico se relaciona con el nivel de educación de los padres y las expectativas, pero incorpora otros elementos a nivel escuela que se deben considerar, y que están representados por la vulnerabilidad social de las familias, el acceso a servicios y una mejor infraestructura, aspectos enmascarados por dicho estatus socioeconómico.

## **Discusión**

En el estudio se buscaba aportar evidencia en favor de la hipótesis que establece una influencia de factores sociales y económicos (variables contextuales) sobre la comprensión de lectura tal como son considerados en el proceso de recolección de información de la Agencia de Calidad de la Educación.

En términos de las variables contextuales que se han considerado, la comprensión de lectura de los estudiantes se ve influida por factores refrendados por estudios previos. Estas variables sociales y económicas apuntan al hecho de que un mejor apoyo en estos aspectos tiene relación con un mejor rendimiento. El nivel de educación del padre, el nivel de educación de la madre y las expectativas de los padres respecto del futuro de su hijo/a (hasta dónde esperan que llegarán en términos profesionales) representan influencias de relevancia en el desempeño académico y especialmente en la lectura. Esto es consistente con diversos hallazgos que apuntan en la misma dirección: padres con mayor educación tienen más interés en que sus hijos alcancen un nivel educativo similar y por tanto, suelen estar más pendientes de sus estudios (por ejemplo, que hagan sus tareas), siendo además dicha influencia, mucho más fuerte en su infancia que en su adolescencia (Pérez, Betancort y Cabrera, 2011). Las expectativas de los padres también denotan una influencia esperable, pues hay una relación con el compromiso de estos en la educación de sus hijos y con las acciones de control que estos ejercen para que los estudiantes logren sus metas (Davis-Kean, 2005; Pérez, Betancort y Cabrera, 2011; Mizala y Torche, 2014).

Aun cuando estos factores denotan una influencia estadísticamente significativa, debe tenerse presente que sustantivamente el efecto es marginal. Sin embargo, no ocurre lo mismo cuando se considera el nivel socioeconómico, pues este factor explica una parte relevante de las diferencias entre las escuelas. Si bien es cierto que el nivel socioeconómico enmascara las variables que tienen directa relación con estas diferencias, se sabe de las implicaciones que este factor tiene para el desempeño en lenguaje. Aspectos como peor infraestructura y recursos materiales y humanos en la escuela, menores oportunidades futuras para los estudiantes, especialmente de carácter profesional (lo que por cierto se refleja en menores expectativas futuras si estos forman parte de familias con menor nivel socioeconómico), y menor acceso a redes de apoyo (Donoso-Díaz, 2014;

Mizala y Torche, 2012), configuran una realidad social que tiene consecuencias profundas en los estudiantes.

Desde la perspectiva de la comprensión, esto puede parecer fuera del contexto académico, pues estas diferencias debiesen estar vinculadas principalmente a atributos individuales, como las habilidades cognitivas a la base de comprensión de lectura (Svetina, Gorin y Tatsuoka, 2011), sin embargo al aportar evidencia relativa a un desempeño comprensivo lector dependiente en parte de las diferencias socioeconómicas entre escuelas, se está también frente a un factor de difícil tratamiento e intervención y que en Chile ha llegado a ser un problema sistémico.

El resultado obtenido también lleva a interrogantes relacionadas con la inclusión o al menos consideración de este factor (y sus implicaciones) al momento de llevar a cabo evaluaciones estandarizadas, que aparentemente en cada cohorte vuelve a evidenciar las mismas diferencias, al margen de este contexto. En efecto, y puede consultarse en estudios previos, año con año, se observan similares diferencias de desempeño en comprensión entre escuelas, mientras se sigue asumiendo que estas son consecuencia del logro individual y la formación recibida, sin incorporar explícitamente el factor socioeconómico.

En cuanto a las limitaciones, el presente estudio solo trabajó con una cohorte correspondiente a la evaluación del año 2014 y no se incluyó a todas las escuelas, pues se fijaron criterios de selección. Además, los resultados son solo atribuibles a estudiantes de 4° de Enseñanza General Básica, pues no se consideraron otros niveles.

En términos de proyecciones, el presente estudio contribuye a resaltar la necesidad de utilizar una aproximación multinivel en evaluaciones estandarizadas de grandes poblaciones de estudiantes y refuerza la inclusión de diferencias entre centros escolares. Es claro que para el contexto nacional, no tiene sentido establecer modelos que no consideren este factor. Será necesario continuar con trabajos que incorporen nuevas variables individuales pero de carácter cognitivo y actitudinal, por su potencial poder explicativo. En cualquier caso, no debe perderse de vista que, en virtud de estos resultados, una parte importante del rendimiento depende de las escuelas. Pensemos lo que significa un 15% poniéndolo en relación con las puntuaciones PISA de 2013 en las que Chile obtuvo en habilidad lectora 441 puntos y ocupó el puesto 52; si se hubiera

logrado una mejora del 15%, Chile habría pasado a ocupar el 4° puesto. En consecuencia, disminuyendo la brecha entre escuelas se pueden lograr grandes resultados.

## **Estudio 2: Evaluación de la comprensión lectora mediante un modelo de diagnóstico cognitivo**

### **Introducción**

#### *Relevancia de los modelos de diagnóstico cognitivo*

No es un misterio la relevancia que tiene la comprensión de lectura en el desempeño de las diversas materias escolares y en el desarrollo académico general de los estudiantes. Se la considera de hecho, una de las competencias genéricas clave incluso a nivel universitario (Washer, 2007). Tal es su importancia, que se ha incluido en muchas evaluaciones internacionales, como ocurre con PISA (Programme for International Student Assessment), PIRLS (Progress in International Reading Literacy Study) (Mullis, Martin, Foy, y Hooper, 2017) o TERCE (Tercer Estudio Regional Comparativo y Explicativo).

Sin embargo estas evaluaciones, así como aquellas que se realizan en diversos países ya sea a nivel nacional o local, ofrecen medidas sobre el desempeño de los estudiantes cuyo propósito es posicionar a los sujetos en un continuo o escala, de modo que las mediciones son entendidas como una suma de puntuaciones (de la Torre y Minchen, 2017) para tomar decisiones relevantes tales como aprobar o reprobar examinados o seleccionar a los sujetos para algún propósito, entre otras opciones. En este tipo de evaluación clásica los ítems miden una única dimensión latente (Wu, Tam y Jen, 2016). En estas medidas se aplica Teoría Clásica de los Tests (TCT) o bien alguno de los modelos de Teoría de Respuesta al Ítem (TRI) para estimar los parámetros de los ítems y de los individuos. En ambos casos se requiere el cumplimiento del supuesto de unidimensionalidad (Baker y Kim, 2017).

Cuando la evaluación es más compleja o bien se requiere información más detallada de los evaluados en cuanto a las posibles deficiencias que explican sus respuestas incorrectas, una interesante alternativa a los modelos basados en el supuesto de unidimensionalidad son los modelos de diagnóstico cognitivo (CDM, por su acrónimo en inglés). Los CDM permiten evaluar el dominio de más de una variable latente, entendidas como habilidades o atributos cognitivos medidos por los ítems que componen el test (de la Torre, 2009; de la Torre, 2011; DiBello, Roussos y Stout, 2007). Son modelos

probabilísticos multidimensionales confirmatorios de clase latente (Rupp y Templin, 2008) que además de una evaluación más detallada del evaluado, también ayudan a proveer una retroalimentación a los estudiantes y permitir mejorar el subsecuente aprendizaje (Kuo, Chen, Yang y Mok, 2016). Esta perspectiva es lo que se comprende como evaluación centrada en el estudiante, que implica un análisis detallado de los atributos medidos, permitiendo que el examinado conozca los factores que influyen en su desempeño, lo que a su vez da al profesor la oportunidad de trabajar fortalezas y debilidades específicas, quienes son de hecho quienes mejor puede dar cuenta de las habilidades de sus estudiantes, diagnosticar sus dificultades y proponer, eventualmente, medidas oportunas (Boora, Pasiphol y Tangdhanakanond, 2015; López-Higes y Rubio-Valdehita, 2014).

Los Modelos de Diagnóstico Cognitivo (CDM) parten del supuesto de que la población está formada por mixturas o distribuciones de sujetos con distintas características o habilidades. El objetivo de los CDM es confirmar la estructura teórica en un test en el que se establecen las habilidades que explican el desempeño de los ítems y que permiten diferenciar las distintas clases de sujetos en función de las habilidades que estos dominan, conformando patrones de dominio de un constructo que se asume multidimensional.

Las principales diferencias entre los CDM con respecto a la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI) son las suposiciones de variable latente discreta y multidimensionalidad. Tanto TCT como TRI unidimensional, incluidos los modelos multicomponenciales como el Modelo Logístico Lineal (LLTM, ver Alvarado, Puente, Fernández, y Jiménez, 2015) asumen que la variable latente es continua y unidimensional, al no existir clases de sujetos se hace posible ordenarlos en una escala unidimensional según su mayor o menor nivel en el rasgo.

El objetivo del presente trabajo es determinar la composición de atributos cognitivos que está a la base de una popular medida de comprensión lectora aplicada en todo Chile, identificando los patrones de dominio cognitivo que prevalecen entre estudiantes de Cuarto Año Básico, contribuyendo a dilucidar las habilidades que parecen concursar en una aptitud tan relevante como la comprensión. Se hace una breve revisión de las aplicaciones de los CDM a la lectura y comprensión lectora, explicándose la lógica

del procedimiento y las características de los modelos más frecuentemente utilizados, para finalmente realizar una aplicación de los CDM, mostrándose las ventajas de esta perspectiva frente a otras alternativas más clásicas (TCT o TRI).

### ***Modelos de diagnóstico cognitivo aplicados en el área de lenguaje***

Los modelos de diagnóstico cognitivo son desarrollados con el propósito de medir una colección de habilidades o atributos cognitivos detallados, que se supone están a la base del desempeño de las tareas propuestas en los ítems de un test (de la Torre y Minchen, 2014). Gierl, Leighton y Hunka (2000) han definido los atributos como una descripción del conocimiento procedimental o declarativo necesario para llevar a cabo una tarea en un dominio específico. Los modelos de esta clase generalmente se inician descomponiendo la habilidad general, por ejemplo, la comprensión de lectura, en un conjunto finito de atributos discretos (Köhn, Chiu y Brusco, 2015). El proceso implica usualmente que un grupo de jueces estudie los ítems y establezca en función de alguna perspectiva teórica, cuáles son las habilidades o atributos base que el examinado debe dominar para poder responder correctamente un ítem (García, Olea y de la Torre, 2014; Hemmati, Baghaei y Bemani, 2016; Jang, 2009; Rupp, Templin y Henson, 2010). La aptitud para comprender un texto se entiende entonces como un compuesto de habilidades o atributos que se puede o no dominar. La consecuencia es que se obtendrán perfiles de dominio conformados por dichos atributos.

Una vez se han definido qué atributos cognitivos condicionan el desempeño en cada uno de los ítems, es posible construir la denominada matriz  $Q$ , que resume la relación entre los ítems y los atributos, indicando qué atributos deben dominarse o no para responder correctamente dichos ítems. En la matriz  $Q$ , que se estructura como una tabla de doble entrada con atributos cognitivos en columnas e ítems en filas, el dominio del atributo se representa mediante el valor 1, mientras que el no dominio con el valor 0 en las celdas que cruzan un atributo con un ítem (de la Torre, 2009; de la Torre y Chiu, 2016; García, Olea y de la Torre, 2014).

Una vez obtenida la matriz, se debe seleccionar uno de los modelos de diagnóstico disponibles. En particular aquí se aboga por aquellos que han sido usados en el estudio del lenguaje (Li, Hunter y Lei, 2015). Es necesario decidir si el modelo de diagnóstico ha de ser de naturaleza conjuntiva o compensatoria. Un modelo conjuntivo implica que

para responder correctamente a un ítem, se deben dominar todos los atributos que condicionan su desempeño. Por su parte, los modelos compensatorios implican que un pobre dominio en un atributo cognitivo, puede ser compensado con un elevado dominio en otro atributo (Li, Hunter y Lei, 2015; Rupp, Templin y Henson, 2010).

En el ámbito del lenguaje, los modelos más utilizados son los llamados modelos conjuntivos, lo que se sustenta en los modelos teóricos sobre lenguaje que explican el desempeño en tareas de comprensión (Buck, Tatsuoka y Kostin, 1997; Hemmati, Baghaei y Bemani, 2016; Jang, 2009; Ravand, Barati y Widhiarso, 2013; Ravand, 2016; von Davier, 2005). Algunos de estos modelos son el Rule Space Model (Svetina, Gorin y Tatsuoka, 2011; Tatsuoka, 2009), el modelo Attributes Hierarchy Model o AHM (Wang y Gierl, 2011), el modelo Deterministic-Input, Noisy-And-Gate, DINA, el modelo Reparameterized Unified Model (RUM), una derivación de este el Fusion Model (Jang, 2009) y el modelo llamado Reduced Reparameterized Unified Model, o RRUM (Jang, 2009; Jang, Dunlop, Wagner, Kim y Gu, 2013; Li, Hunter y Lei, 2015), también reconocido como Noncompesatory Reparameterized Unified Model, NC-RUM (Ravand y Robitzsch, 2015). Por su parte, un conocido modelo compensatorio usado en esta área es el Deterministic-Input, Noisy-OR-Gate, DINO (Liao, Kuo y Deenang, 2015). Entre los modelos generales o saturados, está el modelo Generalized Deterministic-Input, Noisy-And-Gate Model, G-DINA (Chen y Chen, 2016; Ravand, 2016) que agrupan otros como DINA, DINO o RRUM.

El modelo DINA es el más parsimonioso, en el que las iniciales DA “Deterministic Input” indican si los examinados en la clase latente  $c$ , han dominado o no todos los atributos medidos para el ítem  $i$ . “And Gate” se refiere a la relación conjuntiva entre los atributos y “Noisy” se refiere al elemento estocástico del modelo DINA que enfatiza que las respuestas de los examinados no son deterministas, sino probabilísticas. Lo estocástico, en este caso está representado por la inclusión de los parámetros distracción ( $s_{ij}$ ) y conjetura ( $g_{ij}$ ), que implica que un examinado que domine todos los atributos requeridos en un ítem, puede no responderlo por distracción o bien, que aún sin tener dominio sobre dichos atributos responda correctamente solo por azar (Ravand, Barati y Widhiarso, 2013). A partir del modelo DINA se pueden definir otros modelos más sencillos como el modelo DINO en el que solo es preciso un atributo de los especificados en la matriz  $Q$  para que la respuesta al ítem sea correcta. Como en DINA,

también incluye los parámetros de desliz/distracción y conjetura. Un modelo compensatorio que se sitúa entre DINA y DINO es el modelo RRUM (Kim, 2011) que pertenece a la familia de los modelos de Teoría de Respuesta al Ítem multidimensionales. En el modelo RRUM puede obtenerse respuesta correcta aunque no se dominen todos los atributos especificados, pero se penaliza su falta de forma progresiva.

En el modelo DINA -en un test compuesto de  $J$  ítems que miden  $K$  atributos- se dividen los  $2^K$  vectores de atributos latentes o clases, en dos grupos. En función de los requerimientos del dominio de atributos para un ítem  $j$ , se establece de modo determinista la clasificación de cada uno de los  $2^K$  vectores de atributos en uno de estos dos grupos: a)  $\eta_{j1}$  que corresponde a la agrupación de los vectores de atributos que cumplen con todas las habilidades requeridas para el ítem  $j$ , b) y  $\eta_{j0}$ , que es el grupo que reúne los vectores de atributos que carecen de al menos uno de los atributos requeridos para responder correctamente el mismo ítem (de la Torre, 2011).

Con el modelo DINA, los vectores de atributos que han sido reunidos en el mismo grupo se asume que tienen la misma probabilidad de responder correctamente el ítem. Pero este supuesto no necesariamente se mantiene para el grupo  $\eta_{j0}$  pues debido a que los vectores de atributos que lo conforman pueden llegar a tener grados variables de deficiencia con respecto al dominio de las habilidades requeridas para dar una respuesta correcta, sus probabilidades de éxito podrían no ser idénticas. La generalización del modelo DINA, o G-DINA relaja el supuesto de igual probabilidad de éxito para todos los vectores de atributos en  $\eta_{j0}$  (de la Torre, 2011).

Para cada uno de los ítems, el modelo G-DINA divide las clases latentes en  $2^{K_j^*}$  grupos latentes, donde  $K_j^* = \sum_{k=1}^K q_{jk}$  representa el número de atributos requeridos para el ítem  $j$ . Se define  $\alpha_{lj}^*$  como el vector de atributos reducido cuyos elementos son los atributos requeridos para el ítem donde  $l = 1, \dots, 2^{K_j^*}$ , que representa el número total de grupos latentes o patrones de atributos latentes requeridos (patrones de dominio). En el modelo G-DINA, la probabilidad de que un examinado perteneciente a la clase latente  $l$ , con un patrón de atributos  $\alpha_{lj}^*$  responda correctamente un ítem  $j$  se denota como  $P(X_j = 1 | \alpha_{lj}^*) = P(\alpha_{lj}^*)$  (García, Olea y de la Torre, 2014). La formulación original del

modelo G-DINA basada en  $P(\alpha_{ij}^*)$  puede ser descompuesta en la suma de los efectos debido a la presencia de atributos específicos y sus interacciones, así:

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'} \alpha_{lk} \alpha_{lk'} \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk} \quad (1)$$

Donde  $\delta_{j0}$  es el intercepto para el ítem  $j$ ,  $\delta_{jk}$  es el efecto principal debido a  $\alpha_k$ ,  $\delta_{jkk'}$  es el efecto de interacción debido a  $\alpha_k$  y  $\alpha_{k'}$  y  $\delta_{j12\dots K_j^*}$  es el efecto de interacción debido a  $\alpha_1 \dots \alpha_{K_j^*}$ . De los parámetros expuestos, se puede señalar que  $\delta_0$  es la probabilidad basal o probabilidad de respuesta correcta cuando ninguno de los atributos está presente,  $\delta_k$  es el cambio en la probabilidad de una respuesta correcta como resultado de dominar un único atributo (por ejemplo  $\alpha_k$ ),  $\delta_{kk'}$  es un efecto de interacción de primer orden y representa el cambio en la probabilidad de una respuesta correcta debido al dominio tanto en  $\alpha_k$  y  $\alpha_{k'}$  que está más allá del impacto aditivo de dominio de ambos atributos, y  $\delta_{12\dots K_j^*}$  representa el cambio en la probabilidad de una respuesta correcta debido al dominio de todos los atributos requeridos más allá del impacto aditivo de los efectos principales y de interacción. El intercepto es siempre no negativo, los efectos principales son usualmente no negativos, pero los efectos de interacción pueden tomar cualquier valor. Cabe notar que los efectos principales son no negativos si  $P(\mathbf{0}_{K_j^*}) \leq P(\alpha_{ij}^*)$  para  $\sum_{k=1}^{K_j^*} \alpha_{ljk}^* = 1$ , donde  $\mathbf{0}_{K_j^*}$  es el vector nulo de largo  $K_j^*$ . Esto implica que tener dominio sobre cualquiera de los atributos cognitivos requeridos para un determinado ítem representa un incremento en la probabilidad individual de obtener un resultado exitoso en el ítem (de la Torre, 2011).

## **Método**

### ***Muestra***

Corresponde a una muestra no probabilística intencional en el que participaron tanto estudiantes de escuelas públicas como privadas situadas en las regiones metropolitana y de Valparaíso de Chile. Los examinados cursaban 4º Básico, con edades comprendidas entre los 9 y los 10 años. Se evaluó a un total de 501 sujetos que se redujo a 434 estudiantes, una vez eliminados los protocolos no respondidos, inaceptablemente incompletos o erróneos. De los 434 estudiantes, 256 (59%) eran de sexo femenino y 178

(41%) de sexo masculino (41%). El número de examinados de escuelas públicas fue de 129 (29.7%) y el de escuelas privadas de 305 (70.3%) estudiantes.

### ***Diseño***

El estudio es cuantitativo, ex post facto y de carácter instrumental ya que se enfoca en las propiedades psicométricas del test desde la perspectiva cognitiva.

### ***Instrumento***

Se aplicó la prueba de Comprensión Lectora y Producción de Textos (CP-PT) que evalúa comprensión lectora, producción de textos y manejo de la lengua, con fiabilidades test-retest por dimensión que fluctúan entre 0.61 a 0.82, dependiendo del rango de edad a aplicar. Su validez referida a criterio se estableció comparando el desempeño de esta prueba con las calificaciones escolares finales, con correlaciones que fluctuaron entre 0.32 a 0.51 dependiendo del rango de edad (Medina y Gajardo, 2014). Es una prueba comercial estandarizada diseñada para estudiantes de 5 a 14 años, se realiza a partir de 45 ítems agrupados en 14 ejercicios sobre una diversidad de contenidos. En el presente estudio se puntuaron con cero cuando la respuesta al ítem era incorrecta y uno cuando era correcta.

### ***Procedimiento***

Siguiendo las recomendaciones de Ravand y Robitzsch (2015) el estudio CDM comienza por la definición de la matriz  $Q$  en la que se indican los atributos que requiere cada ítem para poder contestarlo correctamente, a continuación, se procedió a la depuración de la base de datos para la selección de los ítems más discriminativos, para finalmente especificar los modelos y comparar su bondad de ajuste.

### ***Definición de la matriz $Q$***

Para construir la matriz, en primer lugar, se definieron los atributos a partir de estudios previos que habían utilizado esta metodología en lectura y comprensión lectora (Buck, Tatsuoka y Kostin, 1997; Hemmati, Baghaei y Bemani, 2016). En segundo lugar, se instruyó a una experta en comprensión de textos sobre el uso de estos modelos y se le pidió que elaborase un modelo atribucional que se adecuara al test utilizado, procediendo por retroajuste, mediante un análisis post hoc del contenido del instrumento (Ravand y

Robitzsch, 2015). En tercer lugar, una vez que la experta definió los atributos cognitivos con base en la literatura provista, se envió el resultado a tres jueces independientes, todos académicos de área de lingüística, quienes evaluaron la selección de atributos y su asignación a los ítems del test. Una vez obtenidas las evaluaciones de los jueces, se le pidió a la experta que volviese a reconsiderar su selección y la asignación realizada. Finalmente, se efectuó una última consulta con los jueces. Con base en esta revisión final se obtuvo la matriz  $Q$  definitiva.

#### *Especificación de los modelos*

Primero fue necesario definir si los atributos son de naturaleza compensatoria o conjuntiva (Lee y Sawaki, 2009). Los jueces se inclinaron por la naturaleza no compensatoria de los atributos, existiendo además mayor evidencia en favor de estos últimos en el contexto de la comprensión de lectura (Li, Hunter y Lei, 2015). En consecuencia, se trabajó con los modelos G-DINA, DINA, DINO (Chen y Chen, 2016; Liao, Kuo y Deenang, 2015; Ravand, 2016) y RRUM (Kim, 2011).

#### *Comparación de los modelos*

La comparación se realizó con el fin de establecer cuál era el que mejor bondad de ajuste mostraba y en consecuencia permitía describir mejor los resultados obtenidos, para juzgar la bondad de ajuste se utilizaron la razón de verosimilitud, el criterio de información de Akaike (AIC) y el criterio de información bayesiano (BIC) (Hu, Miller, Huggins-Manley y Chen, 2016), en los que cuanto menor es el valor del índice mejor es la bondad de ajuste (Ravand y Robitzsch, 2015). También se utilizaron índices absolutos, que comparan el ajuste del modelo a los datos reales. Entre ellos, se incluyó MADcor, o la media de las desviaciones absolutas entre las correlaciones observadas y predichas para cada par de ítems  $i$  y  $j$ , esperando valores iguales o menores a 0.05 (DiBello, Roussos y Stout, 2007; Lei y Li, 2016), el índice  $\max\chi^2$  que de resultar significativo ( $p < 0.05$ ) indica que se viola la independencia estadística entre pares de ítems y por tanto el modelo no ajusta bien (Hu, Miller, Huggins-Manley y Chen, 2016), la raíz cuadrada de la media de residuos estandarizados (SRMSR, de su acrónimo en inglés), que se le considera un promedio de correlaciones residuales, donde esta última es la correlación muestral menos la correlación esperada entre cada par de ítems  $i$  y  $j$  (Maydeu-Olivares y Joe, 2014), siendo valores aceptables aquellos que sean iguales o menores a 0.05 y por último el índice

100\*MADRESIDCOV o MADRCOV, que compara covarianzas o correlaciones observadas y predichas de pares de ítems. A menor tamaño de efecto, mejor ajuste (Ravand y Robitzsch, 2015).

Para la comparación de modelos y evaluar la bondad de ajuste se utilizó el entorno de programación R (R Core Team, 2017) y específicamente el paquete Cognitive Diagnosis Modeling, CDM (George, Robitzsch, Kiefer, Groß y Ünlü, 2016).

### **Resultados**

El primer paso fue la construcción de la matriz  $Q$ . Se definieron seis habilidades presentes en los ítems. Una definición breve de los atributos se dispone en la tabla 6.

Tabla 6. Atributos subyacentes en la comprensión de lectura en el test CL-PT.

Parámetro	Atributo	Descripción
$\alpha_1$	Código de superficie	Es un nivel representacional y de procesamiento en el que se mantiene en memoria de modo exacto lo que el texto entrega. Se mantiene la forma (rasgos formales) y la representación lingüística. El tipo de proceso es el básico de decodificación y reconocimiento de palabras. Implica solo reproducción de contenidos y uso de recuerdo. (Riffo, 2016)
$\alpha_2$	Base de texto	Nivel de procesamiento oracional. Los elementos de una oración leída por un usuario son estructurados mediante micro-proposiciones que están conectadas a nivel local, generando un microestructura textual o coherencia local. Permite realizar inferencias básicas que permiten conectar una idea con otra. Es un nivel representacional que permite parafrasear el texto a través del uso de macro-estrategias. (Hemmati, Baghaei y Bemani, 2016; Riffo, 2016).
$\alpha_3$	Inferencia	Operación que ayuda a agregar la información faltante de aquello que no expresa el nivel de superficie. Las inferencias participan del proceso de comprensión y provienen de conocimiento general, dominios específicos y representaciones basadas en experiencias personales o lecturas previas. Tienen lugar durante la construcción del modelo de situación (Hemmati, Baghaei y Bemani, 2016; Li, Hunter y Lei, 2015; Vieiro y Gómez, 2004).
$\alpha_4$	Modelo de situación	Nivel representacional que permite asumir aprendizaje de lo leído. Hace uso del conocimiento previo y el proceso de inferencia. Representa un momento de integración entre el conocimiento explícito y el modelo de mundo que tiene quien comprende (Gernsbacher y Kaschak, 2013; Piacente, 2009).
$\alpha_5$	Instrucciones retóricas	Representan ayudas a la comprensión del texto. Estas ayudas no se refieren a la temática tratada sino al texto en sí y a la relación que se invita a tener con él. Son marcadores en el texto que deben ser interpretados como instrucciones de procesamiento (Sánchez, 1996).
$\alpha_6$	Habilidades lingüísticas	Involucran procesos cognoscitivos y son parte del desarrollo de un individuo, permitiendo la comprensión y expresión de mensajes, elaboración de ideas, reflexión y solución de problemas (Guarneros, y Vega, 2014). Son consideradas generales y por tanto más allá de la modalidad escrita.

Se tuvo en cuenta la recomendación de identificar entre tres a 10 atributos cognitivos (DiBello, Roussos y Stout, 2007; Rupp y Templin, 2008).

Tabla 7. Matriz Q de la prueba de comprensión lectora CL-PT.

Ítems	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	Ítems	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$
P01a	1	1	1	0	1	0	P06a	1	0	0	0	1	0
P01b	1	1	1	0	1	0	P06b	1	0	0	0	1	0
P01c	1	1	1	0	1	0	P06c	1	0	0	0	1	0
P01d	1	1	1	0	1	0	P06d	1	0	0	0	1	0
P01e	1	1	1	0	1	0	P08	1	0	0	1	0	1
P01f	1	1	1	0	1	0	P09	1	0	1	0	0	0
P01g	1	1	1	0	1	0	P10a	1	1	1	0	0	0
P01h	1	1	1	0	1	0	P10b	1	1	1	0	0	0
P01i	1	1	1	0	1	0	P10c	1	1	1	0	0	0
P01j	1	1	1	0	1	0	P11	1	1	1	0	0	0
P01k	1	1	1	0	1	0	P12a	0	0	0	0	0	1
P01l	1	1	1	0	1	0	P12b	0	0	0	0	0	1
P01m	1	1	1	0	1	0	P12c	0	0	0	0	0	1
P01n	1	1	1	0	1	0	P13a	1	1	0	0	0	0
P01o	1	1	1	0	1	0	P13b	1	1	0	0	0	0
P01p	1	1	1	0	1	0	P13c	1	0	1	0	0	0
P01q	1	1	1	0	1	0	P13d	1	0	0	0	0	0
P01r	1	1	1	0	1	0	P14a	1	1	0	0	0	0
P02	1	0	0	1	0	1	P14b	1	1	0	0	0	0
P03	1	1	0	0	1	0	P14c	1	1	0	0	0	0
P04	1	1	1	0	0	0	P15a	1	0	1	1	0	0
P05	1	0	1	0	0	0	P15b	1	1	1	0	0	0
							P15c	1	1	1	0	0	0

En la tabla 7, se dispone la matriz  $Q$  para todos los ítems de la prueba de comprensión lectora. Código de superficie es un atributo básico en casi todos los ítems, mientras que modelo de situación o habilidades lingüísticas tienen un papel menos relevante en el proceso comprensivo.

#### *Estimación de la dificultad y discriminación de los ítems mediante CDM*

En la tabla 8 se han dispuesto las medias ( $M$ ) y desviaciones típicas ( $s$ ) de los ítems, en las dos primeras columnas desde la izquierda. Se eliminaron los ítems 14a, 14b y 14c ya que presentaban un alto porcentaje de no respuesta, mayor al 20%.

Junto con estos estadísticos descriptivos y gracias al modelo DINA, fue posible obtener además los valores de la dificultad y discriminación estimadas mediante este modelo (de la Torre, 2009). Sea  $s_j$  el parámetro distracción que representa la probabilidad

de fallo al responder un ítem aun cuando el examinado domine todos los atributos necesarios para brindar una respuesta correcta y sea  $g_j$ , el parámetro conjetura que representa la probabilidad de acierto al responder un ítem aun cuando el examinado no posea los atributos necesarios para brindar una respuesta correcta, entonces el valor de discriminación de un ítem se establece como  $\omega_{1j} = 1 - g_j - s_j$ , que se interpreta como la capacidad del ítem para separar adecuadamente a los examinados que dominan los atributos relevantes para dicha tarea respecto de los que no (George y Robitzsch, 2015). El valor de dificultad, por su parte, puede ser reinterpretado como  $\omega_{2j} = [g_j + (1 - s_j)]/2$ , que informa de la facilidad del ítem o probabilidad promedio de responder correctamente un ítem (George y Robitzsch, 2015).

Tabla 8. Parámetros de los ítems obtenidos mediante ajuste del modelo DINA.

Ítems	M	S	$g_j$	$s_j$	$\omega_{1j}$	$\omega_{2j}$	Ítems	M	S	$g_j$	$s_j$	$\omega_{1j}$	$\omega_{2j}$
p01a	0.81	0.15	0.75	0.13	0.12	0.81	p05	0.60	0.24	0.00	0.38	0.62	0.31
p01b	0.66	0.22	0.41	0.17	0.42	0.62	p06a	0.68	0.22	0.67	0.33	0.00	0.67
p01c	0.92	0.07	0.83	0.01	0.16	0.91	p06b	0.92	0.07	0.96	0.10	-0.06	0.93
p01d	0.74	0.19	0.46	0.08	0.46	0.69	p06c	0.70	0.21	0.75	0.33	-0.08	0.71
p01e	0.82	0.15	0.60	0.04	0.36	0.78	p06d	0.67	0.22	0.70	0.35	-0.04	0.68
p01f	0.86	0.12	0.69	0.02	0.29	0.84	p08	0.73	0.20	0.34	0.11	0.56	0.62
p01g	0.91	0.08	0.81	0.02	0.17	0.89	p09	0.45	0.25	0.45	0.54	0.00	0.46
p01h	0.68	0.22	0.33	0.09	0.59	0.62	p10a	0.38	0.24	0.41	0.62	-0.03	0.39
p01i	0.78	0.17	0.48	0.02	0.50	0.73	p10b	0.71	0.21	0.57	0.28	0.15	0.64
p01j	0.95	0.05	0.93	0.03	0.04	0.95	p10c	0.75	0.19	0.75	0.25	0.00	0.75
p01k	0.77	0.18	0.51	0.06	0.44	0.72	p11	0.54	0.25	0.12	0.44	0.44	0.34
p01l	0.92	0.07	0.81	0.01	0.18	0.90	p12a	0.93	0.07	0.77	0.02	0.21	0.87
p01m	0.67	0.22	0.29	0.07	0.64	0.61	p12b	0.68	0.22	0.22	0.18	0.60	0.52
p01n	0.75	0.19	0.50	0.07	0.44	0.71	p12c	0.66	0.22	0.19	0.20	0.62	0.49
p01o	0.78	0.17	0.56	0.07	0.37	0.75	p13a	0.77	0.18	0.43	0.21	0.36	0.61
p01p	0.86	0.12	0.72	0.05	0.23	0.84	p13b	0.87	0.11	0.72	0.12	0.16	0.80
p01q	0.67	0.22	0.25	0.07	0.68	0.59	p13c	0.53	0.25	0.36	0.46	0.18	0.45
p01r	0.66	0.22	0.24	0.07	0.69	0.59	p13d	0.27	0.20	1.00	0.74	-0.74	0.63
p02	0.31	0.21	0.05	0.58	0.37	0.24	p15a	0.78	0.17	0.66	0.20	0.15	0.73
p03	0.72	0.20	0.71	0.27	0.02	0.72	p15b	0.41	0.24	0.11	0.57	0.33	0.27
p04	0.65	0.23	0.15	0.31	0.54	0.42	p15c	0.67	0.22	0.07	0.28	0.65	0.39

La mayoría de los ítems denota dificultades moderadas a bajas. La dificultad juega un papel en la discriminación, pues ítems muy fáciles o difíciles tienden a discriminar poco y en consecuencia son menos informativos del nivel del rasgo (Ferrando, 1996) en

los modelos de variable latente continua. En los modelos de variable latente discreta, sin embargo, pueden ayudar a identificar patrones de respuestas de alguna clase concreta de sujetos.

Los valores de  $g_j$  indican la proporción de examinados que responden correctamente aunque no posean el dominio de todos los atributos relevantes para el ítem en cuestión. Los valores de  $g_j$  son bastante grandes, en algunos casos llegando a 1, como ocurre con el ítem p13d. En otros casos, como en los ítems p02 o p15c el nivel de conjetura es casi inexistente.

En el caso de  $s_j$ , a pesar de tener todos los atributos relevantes para un determinado ítem, sugiere en general una baja proporción de respuesta incorrecta cuando se tiene el dominio sobre los atributos cognitivos requeridos, con excepciones como en el ítem p02, p10a o p13d.

El valor de discriminación  $\omega_{1j}$ , depende de los dos parámetros anteriores, por lo tanto, si estos muestran valores bajos, se puede esperar un buen nivel de discriminación. No ocurre con todos los ítems y se aprecian incluso valores negativos. En este último caso, el ítem viola la condición de monotonicidad  $g_j < 1 - s_j$ , que asegura que la probabilidad de responder correctamente un ítem, cuando se dominan todos los atributos cognitivos y sin que haya distracción, es mayor que la probabilidad de responder correctamente mediante conjetura aun cuando se carezca del dominio sobre al menos uno de los atributos requeridos (George, Robitzsch, Kiefer, Groß y Ünlü, 2016). Los ítems con valores negativos deben ser eliminados del análisis.

La probabilidad promedio de responder correctamente el ítem  $\omega_{2j}$ , está vinculada a la discriminación, pues si el valor promedio es alto, entonces es esperable que el ítem no discrimine adecuadamente. Su inclusión ayuda a la interpretación de estos resultados. Nótese la relación con la media de los ítems.

Producto de este primer análisis, incluir ítems que no discriminen adecuadamente puede comprometer seriamente el ajuste de los modelos de diagnóstico cognitivo. En consecuencia, para proceder a estimar la bondad de ajuste de los modelos CDM se conservaron solo aquellos ítems con un nivel de discriminación mayor a 0.3 (George y

Robitzsch, 2015). El análisis de esta y las siguientes secciones se ejecuta sobre los 21 ítems que se conservan.

### *Índices de ajuste de los modelos*

El análisis de ajuste incluyó los índices relativos y absolutos comentados previamente. En la tabla 9 se muestran los índices para los modelos incluidos.

Tabla 9. Índices de ajuste relativos y absolutos para los modelos de diagnóstico cognitivo.

<b>Índices</b>	<b>DINA</b>	<b>DINO</b>	<b>RRUM</b>	<b>G-DINA</b>
Número de parámetros estimados de los ítems	42	42	112	266
Log-likelihood	-4702.521	-4723.926	-4540.060	-4412.947
AIC	9615.042	9657.851	9304.120	9357.894
BIC	10044.867	10087.676	9762.599	10446.784
MADcor	0.055	0.057	0.035	0.029
max $\chi^2$ *	17.898 (0.004)	26.852 (0.000)	9.791 (0.368)	9.086 (0.540)
SRMSR	0.071	0.073	0.046	0.038
MADRCOV	1.105	1.147	0.712	0.596

\* En max $\chi^2$  los valores-p están entre paréntesis

Los modelos DINA y DINO muestran índices inadecuados de bondad de ajuste, debido a que son los más restrictivos frente a los modelos más flexibles RRUM y G-DINA. RRUM es el que mejor bondad de ajuste muestra en los índices bayesianos, especialmente en BIC que penaliza la falta de parsimonia, no obstante, si comparamos los modelos G-DINA y RRUM en términos de devianza, se observa que una mejora estadísticamente significativa en la bondad de ajuste a favor de G-DINA:  $\chi^2 (154) = 254.223 (p < 0.01)$ .

### *Patrones de atributos obtenidos de acuerdo al modelo*

En la Tabla 10 se muestran los porcentajes de sujetos que presentan alguno de los 16 patrones de atributos. Solo se mantienen los patrones con al menos un 1% de respuestas para el modelo que G-DINA. Dichos patrones están representados por un vector de unos y ceros que indican el dominio o no de cada uno de los atributos que aparecen en las columnas de la matriz  $Q$ . Así por ejemplo, un examinado  $i$  que tenga un patrón de atributos igual a  $\alpha_i = \{1,0,0,0,0,1\}$  es alguien que domina los atributos código

de superficie y habilidades lingüísticas, pero no los atributos restantes que han sido descritos para la prueba CL-PT.

Tabla 10. Patrones de atributos en porcentajes estimados por el modelo G-DINA

Nº	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	G-DINA
1	0	0	0	0	0	0	1.45
2	0	0	1	0	1	0	2.63
3	0	1	1	0	0	0	2.11
4	0	1	1	0	1	0	10.56
5	0	1	1	0	1	1	3.8
6	1	0	0	0	0	1	1.83
7	1	0	0	1	1	1	2.01
8	1	0	1	0	0	1	2.47
9	1	0	1	0	1	1	5.76
10	1	1	0	0	0	1	1.35
11	1	1	0	1	1	0	1.49
12	1	1	0	1	1	1	8.69
13	1	1	1	0	0	1	8.64
14	1	1	1	0	1	0	1.03
15	1	1	1	0	1	1	32.39
16	1	1	1	1	1	1	3.97

En la Tabla 10 se aprecia que casi un tercio de las respuestas (32.4%) se corresponden con el patrón  $\alpha_i = \{1,1,1,0,1,1\}$ , es decir, son examinados que dominan todos los atributos excepto el cuarto: modelo de situación. El segundo patrón más frecuente es  $\alpha_i = \{0,1,1,0,1,0\}$  que indica no dominio del código de superficie y de las habilidades lingüísticas, además del desconocimiento del modelo de situación. Son pocos los sujetos que dominan todos los atributos un 3.97% y menos aún las respuestas que implican desconocimiento de todos los atributos 1.45%.

### Discusión

La evaluación de la bondad de ajuste de los modelos muestra que el modelo G-DINA es el que mejor ajusta a los datos de la prueba de comprensión evaluada, mientras que DINO es el modelo que denota el peor ajuste, lo que revela que un modelo compensatorio no es adecuado. Del mismo modo, tampoco un modelo no compensatorio en el que sea preciso dominar todos y cada uno de los atributos, como DINA, sería el más inadecuado.

En cuanto a los porcentajes por patrón de atributos, G-DINA explora un mayor número de parámetros que el resto de modelos, por su mayor flexibilidad para ajustarse al comportamiento evidenciado de los sujetos, agrupando a los examinados en 16 patrones destacados. Se aprecia que un porcentaje importante denota dominio sobre las habilidades código de superficie y base de texto y que dominar la primera parece incluir la segunda. Del mismo modo, modelo de situación parece estar vinculado a la aparición de las instrucciones retóricas. Por el contrario, un porcentaje relevante de los sujetos no domina el atributo modelo de situación, que está ausente en gran parte de los patrones con mayor frecuencia de sujetos.

En cuanto a comprensión, atributos como base de texto, inferencia o instrucciones retóricas tienen sustento en estudios previos (Hemmati, Baghaei y Bemani, 2016; Li, Hunter y Lei, 2015) lo que fortalece la noción de que comprender un texto requiere el concurso de estas habilidades de base. Otros atributos como las habilidades lingüísticas están representados y son importantes para alcanzar un desempeño adecuado en la evaluación, pero requieren mejor sustento teórico, pues al implicar procesos cognitivos generales, su identificación y definición varía considerablemente en la literatura (Guarneros, y Vega, 2014). Por último, en la medición está poco representada el modelo de situación. Dado que estos atributos emergen por retroajuste, no significa que este atributo sea menos relevante, sino que el instrumento no lo incluyó como habilidad significativa. Teóricamente hay una relación entre el modelo de situación, el proceso de inferencia y los conocimientos previos, pero entre los patrones rescatados parece haber una necesaria relación con las instrucciones retóricas. Quienes las dominan también dominan el modelo de situación.

Las proyecciones de este trabajo pasan por profundizar la aplicación de estos modelos en poblaciones más diversas y de mayor magnitud, así como estudiar con mayor detalle los factores que pueden llevar a este tipo de comportamiento en las evaluaciones de comprensión. Una limitación del estudio es la posible generalización de los resultados a otras poblaciones, lo que requeriría replicar el estudio en otras muestras y con otros instrumentos de comprensión lectora.

## **Estudio 3: Análisis de la Comprensión Lectora en función de la condición pública o privada del centro educativo**

### **Introducción**

#### *Escuelas vulnerables y efectos sobre la comprensión lectora*

Matemáticas y Lenguaje son las áreas fundamentales del currículo educativo. Tanto para la OECD (2012, 2014, 2016) como para TIMMS (Mullis, Martin, Foy y Hooper, 2016) o las mediciones a nivel nacional realizadas en Chile representadas por el Simce (Agencia Calidad de Educación, 2012, 2016). Todas ponen el foco en estas áreas clave, pues su desarrollo está vinculado al futuro progreso académico y técnico-profesional de los estudiantes (Agencia de Calidad de la Educación, 2015).

La UNICEF afirma que “entre las claves de los buenos resultados que tienen raíz en las prácticas pedagógicas en el aula, destacan clases motivadoras y cercanas a la vida cotidiana de los alumnos, con propósitos claros, estructura, ritmo, alto aprovechamiento del tiempo, exigentes, con predominio de refuerzos positivos, evaluación y retroalimentación regular del aprendizaje de los alumnos, dan prioridad a la comprensión lectora, expresión de ideas, razonamiento lógico, autonomía y creatividad de los niños” (Fondo de Naciones Unidas para la Infancia, 2004, p. 14). En la cita se menciona explícitamente dos factores esencialmente cognitivos: la comprensión lectora y el razonamiento lógico. De ellas, la comprensión lectora ha sido tema recurrente de investigación por ser una de las variables a la base del logro académico en muchas disciplinas: “En los alumnos, la aplicación de la evaluación estandarizada de comprensión lectora en enseñanza básica permite medir los avances y vacíos de un área que impacta fuertemente los logros alcanzables en las asignaturas restantes.” (Fondo de Naciones Unidas para la Infancia, 2004, p. 297). La comprensión lectora es la base del aprendizaje posterior del estudiante y se encuentra presente en todos los contenidos y/o disciplinas que este enfrente (Agencia de Calidad de la Educación, 2015).

En efecto, la comprensión lectora tiene un sitio destacado en la explicación del logro académico, como siguen demostrándolo numerosas investigaciones particularmente en niños de educación primaria y secundaria (Bunning, Schraw, Norby y Ronning, 2004; Cunningham y Stanovich, 2007; Fuentes, 2009; Gaviria, Biencinto y

Navarro, 2009; Gil, 2009; Kendeou, van den Broek, Helder y Karlsson, 2014; Marchant, Lucchini y Cuadrado, 2007; Núñez y Donoso, 2000; Ramos, 2006).

Aunque puede ser aventurado indicar que el nivel de comprensión lectora es un factor que explique la desigualdad a futuro, lo cierto es que esta noción proviene del hecho por el cual el éxito o fracaso del aprendizaje depende del ambiente en que el niño se desenvuelve, visión conocida como teoría sociocultural (McIntyre, 2011). En este enfoque, el aprendizaje depende de las interacciones que establece el estudiante con otros, en el contexto de su base histórica y cultural. Es el ambiente el que interactúa con el niño e influye en el desarrollo de su aprendizaje. Dicho ambiente está conformado por muchos agentes, pero de ellos es la escuela y la familia las que ocupan roles principales (Dehqan y Samar, 2014; McIntyre, 2011).

Este trabajo se centra en la escuela, considerándolo el contexto social que favorece el que los estudiantes puedan abordar el aprendizaje de la lectura como un proceso social formalizado, en el que las interacciones profesor-estudiante juegan un rol fundamental. A ello se le suma las expectativas del profesor, que al parecer también tienen un importante efecto sobre el logro del alumno: Docentes que piensan que sus estudiantes deben ser rescatados de la influencia negativa de sus comunidades y/o familias, o bien docentes que creen que alumnos en condiciones de pobreza tienen pocas probabilidades de salir adelante (Dehqan y Samar, 2014).

Es interesante notar que esta clase de expectativas ha llevado, en el contexto anglosajón americano, a que los estudiantes de escuelas pobres reciban educación estructurada, programada, mientras que los de escuelas pudientes reciben instrucción enfocada en procesos cognitivos complejos como el pensamiento crítico o la creatividad (McIntyre, 2011). Por su parte, en escuelas vulnerables hay menores recursos, las metodologías al uso pueden ser inadecuadas o atrasadas y los profesores estar menos capacitados para tratar las dificultades presentes en la enseñanza de la lecto-escritura. El efecto pares también debe tomarse en cuenta pues en establecimientos vulnerables, con una composición social homogénea, los estudiantes pueden denotar un vocabulario reducido con pocas posibilidades de mejora, lo que también tiene efecto sobre el aprendizaje inicial de la lecto-escritura (Compton et al., 2009; Marchant, Lucchini y Cuadrado, 2007).

Esta diferencia entre escuelas públicas y privadas es un hecho bien establecido en la literatura a nivel latinoamericano. En el estudio de Diuk y Ferroni (2012) -llevado a cabo en escuelas argentinas- se habla del Efecto Mateo en lectura. Consiste en un fenómeno en que aquellos que ingresan al sistema educativo con mayores conocimientos, desarrollan tasas de aprendizaje superiores a los que denotan al inicio menos habilidades. En el caso de estudiantes de nivel socioeconómico medio y alto, estas desventajas son rápidamente compensadas por una compleja red de apoyo familiar, atención de profesores particulares, psicopedagogos, entre otros. En el caso de los estudiantes de nivel socioeconómico bajo, estos apoyos no suelen estar presentes y por tanto es esperable que no logren avanzar en su aprendizaje. Sotelo et al. (2012) han establecido que las diferencias de dominio lector son notorias entre niños y niñas peruanas de zonas desfavorecidas. Se piensa que esta desventaja está asociada con las menores oportunidades de explorar el lenguaje escrito en estos contextos. Para Fonseca et al. (2014) los niños que provienen de sectores desfavorecidos denotan un rendimiento significativamente más bajo en comprensión lectora, impactando fuertemente en el rendimiento académico y en su desarrollo cognitivo. Uno de los aspectos de su investigación aportó evidencia en torno a la importancia de la adquisición de vocabulario para compensar en algún grado esta desventaja comparativa, aunque también se fundamentó la capacidad para realizar inferencias, comprender la estructura textual y el dominio de mecanismos reguladores: planificación, supervisión y evaluación. En Chile también se han realizado investigaciones en torno a estas diferencias en escuelas de distinta condición socioeconómica (Donoso-Díaz, 2014; Mizala y Torche, 2012). Una de ellas ha sido elaborada previamente en la línea de este estudio por Cáceres-Serrano y Alvarado-Izquierdo (2017), quienes corroboraron la importancia de la diferencia socioeconómica, que representa infraestructura y materiales inadecuados y menos capacitación de los profesores en las escuelas vulnerables. Todos estos aspectos tienen un efecto sobre el aprendizaje y en consecuencia sobre el desarrollo de la comprensión.

### ***Diferencias en habilidades cognitivas mediante funcionamiento diferencial del ítem***

En este estudio se quiere avanzar en la determinación de los efectos que la condición socioeconómica tiene no ya sobre la comprensión de lectura, sino más específicamente, en las habilidades cognitivas que sustentan dicha comprensión. Para ello

se requiere el uso de metodologías multivariantes que permitan establecer de forma flexible las variables subyacentes responsables de los resultados en comprensión.

Los modelos de diagnóstico cognitivo (CDM, de su acrónimo en inglés) pueden ayudar con este fin al permitir un análisis detallado de las habilidades a nivel de los ítems de un test. Los CDM son modelos de variable latente como la Teoría de Respuesta al Ítem (TRI) pero en los que se definen "clases de sujetos" en función de sus patrones de dominio de habilidades. En este sentido son modelos que en lugar de intentar estudiar las características de los ítems, lo que intentan es clasificar a los sujetos según sus respuestas. Es decir, se trata de un cambio de enfoque desde el ítem al sujeto (se pueden denominar perspectivas "centradas en el sujeto", en lugar de en los ítems). Los distintos modelos CDM determinan los parámetros que hay que ajustar (igual en un modelo TRI) y la bondad de ajuste determina qué modelo es el que más se adecúa a las habilidades a la base del desempeño (George, Robitzsch, Kiefer, Groß y Ünlü, 2016).

Los CDM son modelos confirmatorios, puesto que requieren confirmar empíricamente una propuesta teórica. Para ello, un grupo de expertos debe determinar – sobre la base de alguna teoría cognitiva apropiada para el fenómeno evaluado- cuáles son las habilidades o atributos que el examinado debe dominar para poder responder correctamente a cada ítem (Rupp, Templin y Henson, 2010). Una vez se han identificado los atributos cognitivos que condicionan el desempeño en cada ítem, a continuación se construye una matriz Q que asocia las habilidades y los ítems, indicando qué atributos deben o no dominarse para responder correctamente los ítems. En la matriz Q, el dominio del atributo se representa mediante el valor 1, mientras que el no dominio con el valor 0, en aquellas celdas que cruzan un atributo con un ítem (ver tabla 11) (de la Torre, 2009; de la Torre y Chiu, 2016).

Varios son los estudios que se han efectuado en el contexto de la lectura aplicando diversos modelos cognitivos disponibles (Buck, Tatsuoka y Kostin, 1997; Hemmati, Baghaei y Bemani, 2016; Ravand, Barati y Widhiarso, 2013). Estos estudios han permitido reconocer que el llamado modelo G-DINA es el que mejor ajusta una evaluación de comprensión como esta, lo que es posible porque es el modelo menos restrictivo. La representación de G-DINA es esencialmente un modelo estadístico, pero para exponerlo en términos resumidos es necesario tener presente que en la matriz Q, por

cada ítem  $j$ , es posible definir un patrón de dominio de atributos. Si para responder un ítem se requieren tres habilidades, entonces puede haber  $2^3$  patrones u ocho clases latentes  $\alpha_k = \{(0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (0,1,1), (1,0,1), (1,1,1)\}$ , desde no dominar ninguna de las habilidades (lo que implica una probabilidad nula o muy baja de responder correctamente el ítem  $j$ ), dominar solo una, dominar dos o bien las tres habilidades (que representa la mayor probabilidad de responder correctamente el ítem  $j$ ). En el modelo G-DINA, para cada uno de los ítems, el modelo divide las clases latentes en  $2^{K_j^*}$  grupos latentes, donde  $K_j^* = \sum_{k=1}^K q_{jk}$  representa el número de atributos requeridos para el ítem  $j$ . Se define  $\alpha_{lj}^*$  como el vector de atributos reducido cuyos elementos son los atributos requeridos para el ítem donde  $l = 1, \dots, 2^{K_j^*}$ , que como se ha señalado representa el número total de grupos latentes o patrones de atributos latentes posibles. En el modelo G-DINA, la probabilidad de que un examinado perteneciente a la clase latente  $l$ , con un patrón de atributos  $\alpha_{lj}^*$  responda correctamente un ítem  $j$  se denota como  $P(X_j = 1 | \alpha_{lj}^*) = P(\alpha_{lj}^*)$  (García, Olea y de la Torre, 2014). En consecuencia, la probabilidad de responder correctamente el ítem depende del patrón de atributos que lo condiciona y el dominio que tiene el examinado de dicho patrón.

El uso de estos modelos cognitivos representa un avance respecto de las tradicionales y conocidas comparaciones de resultados entre escuelas públicas y privadas (Muñoz-Chereau y Thomas, 2015; Orellana y Merino, 2013, Correa, Parro y Reyes, 2014), pues en este caso más allá de una diferencia en el desempeño total obtenido por los estudiantes en comprensión de lectura, es posible conocer las diferencias entre las aptitudes cognitivas básicas que determinan dicho desempeño.

Siguiendo la lógica de los procedimientos confirmatorios de análisis de la invarianza de la medida (Blanco-Canitrot, Alvarado y Ondé, 2018; Meredith, 1993; Millsap, 2011), una vez establecida la bondad de ajuste del modelo es posible estudiar si los distintos grupos que forman parte de la población muestran patrones de resultados diferentes y en tal caso si existe un problema de funcionamiento diferencial del ítem (DIF). En los modelos CDM, se entiende el DIF como un efecto donde las probabilidades de responder correctamente un ítem varía entre examinados que tienen el mismo patrón de dominio de atributos cognitivos, pero que pertenecen a distintos grupos (Hou, de la Torre y Nandakumar, 2014; Ravand y Robitzsch, 2015).

Una vez comprobada la bondad de ajuste al modelo CDM seleccionado, en nuestro caso G-DINA, se procede a estudiar las diferencias entre los patrones de respuesta de los estudiantes en establecimientos públicos y privados aplicando un análisis de DIF, esto es, se evalúa si las probabilidades de respuesta a los ítems dependen únicamente del nivel en el rasgo latente (Fidalgo, 1996), lo que implica ausencia de DIF, o bien si la pertenencia a alguno de los grupos supone una mayor o menor probabilidad de respuesta correcta, en cuyo caso dicha pertenencia supondría una ventaja ilegítima (por ejemplo, una mayor familiaridad de los ítems) y no debería aplicarse esta prueba para comparar entre estos grupos. Si como se espera, el test es invariante se puede evaluar la razón de las diferencias en comprensión lectora relacionándolas con el mayor o menor dominio de los atributos o habilidades necesarias para la comprensión lectora.

El objetivo principal de esta investigación es analizar las diferencias en comprensión lectora a nivel de atributos o habilidades cognitivas entre centros públicos (establecimientos municipales) y privados chilenos (particulares y particulares subvencionados) una vez evaluada la presencia de DIF en el test utilizado.

## **Método**

### ***Participantes***

Corresponde a una muestra no probabilística intencional, representando estudiantes provenientes de escuelas públicas y privadas. El rango de edad estaba entre nueve y 10 años y todos cursaban 4° Básico. Se analizaron las respuestas de 434 estudiantes de ellos 256 eran niñas (59%) y 178 eran niños (41%). El número de examinados de escuelas públicas fue de 129 (29.7%), mientras que los de escuelas privadas fue de 305 sujetos (70.3%).

### ***Diseño***

El estudio es cuantitativo, ex post facto retrospectivo, donde solo se miden variables en un único grupo al que se le ha evaluado en un período corto de tiempo (Montero y León, 2014).

### ***Instrumento***

Se aplicó la prueba Comprensión Lectora y Producción de Textos (CL-PT) elaborada por Medina y Gajardo (2014). CL-PT es una prueba comercial estandarizada

de amplio uso en Chile para ser aplicada en estudiantes de los cinco a los 14 años y que ha mostrado tener una adecuada fiabilidad y validez predictiva respecto a su capacidad para diferenciar entre estudiantes de distinto nivel socioeconómico. Del total de 47 ítems, agrupados en 16 tareas o ejercicios de comprensión lectora se analizaron los 21 ítems que en estudios previos han mostrado niveles de discriminación aceptables (mayores a 0.30). Para facilitar su identificación se mantuvo la numeración original en las tablas de resultados. Cabe hacer notar que a pesar de la reducción en el número de ítems, las diferencias en el desempeño entre escuelas públicas y privadas se mantuvo, lo que se corroboró mediante Prueba  $t$  ( $t_{(441)}=2.51$ ,  $p=0.013$ ,  $d=0.260$ ) cumpliéndose el supuesto de homocedasticidad para el contraste ( $F_{(1)}=0.423$ ,  $p=0.516$ ).

### ***Procedimiento***

Se tuvo acceso a los establecimientos escolares mediante solicitud directa y obteniendo permiso escrito y/o verbal de las direcciones de cada institución. En cada caso ayudantes en la investigación aplicaron los cuestionarios originales impresos, dando las instrucciones pertinentes que habían sido previamente ensayadas con el investigador a cargo, pues se intentó controlar que todos los estudiantes recibieran la misma retroalimentación.

El primer paso del procedimiento consistió en determinar la composición de la matriz  $Q$ , en donde se establecen los atributos cognitivos que requiere cada ítem para ser respondido correctamente (Ravand y Robitzsch, 2015). La construcción de la matriz implicó la definición de los atributos a partir de estudios previos en comprensión lectora (Buck, Tatsuoka y Kostin, 1997; Chen y de la Torre, 2014; Hemmati, Baghaei y Bemani, 2016), elaborándose un modelo de atributos acorde al contenido del instrumento (Ravand y Robitzsch, 2015), definiéndose seis atributos cognitivos: a)  $\alpha_1$ : Código de superficie (proceso básico de decodificación y reconocimiento de palabras, Riffo, 2016), b)  $\alpha_2$ : Base de texto (Procesamiento oracional en el que los elementos son estructurados mediante micro-proposiciones generando una microestructura textual o coherencia local, Riffo, 2016), c)  $\alpha_3$ : Inferencia (Operación que ayuda a agregar la información faltante de aquello que no expresa el nivel de superficie, Vieiro y Gómez, 2004), d)  $\alpha_4$ : Modelo de situación (Nivel de representación que hace uso del conocimiento previo y el proceso de inferencia, Gernsbacher y Kaschak, 2013), e)  $\alpha_5$ : Instrucciones retóricas (ayudan a la

comprensión del texto, son marcadores presentes en el texto que deben ser interpretados, Sánchez, 1996) y f)  $\alpha_6$ : Habilidades lingüísticas (Involucran procesos cognoscitivos y son parte del desarrollo de un individuo, permitiendo la comprensión y expresión de mensajes, elaboración de ideas, reflexión y solución de problemas. Hay evidencia que en contextos empobrecidos el desempeño en este dominio suele ser deficiente (Guarneros y Vega, 2014; Guevara, García, López, Delgado y Hermosillo, 2007).

Una vez definidas los atributos, tres jueces independientes todos académicos del área de lingüística, evaluaron la asignación de atributos a los ítems de comprensión lectora. Con las evaluaciones de los jueces, se revisa la asignación inicial en aquellos en los que había discrepancia hasta alcanzar unanimidad. Finalizado este proceso quedó definida la matriz  $Q$  (ver Tabla 11).

Tabla 11. Matriz Q de la prueba de comprensión de lectura CL-PT.

Nº	Ítems	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$
1	P01b	1	1	1	0	1	0
2	P01d	1	1	1	0	1	0
3	P01e	1	1	1	0	1	0
4	P01f	1	1	1	0	1	0
5	P01h	1	1	1	0	1	0
6	P01i	1	1	1	0	1	0
7	P01k	1	1	1	0	1	0
8	P01m	1	1	1	0	1	0
9	P01n	1	1	1	0	1	0
10	P01o	1	1	1	0	1	0
11	P01q	1	1	1	0	1	0
12	P01r	1	1	1	0	1	0
13	P04	1	1	1	0	0	0
14	P05	1	0	1	0	0	0
15	P08	1	0	0	1	0	1
16	P11	1	1	1	0	0	0
17	P12b	0	0	0	0	0	1
18	P12c	0	0	0	0	0	1
19	P13a	1	1	0	0	0	0
20	P15b	1	1	1	0	0	0
21	P15c	1	1	1	0	0	0

En la tabla 11 se aprecia que código de superficie ( $\alpha_1$ ) es una habilidad básica para la mayor parte de los ítems, mientras que modelo de situación ( $\alpha_4$ ) solo está presente en un único ítem. Son también relevantes base de texto ( $\alpha_2$ ) y proceso inferencial ( $\alpha_3$ ).

Para evaluar la bondad de ajuste al modelo G-DINA y realizar los análisis de funcionamiento diferencial del ítem (DIF) se utilizó el entorno de programación R (R Core Team, 2018) y específicamente el paquete CDM de George, Robitzsch, Kiefer, Groß y Ünlü (2016).

## Resultados

Antes de proceder al estudio de las diferencias entre atributos, y la presencia de DIF, se obtuvo la bondad de ajuste del test. El mejor ajuste se logró con el modelo G-DINA que mostró un buen ajuste en los diferentes índices evaluados: a) MADcor=0.029, que corresponde a la media de las desviaciones absolutas entre las correlaciones observadas y predichas para cada par de ítems  $i$  y  $j$ , esperando valores iguales o menores a 0.05 (DiBello, Roussos y Stout, 2006; Lei y Li, 2016); b)  $\max\chi^2$  ( $\chi^2(266)=9.09$ ,  $p=0.540$ ), este índice no debe ser significativo ( $p > 0.05$ ) ya que de lo contrario indicaría una violación de la independencia estadística entre pares de ítems (Hu, Miller, Huggins-Manley y Chen, 2016); c) La raíz cuadrada de la media de residuos estandarizados SRMSR=0.038, que corresponde a un promedio de correlaciones residuales, fruto de la diferencia entre la correlación muestral y la correlación esperada entre cada par de ítems  $i$  y  $j$  (Maydeu-Olivares y Joe, 2014), siendo valores aceptables aquellos que sean iguales o menores a 0.05; d) Por último, se obtuvo el valor más bajo para G-DINA en el índice MADRCOV=0.596, que es un indicador de tamaño del efecto que compara covarianzas observadas y predichas, del que se espera un valor bajo; Mientras menor, mejor es el ajuste (Ravand y Robitzsch, 2015).

En la Tabla 12 se observa que la media (M) tomó valores entre 0.41 y 0.86 con una alta variabilidad entre los examinados, como muestra la Desviación Típica (DT). Junto con estos descriptivos, se informa de los parámetro conjetura, o probabilidad de acertar sin tener el dominio cognitivo adecuado ( $g_j$ ), distracción o desliz, o la probabilidad de no acertar a pesar de tener el dominio cognitivo adecuado ( $s_j$ ), discriminación  $\omega_{1j} = 1 - g_j - s_j$  y dificultad  $\omega_{2j} = [g_j + (1 - s_j)]/2$ , todos obtenidos a partir del ajuste del

modelo cognitivo. Su análisis es relevante ya que la presencia de conjetura o distracción disminuye la discriminación y dificultad de los ítems. Los resultados denotan escasa conjetura -a excepción de algunos ítems relacionados con la pregunta 1- y bajo desliz salvo en los ítems 11 y 15b. Dado que la discriminación y dificultad dependen de la conjetura y distracción, se aprecian adecuados niveles en ambos parámetros.

Tabla 12. Parámetros estimados Modelo G-DINA.

Ítems	M	DT	$g_j$	$s_j$	$\omega_{1j}$	$\omega_{2j}$	Ítems	M	DT	$g_j$	$s_j$	$\omega_{1j}$	$\omega_{2j}$
p01b	0.66	0.22	0.41	0.17	0.42	0.62	p01q	0.67	0.22	0.25	0.07	0.68	0.59
p01d	0.74	0.19	0.46	0.08	0.46	0.69	p01r	0.66	0.22	0.24	0.07	0.69	0.59
p01e	0.82	0.15	0.60	0.04	0.36	0.78	p04	0.65	0.23	0.15	0.31	0.54	0.42
p01f	0.86	0.12	0.69	0.02	0.29	0.84	p05	0.60	0.24	0.00	0.38	0.62	0.31
p01h	0.68	0.22	0.33	0.09	0.59	0.62	p08	0.73	0.20	0.34	0.11	0.56	0.62
p01i	0.78	0.17	0.48	0.02	0.50	0.73	p11	0.54	0.25	0.12	0.44	0.44	0.34
p01k	0.77	0.18	0.51	0.06	0.44	0.72	p12b	0.68	0.22	0.22	0.18	0.60	0.52
p01m	0.67	0.22	0.29	0.07	0.64	0.61	p12c	0.66	0.22	0.19	0.2	0.62	0.49
p01n	0.75	0.19	0.50	0.07	0.44	0.71	p13a	0.77	0.18	0.43	0.21	0.36	0.61
p01o	0.78	0.17	0.56	0.07	0.37	0.75	p15b	0.41	0.24	0.11	0.57	0.33	0.27
							p15c	0.67	0.22	0.07	0.28	0.65	0.39

En la Tabla 13 se disponen los resultados del contraste entre los grupos comparados con objeto de evaluar DIF de manera global. Para llevar a cabo el contraste se aplica el test de Wald (Li y Wang, 2015), que brinda un valor  $\chi^2$  para contrastar la hipótesis nula de que los parámetros del ítem son iguales para el grupo focal y de referencia (Hou, de la Torre y Nandakumar, 2014). Un valor no significativo en el contraste permite aceptar  $H_0$ .

El grupo focal corresponde a los establecimientos públicos o municipales y el grupo de referencia, que se asume alcanza mejor los objetivos de aprendizaje en comprensión, corresponde a los establecimientos privados.

Tabla 13. Análisis general de DIF entre colegios públicos y privados.

Ítem	$\chi^2$	gl	Valor p	Ítem	$\chi^2$	gl	Valor p
p01b	13.11	16	0.66	p01q	2.17	16	1.00
p01d	5.83	16	0.99	p01r	7.09	16	0.97
p01e	17.56	16	0.35	p04	19.77	8	0.01
p01f	7.83	16	0.95	p05	17.29	4	<0.01
p01h	7.37	16	0.97	p08	5.05	8	0.75
p01i	11.18	16	0.80	p11	7.20	8	0.52
p01k	6.53	16	0.98	p12b	0.14	2	0.93
p01m	8.43	16	0.93	p12c	0.07	2	0.97
p01n	21.07	16	0.18	p13a	6.24	4	0.18
p01o	9.64	16	0.88	p15b	3.94	8	0.86
				p15c	1.20	8	1.00

En la tabla 13 se observa que no hay DIF generalizado que evidencie las diferencias en comprensión lectora. La única excepción corresponde a los ítems p04 y p05, en que el valor del test de Wald fue estadísticamente significativo ( $p < 0.05$ ) y que parece estar relacionado con la presencia de una mayor acierto por azar en estudiantes de centros públicos que solo dominan el atributo código de superficie pero no los atributos base de texto y/o inferencia.

En la Tabla 14 se pueden ver los patrones de dominio de habilidades con probabilidades iguales o superiores al 2%. Considerando todos los ítems y todos los atributos que subyacen al buen desempeño en la prueba, se observa que el 39% de los estudiantes pertenecientes a centros privados dominan todos los atributos, mientras que en los centros públicos solo un 13% obtiene el mismo patrón de dominio, siendo los patrones más comunes en centros públicos aquellos en los que están implicados conjuntamente los atributos Código de Superficie, Inferencia e Instrucciones Retóricas.

Tabla 14. Patrones de dominio de competencias más frecuentes

Nº	Patrones	Privado	Público	Diferencia
1	001010	0.02	0.00	0.02
2	011010	0.07	0.02	0.05
3	111010	0.00	0.02	-0.02
4	110110	0.02	0.00	0.02
5	101001	0.02	0.00	0.02
6	111001	0.05	0.00	0.05
7	101011	0.00	0.02	-0.02
8	011011	0.03	0.01	0.02
9	100111	0.04	0.01	0.03
10	110111	0.09	0.01	0.08
11	111011	0.00	0.13	-0.13
12	111111	0.39	0.13	0.26

Finalmente, se realizó un análisis estadístico de diferencia de proporciones sobre las probabilidades de poseer los seis atributos en función del tipo de establecimiento. Los resultados se disponen en la tabla 15.

Tabla 15. Probabilidad de dominio de los atributos en centros privados y públicos

Atributos	Privado	Público	Diferencias	Z
$\alpha_1$	0.80	0.67	0.13	2.83*
$\alpha_2$	0.75	0.71	0.04	0.84
$\alpha_3$	0.75	0.80	-0.05	-1.05
$\alpha_4$	0.68	0.67	0.01	0.23
$\alpha_5$	0.78	0.84	-0.06	-1.58
$\alpha_6$	0.81	0.51	0.30	6.42**

\*p<0.01, \*\*p<0.001

En la Tabla 15 se presentan las diferencias entre los atributos cognitivos que explican el desempeño en comprensión lectora del alumnado de centros públicos y privados, observándose diferencias estadísticamente significativas en los atributos  $\alpha_1$  Código de Superficie y  $\alpha_6$  Habilidades Lingüísticas a favor de los centros privados.

## Discusión

Los resultados obtenidos representan una contribución a la discusión sobre tres aspectos principales; en primer lugar, en el presente trabajo se ha mostrado que el modelo G-DINA tiene una adecuada bondad de ajuste sobre un conjunto seleccionado de ítems de la prueba CL-PT, prueba muy popular en la medición de la comprensión en Chile. A partir de este análisis fue posible identificar y establecer los atributos cognitivos específicos responsables del desempeño de los examinados en esta competencia. El resultado concuerda con otros estudios que han mostrado la utilidad de los CDM y en concreto de G-DINA para explicar la comprensión lectora (Buck, Tatsuoka y Kostin, 1997; Hemmati, Baghaei y Bemani, 2016; Ravand, Barati y Widhiarso, 2013).

En segundo lugar, una vez establecida la bondad de ajuste a G-DINA se evaluó si los atributos responsables de las respuestas a los ítems de comprensión lectora dependen o no del tipo de colegio, público o privado, en el que están escolarizados. Los resultados muestran que, en general, no hay DIF, es decir que las diferencias en ejecución en las tareas de comprensión lectora evaluadas no se deben a un posible sesgo por el tipo de colegio o factores vinculados, como podría ser un entrenamiento específico, recursos, metodologías de enseñanza o expectativas docentes distintas. La ausencia de DIF implica que los estudiantes que tienen un mismo patrón de atributos cognitivos tienen la misma probabilidad de dar respuesta correcta a los ítems, independientemente del tipo de colegio en el que estén escolarizados. Hay no obstante, dos excepciones con los ítems 4 y 5. En ellos aparece algún efecto del DIF. Al revisar con detalle el contenido y la expresión de esas preguntas, se aprecia un mayor compromiso con el primer atributo Código de Superficie, pues al comparar los patrones este parece ser un determinante propio de establecimientos públicos. Con base en los fundamentos teóricos, hay evidencia para pensar que la diferencia puede estar en el acceso al vocabulario, pues este tiene una fuerte relación con el contexto social y con el conocimiento previo que los estudiantes tienen antes de enfrentar un proceso educativo formal. Es posible incluso aventurar que la diferencia presente en estas preguntas podría incluso tener relación con la concomitancia entre Código de Superficie e Inferencia producto de su presencia conjunta en mayores probabilidades de respuesta correcta.

En tercer lugar, una vez evaluado el DIF, se obtienen las diferencias concretas calculadas entre dominios cognitivos, lo que revela que es justamente Código de Superficie y Habilidades Lingüísticas las dos habilidades que denotan una ventaja para escuelas privadas. Especialmente llamativo resulta que un 39% en centros privados dominen todos los atributos para comprensión lectora frente a solo un 13% en centros públicos. Nuevamente Código de Superficie es determinante en la diferenciación a favor de privados, lo que habla de la asociación del vocabulario con las oportunidades sociales que se tienen para explorar la riqueza del lenguaje dependiendo del entorno en el que los estudiantes se encuentren inmersos. Es claro que el contexto social de escuelas vulnerables está menos enriquecido (Fonseca et al., 2014; Sotelo et al., 2012). Las Habilidades Lingüísticas representan un dominio cognitivo más complejo y se constituye como otro factor diferenciador entre estudiantes de centros públicos y privados, lo que puede estar relacionado con las condiciones socioculturales familiares, el nivel educacional de las madres y las expectativas y actitudes de sus miembros hacia la formación (Guevara, García, López, Delgado y Hermosillo, 2007).

La mejora en comprensión lectora es clave para un adecuado rendimiento académico y el acceso a mejores oportunidades en la vida, por ello es fundamental que se estudie y trabaje en mejorar las competencias que determinan el rendimiento lector de los estudiantes. La mejora en las habilidades o atributos que determinan la comprensión es una cuestión tanto del docente como del alumnado y en este último caso, el rendimiento del alumno está estrechamente relacionado con el nivel socioeconómico (Cáceres-Serrano y Alvarado-Izquierdo, 2017). La mejora en el dominio de estas habilidades también está relacionada con el apoyo que el estudiante recibe en casa, y por tanto para superar estas limitaciones en ambientes desfavorecidos, la escuela y sus docentes tendrán que realizar un esfuerzo que compense estas desventajas. Se espera además, que la identificación detallada de los problemas concretos que están a la base de la comprensión lectora contribuya a elaborar planes docentes mejor focalizados y más efectivos.

## DISCUSIÓN FINAL

Para referir los resultados y proceder a su discusión, es necesario tener siempre presente la pregunta de investigación que dio origen a este trabajo: ¿Cuáles son las diferencias en el dominio de la comprensión de lectura entre los centros educativos privados y públicos de la zona central de Chile y cuáles son los atributos cognitivos que mejor explican dicha diferencia?

La dimensión de la pregunta requiere más que una única respuesta y es por ello que el trabajo se compuso de tres estudios que intentaron dar cuenta de los matices presentes en el cuestionamiento inicial. En primer lugar, en la pregunta está presente la intención de establecer diferencias entre centros privados y públicos en el desempeño escolar, precisamente en comprensión de lectura (conviene aclarar que durante el desarrollo de este trabajo se ha hablado en términos de privado-público para simplificar la algo más compleja relación de titularidad de las escuelas en Chile, en las que se reconoce al menos la presencia de centros privados pagados o particulares pagados, privados subvencionados o con apoyo de fondos públicos y municipales o de carácter eminentemente público). La diferencia en desempeño académico (en comprensión lectora u otra área) se ha atribuido a diversas variables pero siempre en la esfera social y económica, como lo corroboran los resultados de muchos estudios hechos en este ámbito y que ya tienen algunos años, evidenciando que el tema ha concitado el interés de los investigadores desde hace un buen tiempo (Bellei, 2007; Borsdorf, Hidalgo y Vidal-Koppmann, 2016; Cabalin, 2012; Campos, 2010; Córdova, Rojas y Azócar, 2016; Donoso, 2012; Donoso, 2012; Donoso y Hawes, 2002; Elacqua, 2012; Matus y Quezada, 2007; Mayol, 2011; Mizala y Romaguera, 1998; Mizala, Romaguera y Ostoić, 2004; Pérez, 2009; Valenzuela, Bellei y de los Ríos, 2013). En segundo lugar, en la pregunta también está contenida la pretensión de incorporar una medida psicológica específica relativa a habilidades cognitivas vinculadas al desempeño de los estudiantes, situación que en este trabajo es recogida a partir de la aplicación de los denominados modelos de diagnóstico cognitivo, lo que permite avanzar un poco más en la evaluación de los resultados, ofreciendo un marco explicativo más detallado del rendimiento escolar. Se ha considerado el rendimiento en comprensión de lectura, en primer lugar por ser esta una

habilidad a la base de los aprendizajes de muchas otras áreas o disciplinas académicas y en segundo lugar porque es una de las variables más frecuentemente medidas en diversas evaluaciones nacionales e internacionales (Brunning, Schraw, Norby y Ronning, 2004; Cunningham y Stanovich, 2007; Fuentes, 2009; Gaviria, Biencinto y Navarro, 2009; Gil, 2009; Kendeou, van den Broek, Helder y Karlsson, 2014; Marchant, Lucchini y Cuadrado, 2007; Núñez y Donoso, 2000; Ramos, 2006; Ugarriza, 2006)

Entre ambas inquietudes presentes en la pregunta está la pretensión implícita de tender un puente entre las diferencias que se observan en el nivel social más amplio y a las cuales se les atribuyen las diferencias observadas entre escuelas y los efectos psicológicos que estas diferencias pueden generar a nivel cognitivo entre estudiantes de las escuelas públicas y privadas. El marco ofrecido por los modelos de diagnóstico cognitivo contribuye al logro de esta pretensión.

Los tres estudios independientes que se han reunido en esta tesis se hacen cargo, cada uno de modo parcial, de las dos grandes inquietudes presentes en el problema de investigación. Con objeto de abordar de modo secuencial y con mayor precisión lo señalado es que cada uno de estos estudios responde a uno de los objetivos de investigación trazados y sus resultados se exponen siguiendo la misma lógica.

El primer objetivo pretendió evaluar la influencia de variables contextuales y económicas sobre el desempeño académico en comprensión de lectura entre estudiantes de escuelas públicas y privadas. El estudio llevado a cabo hizo uso de datos secundarios obtenidos de la aplicación del instrumento de medición de nivel nacional (Simce) efectuada cada año por la Agencia de Calidad de Educación (Agencia Calidad de la Educación, 2015).

Los hallazgos de este primer trabajo permitieron establecer que el factor socioeconómico está relacionado claramente con los desempeños en comprensión lectora. De modo consistente, las escuelas que de forma progresiva muestran menores niveles socioeconómicos, denotan un decreciente desempeño en la variable criterio, siendo el peor resultado aquel que obtienen las escuelas más pobres (Cáceres-Serrano y Alvarado-Izquierdo, 2017).

Al resultado anterior, hay que agregar los obtenidos al incorporar otros factores de contexto social. De estos, fueron estadísticamente significativos: a) El nivel de educación de la madre, b) el nivel de educación del padre y c) el grado de expectativas positivas que los padres o apoderados tienen de su hijo/pupilo.

Estos resultados son consistentes con los hallazgos obtenidos en estudios previos. En el trabajo de Donoso y Hawes (2002) se ratifican las claras desigualdades educativas producto de las inequidades sociales. El componente socioeconómico, según los autores, tiene una enorme influencia sobre los resultados Simce. Aunque no es sorprendente, los autores destacan lo lejos que Simce está de superar estas diferencias y que son justamente estos resultados lo que lleva a un éxodo de estudiantes desde el sistema público al sistema privado. McEwan (2003) por su parte, identifica en la realidad chilena la educación de los padres, especialmente de la madre, como un elemento de relevancia en los efectos sobre la variable dependiente (lenguaje y matemáticas). La misma situación es descrita por Correa, Parro y Reyes (2014). Para Elacqua (2012) las escuelas públicas denotan claras desventajas con respecto a las privadas, evidenciado también mayor grado de segregación. Señala además que la política de “vouchers” tiende a segregar aún más las escuelas y a generar mayores desventajas entre ellas. Para Mizala y Torche (2012) hay una marcada estratificación entre escuelas producto de los resultados que estas obtienen. El nivel socioeconómico de las familias, así como la educación parental, las expectativas familiares y el número de libro en casa, tienen un efecto diferenciador entre las escuelas. Las evidencias en este tema son concordantes (Osses, Bellei y Valenzuela, 2015; Pérez, Betancort y Cabrera, 2011; Thieme, Tortosa-Ausina, Prior y Gempp, 2012; Somers, McEwan y Willms, 2004).

En este primer estudio no se consideran directamente todas las variables, sin embargo, es teóricamente sustentable que el nivel socioeconómico es una variable proxy que enmascara la influencia de muchos factores contextuales que inciden en el desempeño de los estudiantes. La peor infraestructura, la falta de recursos materiales y humanos, las menores expectativas y visión de corto plazo de los padres o apoderados, así como la pobreza propia de un entorno menos diverso, menos rico en estímulos y oportunidades, tiene sin duda su efecto en los desempeños de los estudiantes y en particular en algo tan crítico y fundamental para el desarrollo de otras disciplinas como ocurre con la comprensión de lectura.

El factor socioeconómico y los factores específicos vinculados representan una oportunidad de comprender que las diferencias entre escuelas están presentes y condicionan en buena medida el rendimiento de los escolares. Las mediciones de lenguaje y también de matemáticas evidencian los obstáculos que derivan de la pobreza y de las condiciones desfavorables del contexto en el que se encuentran los estudiantes. La atribución social del desempeño es el punto de partida para concebir que los efectos van más allá y pueden observarse también a niveles cognitivos.

El segundo objetivo de esta tesis planteó establecer los atributos cognitivos que determinan el desempeño en comprensión lectora sobre estudiantes con las mismas características de aquellos que fueron evaluados en el primer estudio y que pertenecen tanto a centros públicos como privados. Para dar cuenta de estos atributos y a pesar de la abundante literatura en torno a los factores que explican la comprensión, en este trabajo se favoreció la aplicación de los emergentes modelos de diagnóstico cognitivo, que permiten determinar las habilidades cognitivas que explican el nivel de comprensión que manifiestan los estudiantes y hacerlo por cada uno de los ítems, reconociendo el concurso simultáneo de más de una habilidad, si corresponde (Rupp, Templin y Henson, 2010; Rupp y Templin, 2008).

Esta fortaleza de los modelos de diagnóstico cognitivo representó un marco de trabajo que fue aprovechado para indagar y profundizar en los procesos cognitivos involucrados en la medición, superando el tradicional modelo de evaluación sumativa que solo otorga evidencias de desempeño contra ciertos estándares de resultados esperados (de la Torre y Minchen, 2014).

Al aplicar modelos de diagnóstico cognitivo y evidenciar directamente los atributos desde la tarea de comprensión propia de la evaluación, no solo se obtiene información útil para orientar mejor el accionar de formadores e instructores, sino que además se ahorra en el número de mediciones a efectuar. En efecto, la noción de que ciertas dimensiones cognitivas influyen en la mejor o peor comprensión de lectura no solo ha sido planteada muchas veces, sino que además es esperable. La desventaja es que estos planteamientos requieren medir, por un lado, el grado de comprensión y por otro, las dimensiones con herramientas o instrumentos adicionales. Al hacerlo sin embargo, hay un divorcio entre lo que efectivamente el examinado ejecuta en la tarea de

comprensión y lo responde en los instrumentos que evalúan los posibles factores. Distintos instrumentos para las variables comprometidas capitalizan los errores de medición que pueden estar presentes en todas esas medidas. Por ello, los modelos de diagnóstico cognitivo ofrecen la posibilidad de extraer -desde las mismas tareas de comprensión- aquellas habilidades que explican su mejor o peor desempeño. Este proceder favorece la posibilidad de incorporar modelos o teorías de carácter cognitivo que representan campos relevantes para el constructo en cuestión.

Dentro del marco de trabajo del segundo estudio, los jueces pudieron retroajustar la prueba de comprensión lectora con base en el modelo situación de Kintsch (Kintsch, 1998; van Dijk y Kintsch, 1983), que presume de la comprensión como una interacción entre las expresiones textuales explícitas y el conocimiento de mundo de quien comprende (Hirsch, 2003; McIntyre, 2011; Sugate, Schaughency y Reese, 2013; Tijero, 2009) Las dimensiones que emergieron del retroajuste y que fueron incorporadas a la matriz Q correspondieron a: a) Código de superficie: Que corresponde a un nivel representacional y de procesamiento en el que se mantiene en memoria de modo exacto lo que el texto entrega. Se mantiene la forma (rasgos formales) y la representación lingüística. El tipo de proceso es el básico de decodificación y reconocimiento de palabras. Implica solo reproducción de contenidos y uso de recuerdo. (Riffo, 2016); b) Base de texto: O nivel de procesamiento oracional, donde los elementos de una oración leída por un usuario son estructurados mediante microproposiciones que están conectadas a nivel local, generando un microestructura textual o coherencia local. La base de texto permite realizar inferencias básicas que permiten conectar una idea con otra. Es un nivel representacional que permite parafrasear el texto a través del uso de macroestrategias. (Hemmati, Baghaei y Bemani, 2016; Riffo, 2016); c) Inferencia: Esto es, una operación que ayuda a agregar la información faltante de aquello que no expresa el nivel de superficie. Las inferencias participan del proceso de comprensión y provienen de conocimiento general, dominios específicos y representaciones basadas en experiencias personales o lecturas previas. Tienen lugar durante la construcción del modelo de situación (Hemmati, Baghaei y Bemani, 2016; Li, Hunter y Lei, 2015; Vieiro y Gómez, 2004); d) Modelo de situación: O nivel representacional que permite asumir aprendizaje de lo leído. Hace uso del conocimiento previo y del proceso de inferencia. Representa un momento de integración entre el conocimiento explícito y el modelo de mundo que tiene

quien comprende. Este atributo da nombre al modelo cognitivo que lo enmarca (Gernsbacher y Kaschak, 2013; Piacente, 2009); e) Instrucciones retóricas: Representan ayudas a la comprensión del texto. Estas ayudas no se refieren a la temática tratada sino al texto en sí y a la relación que se invita a tener con él. Son marcadores en el texto que deben ser interpretados como instrucciones de procesamiento (Sánchez, 1996); y f) Habilidades lingüísticas: Que son parte del desarrollo de un individuo, permitiendo la comprensión y expresión de mensajes, elaboración de ideas, reflexión y solución de problemas (Guarneros, y Vega, 2014). Son consideradas generales y por tanto más allá de la modalidad escrita.

Con la aplicación de un modelo cognitivo y la definición de los atributos cognitivos que gobiernan el desempeño de los ítems de comprensión, se pudo aportar mayor validez a una prueba (CL-PT) que es de uso masivo a nivel nacional, que ha informado con precisión de sus propiedades psicométricas (Medina y Guajardo, 2014) y que ve enriquecida la interpretación de sus resultados en función de estos aportes.

El segundo estudio, además de contribuir con la elucidación de los atributos cognitivos a los cuales responden los ítems que componen la CL-PT, enfocó su atención en determinar cuál de los modelos cognitivos representaba una mejor opción de ajuste a los datos obtenidos. El número de modelos es ingente, y trabajar con todos ellos es poco oficioso, con varios artículos tratando el tema en el ámbito de la comprensión lectora (Buck, Tatsuoka y Kostin, 1997; Hemmati, Baghaei y Bemani, 2016; Jang, 2009; Ravand, Barati y Widhiarso, 2013; Ravand, 2016; von Davier, 2005) la elección de los modelos fue primero de carácter teórico.

Una vez se seleccionaron los modelos más adecuados desde el punto de vista conceptual (a saber DINA, DINO, RRUM y G-DINA) se procedió al análisis de la bondad de ajuste, lo que permitió establecer que G-DINA fue el mejor modelo para los datos obtenidos con la medida de comprensión de lectura, mientras que DINO mostró el peor ajuste de los utilizados. G-DINA es un modelo más flexible, y favorece la exploración de un mayor número de parámetros que los restantes modelos. Hay respaldo teórico para este hallazgo en algunos de los trabajos en comprensión donde se le ha aplicado (Hemmati, Baghaei y Bemani, 2016; Li, Hunter y Lei, 2015; Ravand, 2015). Si bien estos trabajos comparten el mejor ajuste del modelo comentado, no es posible comparar los

atributos entre estudios pues estos dependen del modelo teórico utilizado por los investigadores, aunque en algunos casos es posible reconocer algunas habilidades, tales como Base de Texto, Inferencia o Instrucciones Retóricas, o aptitudes vinculadas como Vocabulario o Reconocimiento de Palabras que están presentes en estos previos estudios (Hemmati, Baghaei y Bemani, 2016; Li, Hunter y Lei, 2015; Ma y Meng, 2014).

Hay que destacar que entre los hallazgos, uno de los causó sorpresa fue la ausencia de representación adecuada del Modelo de Situación. No es menos relevante, pero su presencia en el instrumento de comprensión fue marginal, aun con la importancia que tiene en esta habilidad, de acuerdo al modelo Kintsch. No obstante, es necesario tener presente que la génesis del CL-PT proviene de mucho antes que este modelo teórico fuese de relevancia entre los lingüistas. El sostén teórico de la prueba es otro, pero aquí no fue considerado por los jueces.

Mediante el análisis obtenido se pudieron conservar 16 patrones de atributos cognitivos que configuraron los perfiles de los estudiantes que respondieron al instrumento de comprensión. Casi un tercio de estos denotaron un perfil de dominio en que estaban presentes todos los atributos cognitivos, con excepción del denominado Modelo de Situación (32.4% de los examinados). El segundo perfil más frecuente entre los estudiantes evidenció la ausencia de dominio de Código de Superficie y de las denominadas Habilidades Lingüísticas, además de ya mencionado Modelo de Situación (10.6% de los examinados). El perfil con el dominio de todos los atributos solo reúne a una porción pequeña de los estudiantes (apenas un 4% del total de examinados), pero es aún menor si el perfil implica no dominio en ninguno de los atributos considerados (menos del 2% de los examinados).

El tercer y último objetivo de investigación de la tesis intentó abordar conjuntamente las dos pretensiones previas. Este objetivo buscaba: Determinar las diferencias en el patrón de atributos cognitivos propios de la comprensión lectora en centro de educación pública y privada.

A la base de la comparación se tenía la constancia previa de haber establecido diferencias claras en los niveles de comprensión lectora entre centros públicos y privados. Estas diferencias son a nivel nacional con otro instrumento, la prueba de comprensión de lectura del Sistema de Medición de la Calidad de la Educación, Simce. La tendencia

tiende a mostrar poco cambio entre escuelas a lo largo del tiempo. Al momento del cierre de esta tesis, las bases de datos disponibles seguían evidenciando la misma realidad.

Para dar cuenta de las diferencias, en primer lugar se llevó a cabo una evaluación de la presencia de funcionamiento diferencial del ítem, DIF, obtenida en el contexto del ajuste del modelo cognitivo utilizado, G-DINA (Hou, de la Torre y Nandakumar, 2014). No se halló presencia relevante de DIF entre los ítems, de modo que se procedió a comparar los resultados obtenidos según la titularidad del centro.

Tanto la base de datos así como los atributos cognitivos de la evaluación de comprensión son los mismos que los referidos en el estudio dos.

El resultado de las diferencias entre centros permitió identificar que los atributos Código de Superficie y Habilidades Lingüísticas mostraron mejores desempeños entre las escuelas privadas. Código de superficie, vinculado al mejor manejo de vocabulario muestra consistencia con las evidencias teóricas que se han recopilado en este trabajo, pues un mejor manejo de palabras está relacionado con un mayor conocimiento de mundo, un contexto enriquecido y mejores oportunidades de explorar y reconocer otras realidades sociales y educativas (Compton et al., 2009; Fonseca et al., 2014; Hirsch, 2003; Luo, Pace, Masek, Hirsh-Pasek y Golinkoff, 2016; Marchant, Lucchini y Cuadrado, 2007; Sotelo et al., 2012; Sugate, Schaughency y Reese, 2013).

Por su parte, las denominadas Habilidades Lingüísticas refieren a un atributo cognitivo de mayor complejidad, pero que también determina una diferencia de importancia entre alumnos de escuelas privadas y públicas. Se conjetura que esta situación también puede ser referida a condiciones socioculturales familiares, al nivel educacional de las madres (o de los padres, según el caso) y las expectativas y actitudes de dichos padres hacia la formación y las oportunidades educativas (Guevara, García, López, Delgado y Hermosillo, 2007).

A través de estos tres objetivos específicos se ha intentado responder, es de esperar que con fortuna, a la pregunta de investigación que guio este trabajo. Si bien son proyectos separados, se piensa que hay suficiente evidencia teórica y suficientes nexos empíricos entre los trabajos, como para vincular los factores sociales con las diferencias inter-escuelas llegando incluso a sostener que las diferencias afectan a niveles cognitivos muy

específicos, y que los efectos socioeconómicos y contextuales entrañan mayor relevancia que la sola privación de oportunidades educativas, pues dichas oportunidades en el fondo son cruciales para gestar desde temprana edad las habilidades o atributos necesarios para llegar a tener un dominio suficiente del lenguaje y la comprensión de lectura. Siendo esta última aptitud absolutamente determinante de nuevos aprendizajes formales en la escuela.

Al cierre, es necesario ponderar el hecho de haber tenido que trabajar con muestras separadas, siendo lo idóneo hacer las indagaciones en una muestra única en la que el investigador haya tenido el control total de su recolección. Sin embargo, a pesar de llegar instancias judiciales, fue imposible liberar los ítems de la prueba Simce, debiendo trabajar con bases de datos primarias, pero que solo contenían los resultados finales. Este hecho no perjudicó en absoluto lo que se pretendió realizar con el primer estudio, pues las implicaciones de este primer análisis no iban más allá, pero sin duda impidió darle la continuidad requerida al no poder establecer los atributos cognitivos propios de dicha evaluación, debiendo el investigador tomar muestras particulares, para darle continuidad al proyecto. En este sentido, una segunda limitación correspondió a la muestra utilizada para ajustar el modelo G-DINA que afecta a su generalizabilidad.

Una tercera limitación que atañe a la mayoría de las investigaciones realizadas hasta la fecha con los CDM, es la adaptación de pruebas que inicialmente no fueron pensadas ni diseñadas desde un modelo teórico que sustente la aplicación de los CDM. El problema de esta estrategia es que la prueba en principio no ha sido construida para medir esos atributos cognitivos identificados *a posteriori*, lo que conduce a la aplicación de procedimientos de depuración para la eliminación de aquellos ítems que no ajustan a los modelos, incluso aunque sus propiedades psicométricas originales, estimadas a partir de modelos psicométricos clásicos, fueran aceptables. Es obvio que para el uso de los CDM en pruebas ya construidas se requiere de esta estrategia de “retroajuste”. La otra alternativa es construir nuevas pruebas que sean diseñadas desde un principio bajo las bases de los CDM, opción que si bien es mucho más costosa representa la situación ideal. Es probable que sea el escenario futuro una vez los investigadores tomen consciencia de las ventajas de estos modelos.

Las proyecciones del trabajo, a juicio del investigador se remiten a dos áreas que de hecho dan cuerpo al trasfondo de este trabajo: a) la oportunidad de desarrollar medidas de comprensión de lectura (sin perjuicio de otros ámbitos) con base en estos modelos, permitiendo administrar un instrumento que sirva de apoyo directo a la formación de los estudiantes, pues se tiene la opción de tratar con los atributos específicos que condicionan el desempeño comprensivo lector y b) la oportunidad de nivelar a los estudiantes de contextos empobrecidos atendiendo a las reales y concretas necesidades que deben ser cubiertas con el fin de contribuir a aminorar la segregación y mejorar sus estándares de aprendizaje.

## CONCLUSIONES

La aplicación de los CDM en el ámbito de la evaluación de la comprensión lectora es una poderosa herramienta que en un proceso de retroalimentación debe permitir tanto la mejora en los modelos teóricos sobre la comprensión lectora y sus componentes esenciales, así como la mejora en los procesos de evaluación de dicha competencia. En el presente trabajo se constata y replica el mejor desempeño de los niños y niñas escolarizados en centros de titularidad privada respecto a aquellos escolarizados en centros de titularidad pública. Como se muestra mediante la aplicación de análisis multinivel, el mayor porcentaje de la varianza en los resultados en comprensión son explicados por el centro, y solo una pequeña parte de los resultados en los tests de comprensión lectora depende de factores personales. Pero para profundizar en lo que hace que los mejores centros destaquen por sobre los demás, aquellos de titularidad privada accesibles solo a las clases sociales con un alto nivel de ingresos, es preciso ir más allá de la mera puntuación final que muestra estas diferencias estadísticamente significativas entre centros. Es preciso -como se ha intentado en esta tesis- conocer dónde están las diferencias, mediante el uso de modelos teóricos sólidos y la evaluación mediante los CDM que permite establecer el origen de las diferencias y en consecuencia, aplicar medidas correctoras que permitan disminuir la brecha entre centros.

## REFERENCIAS BIBLIOGRÁFICAS

- Abad, F., Olea, J., Ponsoda, V. y García, C. (2010). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Agencia de Calidad de la Educación (2012). *Resultados TIMSS 2011*. Santiago, CL: Agencia de Calidad de Educación.
- Agencia de Calidad de Educación (2013a). *PISA Evaluación de las competencias lectoras para el siglo XXI*. Santiago, CL: Agencia de Calidad de Educación.
- Agencia de Calidad de Educación (2013b). *Recursos en el hogar y en el establecimiento. Apuntes sobre la calidad de la educación*. Santiago, CL: Agencia de Calidad de educación.
- Agencia de Calidad de la Educación (2014). *Informe nacional resultados Chile PISA 2012*. Santiago, CL: Agencia de Calidad de Educación. Recuperado de: [http://archivos.agenciaeducacion.cl/Informe\\_Nacional\\_Resultados\\_Chile\\_PISA\\_2012.pdf](http://archivos.agenciaeducacion.cl/Informe_Nacional_Resultados_Chile_PISA_2012.pdf)
- Agencia de Calidad de la Educación (2015a) *Factores que influyen en la motivación por la lectura y su relación con logros de aprendizajes e indicadores de desarrollo personal y social*. Recuperado de [http://www.agenciaeducacion.cl/wp-content/uploads/2016/02/Estudio\\_Factores\\_que\\_influyen\\_en\\_motivacion\\_lectora.pdf](http://www.agenciaeducacion.cl/wp-content/uploads/2016/02/Estudio_Factores_que_influyen_en_motivacion_lectora.pdf)
- Agencia de Calidad de la Educación (2015b). Informe nacional de resultados TERCE. Santiago, CL: Agencia de Calidad de la Educación. Recuperado de: [http://archivos.agenciaeducacion.cl/TERCE\\_informefinal.pdf](http://archivos.agenciaeducacion.cl/TERCE_informefinal.pdf)
- Agencia de Calidad de la Educación (2015c). *Informe técnico Simce 2013*. Santiago, CL: Agencia de Calidad de Educación.
- Agencia de Calidad de la Educación (2016a). *Informe técnico Simce 2014*. Santiago, CL: Agencia de Calidad de Educación.

- Agencia de Calidad de la Educación (2016b). *Resultados educativos 2016* [Presentación PowerPoint]. Recuperado de [http://archivos.agenciaeducacion.cl/ResultadosNacionales2016\\_.pdf](http://archivos.agenciaeducacion.cl/ResultadosNacionales2016_.pdf)
- Aguirre, M., Castro, M. y Adasme, A. (2009). Factores que inciden en el rendimiento escolar en Chile. *Estudios de Economía Regional. Centro de Estudios de Opinión Ciudadana, Universidad de Talca*, 2(4), 2-4.
- Almond, P. J., Cameto, R., Johnstone, C. J., Laitusis, C., Lazarus, S., Nagle, K., Sato, E. (2009). White paper: *Cognitive interview methods in reading test design and development for alternate assessments based on modified academic achievement standards (AA-MAS)*. Menlo Park, CA: SRI International.
- American Educational Research Association (AERA) (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Alonso, J. y Mateos, M. (1985) *Comprensión lectora: Modelos, entrenamiento y evaluación*. *Infancia y Aprendizaje*, 31-32, 5-19.
- Alvarado, J. M., Puente, A., Fernández, M. P. y Jiménez, V. (2015) Análisis de los componentes en la adquisición de la lectura en castellano: una aplicación del modelo logístico lineal. *Suma Psicológica*, 22(1), 45-52. <http://dx.doi.org/10.1016/j.sumpsi.2015.05.006>
- Atria, F. (2009) ¿Qué es educación pública? *Revista de Estudios Sociales*, 117, 45-69
- Baghaei, P. y Kubinger, K. D. (2015). Linear logistic test modeling with R. *Practical Assessment, Research y Evaluation*, 20(1), 1-11.
- Baker, F. B. y Kim, S. (2017). *The basics of item response theory using R*. New York, NY: Springer.
- Balladares, J., Marshall, C. y Griffiths, Y. M. (2016) Socio-economic status affects sentence repetition, but not non-word repetition, in Chilean pre-schoolers. *First Language*, 36(3), 338-351. <https://doi.org/10.1177/0142723715626067>

- Bates, D., Machler, M., Bolker, B. y Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Barrada, J. R. (2012). Tests adaptativos informatizados: una perspectiva general. *Anales de Psicología*, 28(1), 289-302.
- Bellei, C. (2007). Expansión de la educación privada y mejoramiento de la educación en Chile. Evaluación a partir de la evidencia. *Revista Pensamiento Educativo*, 40(1), 285-311.
- Blanco-Canitrot, D., Alvarado, J.M. y Ondé, D. (2018) Consequences of disregarding metric invariance on diagnosis and prognosis using psychological tests. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00167>
- Brandt, N. (2010). Chile: Climbing on giants' shoulders: Better schools for all Chilean children. *Economics department working papers*, 784. París, FR: Organisation for Economic Co-operation and Development (OECD). <http://dx.doi.org/10.1787/5kmd41g7x9g0-en>
- Bravo, J. (2011). SIMCE: Pasado, presente y futuro del sistema nacional de evaluación. *Estudios Públicos*, 123, 189-211.
- Brunning, R. H., Schraw, G. J., Norby, M. N. y Ronning, R. R. (2005). *Psicología cognitiva y de la instrucción*. Madrid: Pearson Educación.
- Boora, S., Pasiphol, S. y Tangdhanakanond, K. (2015). Development of cognitive diagnostic testing on basic arithmetic operation. *Procedia - Social and Behavioral Sciences*, 191, 769-772. <https://doi.org/10.1016/j.sbspro.2015.04.533>
- Borsdorf, A., Hidalgo, R. y Vidal-Koppmann, S. (2016). Social segregation and gated communities in Santiago de Chile and Buenos Aires. A comparison. *Habitat International*, XXX, 1-10. <http://dx.doi.org/10.1016/j.habitatint.2015.11.033>
- Bravo, D., Contreras, D. y Sanhueza, C. (1999). *Rendimiento educacional, desigualdad y brecha de desempeño privado/público: Chile 1982-1997. Documento de trabajo*. Santiago, CL: Departamento de Economía Universidad de Chile.

- Bruning, R., Schraw, G., Norby, M. y Ronning, R. (2004). *Cognitive psychology and instruction*. Upper Saddle River, NJ: Prentice Hall.
- Buck, G., Tatsuoka, K. K. y Kostin, I. (1997). The subskills of reading: Rule-space analysis of a multiple-choice test of second language reading comprehension. *Language Learning*, 47(3), 423-466. <https://doi.org/10.1111/0023-8333.00016>
- Buck, G., Van Essen, T., Tatsuoka, K., Kostin, I., Lutz, D. y Phelps, M (1998). *Development, selection and validation of a set of cognitive and linguistic attributes for the SAT I verbal: Analogy section*. Princeton: Educational Testing Service.
- Cabalin, C. (2012). Neoliberal education and student movements in Chile: inequalities and malaise. *Policy Futures in Education*, 10(2), 219-228. <https://doi.org/10.2304/pfie.2012.10.2.219>
- Cáceres-Serrano, P. y Alvarado-Izquierdo, J. M. (2017) The effect of contextual and socioeconomic factors on reading comprehension levels. *Modern Journal of Language Teaching Methods*, 7(8), 76-85.
- Cádiz, J. (2006). Reanalyzing fourth grade math student achievement in Chile: Applying hierarchical linear models (HLMS). *RELIEVE*, 12(1), 75-91.
- Campos, J. (2010). *Las desigualdades educativas en Chile*. Bogotá: Foro Latinoamericano de Políticas Educativas.
- Canales, M., Bellei, C. y Orellana, V. (2016). ¿Por qué elegir una escuela privada subvencionada? Sectores medios emergentes y elección de escuela en un sistema de mercado. *Estudios Pedagógicos*, XLII(3), 89-109.
- Canet-Juric, L. Andrés, M. L. y Ané, A. (2005) *Modelos teóricos de comprensión lectora. Relaciones con prácticas pedagógicas de enseñanza y aprendizaje*. Trabajo presentado en las XII Jornadas de Investigación y Primer Encuentro de Investigadores en Psicología del Mercosur. Buenos Aires: Universidad de Buenos Aires.
- Casassus, J. (2003). La escuela y la (des)igualdad. *Revista de Educación*, 304, 12-17.

- Chall, J. S. y Jacobs, V. A. (1983). Writing and reading in the elementary grades: Developmental trends among low SES children. *Language Arts*, 60(5), 617-626.
- Chen, H. y Chen, J. (2016). Retrofitting non-cognitive-diagnostic reading assessment under the generalized DINA model framework. *Language Assessment Quarterly*, 13(3), 218-230. <https://doi.org/10.1080/15434303.2016.1210610>
- Chen, J. y de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, 37(6), 419-437. <http://dx.doi.org/10.1177/0146621613479818>
- Chen, J. y de la Torre, J. (2014). A procedure for diagnostically modeling extant large-scale assessment data: The case of the programme for international student assessment in reading. *Psychology*, 5(18), 1967-1978. <http://dx.doi.org/10.4236/psych.2014.518200>
- Chiu, C. y Seo, M. (2009) Cluster analysis for cognitive diagnosis: An application to the 2001 PIRLS 137. En EA-ETS Research Institute (Ed.), *Issues and methodologies in large-scale assessments* (pp. 137-160). Princeton, NJ: IERInstitute.
- Chiu, C. (2013). Statistical refinement of the Q-matrix in cognitive diagnosis. *Applied Psychological Measurement*, 37(8), 598-618. <http://dx.doi.org/10.1177/0146621613488436>
- Cisterna, F. (2007). Reforma educacional, capital humano y desigualdad en Chile. *Horizontes Educativas*, 12(2), 43-50.
- Compton, D. L., Elleman, A. M., Olinghouse, N. G., Lawrence, J., Bigelow, E., Gilbert, J. K. y Davis, G. N. (2009) The influence of in-text instruction on declarative knowledge and vocabulary learning in struggling readers. How IQ confounds the story. En R. K. Wagner, C. Schatschneider y C. Phythian-Sence (Eds.) *Beyond decoding the behavioral and biological foundations of reading comprehension* (pp. 46-71), New York, NY: Guilford.
- Córdova, C., Rojas, K. y Azócar, J. (2016). Selección de alumnos y co-pago como factores de segregación escolar: Presentación de tres casos de estudio. *Psicoperspectivas*, 15(1), 90-104.

<https://doi.org/10.5027/PSICOPERSPECTIVASVOL15-ISSUE1-FULLTEXT-642>

- Correa, J., Inostroza, D., Parro, F., Reyes, L. y Ugarte, G. (2013). *The effects of vouchers on academic achievement: Evidence from Chile's conditional voucher program*. Santiago, CL: Ministry of Finance of Chile.
- Correa, J. A., Parro, F. y Reyes, L. (2014). The effects of vouchers on school results: Evidence from Chile's targeted voucher program. *Journal of Human Capital*, 8(4), 351-398. <https://doi.org/10.1086/679282>
- Corvalán, J., Elacqua, G. y Salazar, F. (2009). *El sector particular subvencionado en Chile. Tipologización y perspectivas frente a las nuevas regulaciones*. Santiago, CL: FONIDE.
- Cox, C. (2012). Política y políticas educacionales en Chile 1990-2010. *Revista Uruguaya de Ciencia Política*, 21(1), 13-42.
- Cronbach, L. J. y Meehl, P. E. (1955) Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <http://dx.doi.org/10.1037/h0040957>
- Cunningham, A. E. y Stanovich, K. E. (2007). Los efectos de la lectura en la mente. *Estudios Públicos*. 108, 207-228.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294-304. <http://dx.doi.org/10.1037/0893-3200.19.2.294>
- Deane, P. y Song, Y. (2014). A case study in principled assessment design: Designing assessments to measure and support the development of argumentative reading and writing skills. *Psicología Educativa*, 20(2), 99-108. <http://dx.doi.org/10.1016/j.pse.2014.10.001>
- De Boeck, P., Wilson, M. y Acton, G. S. (2005). A conceptual and psychometric framework for distinguishing categories and dimensions. *Psychological Review*, 112(1), 129-158. <https://doi.org/10.1037/0033-295X.112.1.129>

- Dehqan, M. y Samar, R. G. (2014). Reading comprehension in a sociocultural context: Effect on learners of two proficiency levels. *Procedia*, 98, 404-410.  
<https://doi.org/10.1016/j.sbspro.2014.03.433>
- de la Torre, J. (2008). An empirically based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, 45, 343-362. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34(1), 115-130.  
<https://doi.org/10.3102/10769986073094>
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179-199. <https://doi.org/10.1007/s11336-011-9207-7>
- de la Torre, J. y Chiu, C. (2016) A general method of empirical Q-matrix validation. *Psychometrika*, 81(2), 253-273. <https://doi.org/10.1007/s11336-015-9467-8>
- de la Torre, J. y Minchen, N. (2014). Cognitively diagnostic assessments and the cognitive diagnosis model framework. *Psicología Educativa*, 20(2), 89-97.  
<https://doi.org/10.1016/j.pse.2014.11.001>
- Davoudi, M. y Hashemi, H. R. (2015) Critical review of the models of reading comprehension with a focus on situation models. *International Journal of Linguistics*, 7(5), 172-187. <http://dx.doi.org/10.5296/ijl.v7i5.8357>
- DiBello, L. V., Roussos, L. A. y Stout, W. (2006) Review of cognitively diagnostic assessment and a summary of psychometric models. En C.R. Rao y S. Sinharay (Eds.) *Handbook of Statistics*. 26, 979-1030. Oxford, UK: Elsevier.  
[https://doi.org/10.1016/S0169-7161\(06\)26031-0](https://doi.org/10.1016/S0169-7161(06)26031-0).
- Diuk, B. y Ferroni, M. (2012). Dificultades de lectura en contextos de pobreza: ¿un caso de Efecto Mateo? *Revista Semestral da Associação Brasileira de Psicologia Escolar e Educacional*. 16(2), 209-217
- Dogan, E. y Tatsuoka, K (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational*

*Studies in Mathematics*, 68(3), 263–272. <https://doi.org/10.1007/s10649-007-9099-8>

Dong, Y. y Peng, C. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(222), 1-17. <https://doi.org/10.1186/2193-1801-2-222>

Donoso, S. (2006). Financiamiento y gestión de la educación pública chilena de los años 90. *Cadernos de Pesquisa*, 36(127), 151-172. <https://doi.org/10.1590/S0100-15742006000100007>

Donoso, S. y Hawes, G. (2002). Eficiencia escolar y diferencias socioeconómicas: A propósito de los resultados de las pruebas de medición de la calidad de la educación en Chile. *Educacao e Pesquisa*, 28(2), 25-39.

Donoso-Díaz, S. (2014). La escuela pública y la promoción de la mejora social. Reflexiones a partir de Latinoamérica. *Educar*, 30, 139-155. <http://dx.doi.org/10.5565/rev/educar.694>

Dreyer, L. G. y Katz, L. (1992). An examination of "The Simple View of Reading". *Haskins Laboratories Status Report on Speech Research*, 112, 161-166.

Elacqua, G. (2007). *Enrollment practices in response to market incentives: Evidence from Chile*. Princeton, NJ: Princeton University Press.

Elacqua, G. (2012). The impact of school choice and public policy on segregation: Evidence from Chile. *International Journal of Educational Development*, 32(3), 444-453. <https://doi.org/10.1016/j.ijedudev.2011.08.003>

Embretson, S.E. y Wetzel, C. D. (1987) Component latent trait models for paragraph comprehension tests. *Applied Psychological Measurement*, 11(2), 175-193. <https://doi.org/10.1177/014662168701100207>

Embretson, S. y Gorin, J. (2001). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement*, 38(4), 343-368. <https://doi.org/10.1111/j.1745-3984.2001.tb01131.x>

Embretson, S. E. y Yang, X. (2006). Multicomponent latent trait models for complex tasks. *Journal of Applied Measurement*, 7(3), 335-350.

- Embretson, S. E. y Yang, X. (2012). A multicomponent latent trait model for diagnosis. *Psychometrika*, 78(1), 14-36. <https://doi.org/10.1007/s11336-012-9296-y>
- Eyzaguirre, B. y Fontaine, L. (1999). ¿Qué mide realmente el SIMCE? *Centro de Estudios Públicos*, 75, 107-161.
- Fernández-Alonso, R., Suárez-Álvarez, J. y Muñiz, J. (2012). Imputación de datos perdidos en las evaluaciones diagnósticas educativas. *Psicothema*, 24(1), 167-175.
- Ferrando, P. J. (1996). Evaluación de la unidimensionalidad de los ítems mediante análisis factorial. *Psicothema*, 8(2), 397-410.
- Fidalgo, A. M. (1996). *Funcionamiento diferencial de los ítems*. Madrid: Universitas.
- Filzmoser, P., Garrett, R. G. y Reimann, C. (2005). Multivariate outlier detection in exploration geochemistry. *Computers y Geosciences*, 31, 579-587. <https://doi.org/10.1016/j.cageo.2004.11.013>
- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359-374. [https://doi.org/10.1016/0001-6918\(73\)90003-6](https://doi.org/10.1016/0001-6918(73)90003-6)
- Follmer, D. J. (2018). Executive function and reading comprehension: A meta-Analytic review. *Educational Psychologist*, 53, 42-60. <https://doi.org/10.1080/00461520.2017.1309295>
- Fondo de Naciones Unidas para la Infancia (2004). *¿Quién dijo que no se puede? Escuelas efectivas en sectores de pobreza*. Santiago, CL: Fondo de Naciones Unidas para la Infancia, UNICEF.
- Fonseca, L., Pujals, M., Lasala, E., Lagomarsino, I., Migliardo, G., Aldrey, A., Buonsanti, L., Barreyro, J. P. (2014) Desarrollo de habilidades de comprensión lectora en niños de escuelas de distintos sectores socioeconómicos. *Revista Neuropsicología Latinoamericana*, 6(1), 41-50. <https://doi.org/10.5579/rnl.2014.0151>
- Flotts, M. P., Manzi, J., Jiménez, D., Abarzúa, A., Cayuman, C. y García, M. J. (2015). *Informe de resultados TERCE. Tercer estudio regional comparativo y explicativo*. Santiago, CL: UNESCO.

- Fuentes, L. I. (2009). Diagnóstico de comprensión lectora en educación básica en Villarrica y Loncoche. *Perfiles Educativos*, XXXI (125), 23-37.
- Fuentes, C. (2014) Los límites del enunciado. *Estudios de Lingüística del Español*, 35(1), 137-160
- Furr, R. M. y Bacharach, V. R. (2014). *Psychometry. An introduction*. Thousand Oaks, CA: Sage.
- García, J., Cordero, M. Luque, J. y Santamaría, C. (1996) Intervención sobre la comprensión y recuerdos de textos: un programa de intervención experimental. *Infancia y Aprendizaje*, 74, 67-82
- García, P. E., Olea, J. y de la Torre, J. (2014). Application of cognitive diagnosis models to competency-based situational judgment tests. *Psicothema*, 26(3), 372-377. <https://doi.org/10.7334/psicothema2013.322>
- Gaviria, J. L., Biencinto, M. C. y Navarro, E. (2009). Invarianza de la estructura de covarianzas de las medidas de rendimiento académico en estudios longitudinales en la transición de educación primaria a secundaria. *Revista de Educación*, 348, 153-173.
- George, A. C. y Robitzsch, A. (2015). Cognitive diagnosis models in R: A didactic. *The Quantitative Methods for Psychology*, 11(3), 189-205. <https://doi.org/10.20982/tqmp.11.3.p189>
- George, A. C., Robitzsch, A., Kiefer, T., Groß, J. y Ünlü, A. (2016). The R package CDM for cognitive diagnosis models. *Journal of Statistical Software*, 74(2), 1-24. <https://doi.org/10.18637/jss.v074.i02>
- Gernsbacher, M. y Kaschak, M. (2013). Text comprehension. En D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology* (pp. 462-474). Londres: Oxford University Press.
- Gierl, M. J., Leighton, J. P. y Hunka, S. M. (2000). An NCME instructional module on exploring the logic of Tatsuoaka's rule-space model for test development and analysis. *Educational Measurement: Issues and Practice*, 19(3), 34-44. <https://doi.org/10.1111/j.1745-3992.2000.tb00036.x>

- Gil, J. (2009). Hábitos y actitudes de las familias hacia la lectura y competencias básicas del alumnado. *Revista de Educación*, 350, 301-322.
- Gómez, L. A. (1997a) Dimensión social de la comprensión textual (propuesta para un mejoramiento cualitativo del aprendizaje escolar). En M. Peronard, L. A. Gómez, G. Parodi y P. Núñez (Eds.). *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 43-54) Santiago, CL: Andrés Bello.
- Gómez, L. A. (1997b) Hacia una comprensión textual plena en el contexto de los mundos del hombre. En M. Peronard, L. A. Gómez, G. Parodi, P. Núñez (Eds.). *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 105-120) Santiago, CL: Andrés Bello.
- González, C. M. (2016) Aproximación al concepto de inferencia desde dos modelos de comprensión: Modelo estratégico y modelo de construcción e integración. *Literatura y Lingüística*, 35, 297-314. <http://dx.doi.org/10.4067/S0716-58112017000100295>
- Goldrine, T., Estrella, S., Olfos, R., Cáceres, P., Galdames, X. Hernández, N. y Medina, V. (2015). Conocimiento para la enseñanza del número en futuras educadoras de párvulos: Efecto de un curso de didáctica de la matemática. *Estudios Pedagógicos*, XLI(1), 93-105. <http://dx.doi.org/10.4067/S0718-07052015000100006>
- Gómez-Veiga, I., Vila, J. O., García-Madruga, J. A., Contreras, A. y Elosúa, M. R. (2013). Comprensión lectora y procesos ejecutivos de la memoria operativa. *Psicología Educativa*, 19(2), 103-111. <http://dx.doi.org/10.5093/ed2013a17>
- Gorin, J. S. (2007) Test construction and diagnostic testing. En J. P. Leighton y M. J. Gierl (Eds.) *Cognitive diagnostic assessment for education* (pp. 173-204) New York, NY: Cambridge University Press.
- Guarneros, E. y Vega, L. (2014). Habilidades lingüísticas orales y escritas para la lectura y escritura en niños preescolares. *Avances en Psicología Latinoamericana*, 32(1), 21-35. <https://dx.doi.org/10.12804/apl32.1.2014.02>
- Guerra, J. y Guevara, C. Y. (2017). Variables académicas, comprensión lectora, estrategias y motivación en estudiantes universitarios. *Revista Electrónica de*

*Investigación Educativa*, 19(2), 78-90.  
<https://doi.org/10.24320/redie.2017.19.2.1125>

Guevara, Y., García, G., López, A., Delgado, U. y Hermosillo, A. (2007) Habilidades lingüísticas en niños de estrato sociocultural bajo, al iniciar la primaria. *Acta Colombiana de Psicología*, 10(2), 9-17.

Gutiérrez, F., García, J. A., Elosúa, R., Luque, J. L. y Gárate, M. (2002). Memoria operativa y comprensión lectora: algunas cuestiones básicas. *Revista Acción Psicológica*, (1), 45-68. <https://doi.org/10.5944/ap.1.1.541>

Guzmán, B., Véliz, M. y Reyes, F. (2017). Memoria operativa, comprensión lectora y rendimiento escolar. *Literatura y lingüística*, (35), 379-404.  
<http://dx.doi.org/10.4067/S0716-58112017000100377>.

Hambleton, R. K., Swaminathan, H., y Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hayes, B., Tesar, B., y Zuraw, K. (2003). *OTSoft: Optimality theory software (Version 2.1)* [Software]. Recuperado de <http://www.linguistics.ucla.edu/people/hayes/otsoft/>

Hemmati, S. J., Baghaei, P. y Bemani, M. (2016). Cognitive diagnostic modeling of L2 reading comprehension ability: Providing feedback on the reading performance of iranian candidates for the university entrance examination. *International Journal of Language Testing*, 6(2), 92-100.

Henson, R. y Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement*, 29(4), 262-277.  
<https://doi.org/10.1177/0146621604272623>

Henson, R., DiBello, L., y Stout, B. (2018). A generalized approach to defining item discrimination for DCMs. *Measurement: Interdisciplinary Research and Perspectives*, 16(1), 18-29. <https://doi.org/10.1080/15366367.2018.1436855>

Herrada-Valverde, G. y Herrada, R. I. (2017) Análisis del proceso de comprensión lectora de los estudiantes desde el modelo construcción-integración. *Perfiles Educativos*, XXXIX(157), 181-197.

- Hirsch, E. D. (2003). *Reading comprehension requires knowledge of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores*. American Federation of Teachers: American Educator
- Hou, L., de la Torre, J. y Nandakumar, R. (2014). Differential item functioning assessment in cognitive diagnostic modeling: Application of the Wald test to investigate DIF in the DINA model. *Journal of Educational Measurement*, 51(1), 98-115. <https://doi.org/10.1111/jedm.12036>
- Hu, J., Miller, D., Huggins-Manley, A. C. y Chen, Y. (2016). Evaluation of model fit in cognitive diagnosis models. *International Journal of Testing*, 16(2), 119-141. <https://doi.org/10.1080/15305058.2015.1133627>
- Huebner, A. (2010). An overview of recent developments in cognitive diagnostic computer adaptive assessments. *Practical Assessment, Research y Evaluation*, 15(3), 1-7.
- Ibáñez, R. (2012). La comprensión del discurso escrito: Una propuesta teórico-metodológica para su evaluación. *Signos*, 45 (78), 20-43. <http://dx.doi.org/10.4067/S0718-09342012000100002>
- Jadue, G. (1997). Factores ambientales que afectan el rendimiento escolar de los niños provenientes de familias de bajo nivel socioeconómico y cultural. *Estudios Pedagógicos*, 1997(23), 75-80. <http://dx.doi.org/10.4067/S0718-07051997000100007>
- Jang, E. E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for fusion model application to LanguEdge assessment. *Language Testing*, 26(1), 31-73. <https://doi.org/10.1177/0265532208097336>
- Jang, E. E., Dunlop, M., Wagner, M., Kim, Y. y Gu, Z. (2013). Elementary schools ELL's reading skill profiles using cognitive diagnosis modeling: Roles of length of residence and home language environment. *Language Learning*, 63(3), 400-436. doi: <http://dx.doi.org/10.1111/lang.12016>
- Jang, E. E., Dunlop, M., Wagner, M., Park, G. y van der Boom, E. H. (2015). How do young students with different profiles of reading skill mastery, perceived ability,

and goal orientation respond to holistic diagnostic feedback? *Language Testing*, 32(3), 359-383. <http://dx.doi.org/10.1177/0265532215570924>

Juel, C. (1988) Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80(4), 437-447. <https://doi.org/10.1037/0022-0663.80.4.437>

Kane, M. T. y Bejar, I. I. (2014). Cognitive frameworks for assessment, teaching, and learning: A validity perspective. *Psicología Educativa*, 20(2), 117-123. <https://doi.org/10.1016/j.pse.2014.11.006>

Karreman, J. (2004) *Use and effect of declarative information in user instructions*. New York: Rodopi.

Kendeou, P., van den Broek, P., Helder, A. y Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research y Practice*, 29(1), 10-16. <https://doi.org/10.1111/ldrp.12025>

Kim, Y. (2011) Diagnosing EAP writing ability using the reduced reparameterized unified model. *Language Testing*, 28(4), 509-541. <http://dx.doi.org/10.1177/0265532211400860>

Kim, A. (2015). Exploring ways to provide diagnostic feedback with an ESL placement test: Cognitive diagnostic assessment of L2 reading ability. *Language Testing*, 32(2), 227-258. <http://dx.doi.org/10.1177/0265532214558457>

King, G., Honaker, J., Joseph, A. y Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm or multiple imputation. *American Political Science Review*, 95(1), 49-69.

Kintsch, W. y Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394. <http://dx.doi.org/10.1037/0033-295X.85.5.363>

Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. New York, NY: Cambridge University Press.

- Klein, M., Birenbaum, M., Standiford, S., y Tatsuoka, K. K. (1981). *Logical error analysis and construction of tests to diagnose student 'bugs' in addition and subtraction of fractions* (Technical Report 81-6-NIE). Urbana, IL: University of Illinois, CERL.
- Köhn, H., Chiu, C., Brusco, M. J. (2015). Heuristic cognitive diagnosis when the Q-matrix is unknown. *British Journal of Mathematical and Statistical Psychology*, 68, 268-291. <https://doi.org/10.1111/bmsp.12044>
- Köhn, H. y Chiu, C. (2017) A procedure for assessing the completeness of the Q-matrices of cognitively diagnostic tests. *Psychometrika*, 82(1), 112-132. <https://doi.org/10.1007/s11336-016-9536-7>
- Krashen, S. y Brown, C.L (2005). The ameliorating effects of high socioeconomic status: A secondary analysis. *Bilingual Research Journal*, 29(1), 185-196. <https://doi.org/10.1080/15235882.2005.10162830>
- Kuo, B., Chen, C., Yang, C. y Mok, M. M. (2016). Cognitive diagnostic models for tests with multiple-choice and constructed-response items. *Educational Psychology*, 36, 1115-1133. <http://dx.doi.org/10.1080/01443410.2016.1166176>
- Lazarsfeld, P. F. y Henry, N. W. (1977). *Análisis de la estructura latente*. Madrid: Instituto de Estudios Políticos.
- Lee, Y. y Sawaki, Y. (2009a). Cognitive diagnosis approaches to language assessment: An overview. *Language Assessment Quarterly*, 6(3), 172-189. <http://dx.doi.org/10.1080/15434300902985108>
- Lee, Y. y Sawaki, Y. (2009b). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, 6(3), 239-263. <http://dx.doi.org/10.1080/15434300903079562>
- Lei, P. y Li, H. (2016). Performance of fit indices in choosing correct cognitive diagnostic models and Q-matrices. *Applied Psychological Measurement*, 40(6), 405-417. <http://dx.doi.org/10.1177/0146621616647954>
- Leighton, J. P., Gierl, M. J. y Hunka, S. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule-space approach. *Journal of*

*Educational Measurement*, 41, 205-237. <https://doi.org/10.1111/j.1745-3984.2004.tb01163.x>

Leighton, J. P. y Gierl, M. J. (2007a). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice*, 26(2), 3-16. <https://doi.org/10.1111/j.1745-3992.2007.00090.x>

Leighton, J. P. y Gierl, M. J. (2007b) Why cognitive diagnostic assessment? En J. P. Leighton y M. J. Gierl (Eds.) *Cognitive diagnostic assessment for education* (pp. 3-18) New York, NY: Cambridge University Press.

Leighton, J. P., Gokiert, R. J., Cor, M. K. y Heffernan, C. (2010) Teacher beliefs about the cognitive diagnostic information of classroom- versus large-scale tests: Implications for assessment literacy. *Assessment in Education: Principles, Policy y Practice*, 17(1), 7-21. <http://dx.doi.org/10.1080/09695940903565362>

Leseman, P. P. y de Jong, P. F. (1998). Home literacy: Opportunity, instruction, cooperation and social-emotional quality predicting early reading achievement. *Reading Research Quarterly*, 33(3), 294-318. <http://dx.doi.org/10.1598/RRQ.33.3.3>

Liao, C., Kuo, B. y Deenang, E. (2015). *Exploring the mastery pattern of a grade 4 reading comprehension test using cognitive diagnosis model*. Bangkok: Conference Proceedings of International Conference on Advances in Business y Management (ICABM).

Li, H., Hunter, V. y Lei, P. (2015). The selection of cognitive diagnostic models for a reading comprehension test. *Language Testing*, 33(3), 391-409. <https://doi.org/10.1177/0265532215590848>

Li, X. y Wang, W. (2015) Assessment of differential item functioning under cognitive diagnosis models: The DINA model example. *Journal of Educational Measurement*, 52(1), 28-54 <https://doi.org/10.1111/jedm.12061>

Liu, J. (2015). On the consistency of Q-matrix estimation: A comentary. *Psychometrika*, 82(2), 523-527. <https://doi.org/10.1007/s11336-015-9487-4>

- López-Higes, R. y Rubio-Valdehita, S. (2014) ¿Qué variables determinan el nivel lector de un alumno en el segundo ciclo de Educación Primaria y cuál es su valor diagnóstico? *Revista Iberoamericana de Diagnóstico y Evaluación*, 1(37), 31-52
- Lorenzo, J. R. (2001a). Procesos cognitivos básicos relacionados con la lectura. Primera parte: la conciencia fonológica. *Interdisciplinaria*, 18(1), 1-33
- Lorenzo, J. R. (2001b). Procesos cognitivos básicos relacionados con la lectura. Segunda Parte: La memoria y su implicancia en el rendimiento lector. *Interdisciplinaria*, 18(2), 115-134.
- Lorenzo, F. (2016). Competencia en comunicación lingüística: Claves para el avance de la comprensión lectora en las pruebas PISA. *Revista de Educación*, 374, 142-160.
- Luna, D. A. (2017). Modelos de diagnóstico cognitivo: Fundamentos, didáctica y aplicaciones en ciencias del comportamiento. *Salud y Sociedad*, 8(1), 68-80.  
<https://doi.org/10.22199/S07187475.2017.0001.00005>
- Luo, R., Pace, A., Masek, L. R., Hirsh-Pasek, K. y Golinkoff, R. M. (2016) The family's role in the relation between socioeconomic status and early language development. *Journal of Family Medicine*, 3(6), 1073-1078.
- Ma, W. y de la Torre, J. (2016). A sequential cognitive diagnosis model for polytomous responses. *British Journal of Mathematical and Statistical Psychology*, 69(3), 253-275. <https://doi.org/10.1111/bmsp.12070>
- Ma, X. y Meng, Y. (2014). Towards personalized english learning diagnosis: Cognitive diagnostic modelling for EFL listening. *Asian Journal of Education and e-Learning*, 2(5), 336-348.
- Marchant, T., Lucchini, G. y Cuadrado, B. (2007). ¿Por qué leer bien es importante? Asociación del dominio lector con otros aprendizajes. *Psyche*, 16(2), 3-16.  
<http://dx.doi.org/10.4067/S0718-22282007000200001>
- Martínez-Arias, M. R., Hernández, M. J. y Hernández, M. V. (2006). *Psicometría*. Madrid: Alianza Editorial.

- Matus, T. y Quezada, M. (2007) Los desafíos de trabajar la pobreza con calidad. Análisis de la relación complejidad versus condiciones de calidad en el ámbito social municipal de las 52 comunas de la Región Metropolitana. En T. Matus (Coord.) *La reforma Municipal en la Mira* (pp. 85-132). Santiago, CL: Expansiva.
- Maydeu-Olivares, A. y Joe, H. (2014). Assessing approximate fit in categorical data analysis. *Multivariate Behavioral Research*, 49, 305-328.  
<https://doi.org/10.1080/00273171.2014.911075>
- Mayol, A. (2011). *Investigar la desigualdad: Problemas de iluminación*. Santiago, CL: Centro de Investigación en Estructura Social.
- McEwan, P. J. y Carnoy, M. (2000). The effectiveness and efficiency of private schools in Chile's voucher system. *Educational Evaluation and Policy Analysis*, 22(3), 213-239. <https://doi.org/10.3102/01623737022003213>
- McIntyre, H. (2011). Sociocultural perspectives on children with reading difficulties. En A. McGill-Frazen y R. L. Allington (Eds.) *Handbook of reading disability research* (pp. 45-56) New York, NY: Routledge.
- Medina, A. y Guajardo, A. M. (2014) *Pruebas de comprensión lectora y producción de textos CL-PT. Kinder a 4° Básico*. Santiago, CL: Ediciones Universidad Católica
- Mella, O. (2006). Factores que afectan los resultados de la escuela pública chilena. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 4(1), 29-37.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58(4), 525–543. <https://doi.org/10.1007/BF02294825>
- Merino, J. M. y Maldonado, J. A. (2014). Estudio de efectos contextuales en el rendimiento en matemáticas de alumnos de 8° Básico de la región del Biobío, Chile. *Estudios Pedagógicos*, 15(2), 241-263. <http://dx.doi.org/10.4067/S0718-07052014000300015>
- Messick, S. (1989). Validity. En R. L. Linn (Ed.) *Educational Measurement* (pp. 1–103). New York, NY: American Council on Education/Macmillan.

- Ministerio de Educación (2007). *PISA 2006: Rendimientos de estudiantes de 15 años en ciencias, lectura y matemática*. Santiago, CL: Ministerio de Educación.
- Ministerio de Educación (2012). *Lenguaje y Comunicación. Bases curriculares*. Santiago, CL: Ministerio de Educación.
- Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge.
- Mislevy, R. J, Wilson, M. R., Ercikan, K. y Chudowsky, N. (2003) Psychometric principles in student assessment. En T. Kellaghan, D. L. Stufflebeam y L. A. Wingate (Eds.) *International handbook of educational evaluation. Part One: Perspectives* (pp. 489-532) Dordrecht: Kluwer Academic.
- Mizala, A. y Romaguera. P. (1998). *Desempeño escolar y elección de colegios: La experiencia chilena. Serie Economía N° 36*. Santiago, CL: Centro de Economía Aplicada Universidad de Chile.
- Mizala, A., Romaguera, P. y Ostoic, C. (2004). *A hierarchical model for studying equity and achievement in the Chilean school choice system*. Santiago, CL: Centro de Economía Aplicada Universidad de Chile.
- Mizala, A., Romaguera, P. y Urquiola, M. (2007). Socioeconomic status or noise? Tradeoffs in the generation of school quality information. *Journal of Development Economics*, 84, 61-75. <https://doi.org/10.1016/j.jdeveco.2006.09.003>
- Mizala, A. y Torche, F. (2012) Bringing the schools back in: The stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, 32, 132–144. <https://doi.org/10.1016/j.ijedudev.2010.09.004>
- Moll, L. C., Amanti, C., Neff, D. y González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect home and classrooms. *Qualitative Issues in Educational Research*, 31(2), 132-141. <https://doi.org/10.1080/00405849209543534>

- Montero, I. y León, O. (2002). Clasificación y descripción de las metodologías de investigación en psicología. *International Journal of Clinical and Health Psychology*, 2(3), 503-508.
- Moreno, R., Martínez, R. J. y Muñiz, J. (2006). New guidelines for developing multiple-choice items. *Methodology*, 2(2), 65-72. <https://doi.org/10.1027/1614-2241.2.2.65>
- Morgan, H. (2016). Relying on high-stakes standardized tests to evaluate schools and teachers: A bad idea. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 89(2), 67-72. <https://doi.org/10.1080/00098655.2016.1156628>
- Mullis, I. V., Martin, M. O., Foy, P. y Hooper, M. (2016). TIMSS 2015 international results in mathematics. Recuperado de TIMSS y PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/>
- Mullis, I. V. S., Martin, M. O., Foy, P., y Hooper, M. (2017). PIRLS 2016 International Results in Reading. Recuperado de Boston College, TIMSS y PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/international-results/>
- Muñoz-Chereau, B. y Thomas, S. M. (2015): Educational effectiveness in Chilean secondary education: Comparing different ‘value added’ approaches to evaluate schools. *Assessment in Education: Principles, Policy y Practice* 23(1), 26-52. <http://dx.doi.org/10.1080/0969594X.2015.1066307>
- Nichols, P. (1994). A framework for developing cognitively diagnostic assessments. *Review of Educational Research*, 64, 575-603. <https://doi.org/10.3102/00346543064004575>
- Noble, K. G., Farah, M. J. y McCandliss, B. D. (2006). Socioeconomic background modulates cognition–achievement relationships in reading. *Cognitive Development*, 21(3), 349-368. <https://doi.org/10.1016/j.cogdev.2006.01.007>
- Núñez, P. y Donoso, P. (2000). Evaluación de la propuesta de enseñanza de la comprensión lectora en los textos escolares de la Reforma Educacional Chilena: Un estudio exploratorio. *Revista Signos*, 33(47), 123-150. <http://dx.doi.org/10.4067/S0718-09342000000100011>

- OECD (2009). *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*. París, FR: OECD Publishing.  
<http://dx.doi.org/10.1787/9789264062658-en>
- OECD (2012). *¿Tiene el cheque escolar relación con la equidad en educación?* Paris, FR: OECD Publishing.
- OECD (2014). *Resultados de PISA 2012 en foco. Lo que los alumnos saben a los 15 años y lo que pueden hacer con lo que saben*. Paris, FR: OECD Publishing.
- OECD (2016). *Low-performing students: Why they fall behind and how to help them succeed, PISA*. París, FR: OECD Publishing. <http://dx.doi.org/10.1787/19963777>
- Oliva, M. A. (2008). Política educativa y profundización de la desigualdad en Chile. *Estudios Pedagógicos*, XXXIV(2), 207-226. <http://dx.doi.org/10.4067/S0718-07052008000200013>
- Orellana, R. y Merino, J. M. (2013) Predictores asociados a variaciones en puntajes Simce en la región del Biobío. *CUHSO*, 23(1), 37-54.  
<http://dx.doi.org/10.7770/CUHSO-V23N1-ART445>
- Ortúzar, M .S., Flores, C., Milesi, C. y Cox, C. (2009). Aspectos de la formación inicial docente y su influencia en el rendimiento académico de los alumnos. En Ministerio de Educación (Ed.). *Camino al Bicentenario. Propuestas para Chile* (pp. 155-186) Santiago: Ministerio de Educación.
- Osses, A., Bellei, C. y Valenzuela, J. P. (2015). External technical support for school improvement: Critical issues from the Chilean experience. *Journal of Educational Administration and History*, 47(3), 272-293.  
<https://doi.org/10.1080/00220620.2015.1038699>
- Paris, S. G. y Stahl, S. A. (2005). *Children's Reading Comprehension and Assessment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Park, C. y Cho, S. (2011). Cognitive diagnostic assessment of english grammar for korean EFL learners. *English Teaching*, 66(4), 101-117.  
<https://doi.org/10.15858/engtea.66.4.201112.101>

- Park, C. y Cho, S. (2017). An exploratory analysis of compensatory cognitive diagnosis in EFL reading comprehension. *Korea Journal of English Language and Linguistics*, 17(1), 85-104.
- Parodi, G. (2011). La teoría de la comunicabilidad: Notas para una concepción integral de la comprensión de textos escritos. *Signos*, 44(76), 145-167. <http://dx.doi.org/10.4067/S0718-09342011000200004>
- Paz, M. T. (2015). Educational policies in Chile: Municipality law and LOCE, how they have impacted Chilean education. *Imago*, 8, 53-75.
- Pellegrino, J. W., Chudowsky, N. y Glaser, R. (2001). *Knowing what students know. The science and design of educational assessment*. Washington, DC: National Academic Press.
- Pérez, M. J. (2005). La evaluación de la comprensión lectora: Dificultades y limitaciones. *Revista de Educación, Número especial*, 121-138.
- Pérez, P. (2009). La comprensión lectora y la competencia en comunicación lingüística en el nuevo marco curricular: Algunas claves para su desarrollo. *Educatio Siglo XXI*, 27(1), 13-32.
- Pérez, C., Betancort, M. y Cabrera, L. (2011). Family influences in academic achievement. A study of the Canary Islands. *Revista Internacional de Sociología*, 71(1), 169-187. <http://dx.doi.org/10.3989/ris.2011.04.11>
- Piacente, T. (2009) Especificidad de la evaluación psicológica en los procesos de aprendizaje del lenguaje escrito. Evolución y estado actual. *Revista Iberoamericana de Diagnóstico y Evaluación*, 2(28), 135-148.
- Piccolo, L. D., Arteché, A. X., Fonseca, R. P. y Grassi-Oliveira, R. (2016) Influence of family socioeconomic status on IQ, language, memory and executive functions of brazilian children. *Psicologia: Reflexao e Crítica*, 29(23), 1-10, <http://dx.doi.org/10.1186/s41155-016-0016-x>
- Pigott, T. D. (2001). A review of methods for missing data. *Educational Research and Evaluation*, 7(4), 353-383. <https://doi.org/10.1076/edre.7.4.353.8937>

- Puga, B. (2011). Escuela y estratificación social en Chile: ¿Cuál es el rol de la municipalización y la educación particular subvencionada en la reproducción de la desigualdad social? *Estudios Pedagógicos*, 37(2), 213-232.  
<http://dx.doi.org/10.4067/S0718-07052011000200013>
- Qin, C., Zhang, L., Qiu, D., Huang, L., Teng, T., Jiang, H., Ren, Q. y Zhou, J. (2015). Model identification and Q-matrix incremental inference in cognitive diagnosis. *Knowledge-Based Systems*, 86, 66-76.  
<http://dx.doi.org/10.1016/j.knosys.2015.05.024>
- Ramírez, M., Devia, R. E., Salazar, L. y Ramón, A. (2011). Pobreza y rendimiento escolar: Estudio de caso de jóvenes de alto rendimiento. *Educere*, 15(52), 663-672.
- Ramos, C. (2006). Elaboración de un instrumento para medir comprensión lectora en niños de octavo básico. *Onomázein*, 2(14), 197-210.
- Ravand, H., Barati, H. y Widhiarso, W. (2013). Exploring diagnostic capacity of a high stakes reading comprehension test: A pedagogical demonstration. *Iranian Journal of Language Testing*, 3(1), 11-37.
- Ravand, H. y Robitzsch, A. (2015). Cognitive diagnostic modeling using R. *Practical Assessment Research y Evaluation*, 20(11), 1-12.
- Ravand, H. (2016). Application of a cognitive diagnostic model to a high-stakes reading comprehension test. *Journal of Psychoeducational Assessment*, 34(8), 765-781.  
<http://dx.doi.org/10.1177/0734282915623053>
- Rayner, K. y Reichle, E. D. (2010) Models of the reading process. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 787-799. <http://dx.doi.org/10.1002/wcs.68>
- R Core Team (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Recuperado de <https://www.r-project.org/>.
- Ripoll, J. C. y Aguado, G. (2014). La mejora de la comprensión lectora en español: Un meta-análisis. *Revista de Psicodidáctica*, 19(1), 27-44.

- Riffo, B. (2016). Representaciones mentales en la comprensión del discurso: Del significativo lineal al modelo de situación. *Revista Signos*, 49(1), 205-223. <http://dx.doi.org/10.4067/S0718-09342016000400010>
- Rivkin, S. G., Hanushek, E. A. y Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>
- Robertson, S. L. y Verger, A. (2012) Governing education through public private partnerships. En S. Robertson, K. Mundy, A. Verger y F. Menashy (Eds.) *Public Private Partnerships in Education: New Actors and Modes of Governance in a Globalizing World* (pp. 21-42) Cheltenham, UK: Edward Elgar Publishing.
- Rojas, G., de la Torre, J., y Olea, J. (2012, Abril). *Choosing between general and specific cognitive diagnosis models when the sample size is small*. Paper presentado en el Annual Meeting of the National Council on Measurement in Education, Vancouver, British Columbia, Canada.
- Romero, S. J. (2010). *Propiedades y aplicaciones del método de las distancias mínimo-cuadráticas (LSDM) para la validación y análisis de atributos cognitivos*. Universidad autónoma de Madrid, Facultad de Psicología. Recuperado de <https://repositorio.uam.es/handle/10486/4614?show=full>
- Rosas, R., Medina, L., Meneses, A., Guajardo, A., Cuchacovich, S. y Escobar, A. (2011). Construcción y validación de una prueba de evaluación de competencia lectora inicial basada en computador. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*, 48(1), 43-61. <http://dx.doi.org/10.7764/PEL.48.1.2011.4>
- Ruíz de Miguel, C. (2001). Factores familiares vinculados al bajo rendimiento. *Revista Complutense de Educación*, 12(1), 81-113.
- Rupp, A. A., Templin, J. y Henson, R. A. (2010). *Diagnostic measurement. Theory, methods, and applications*. New York, NY: Guilford Press.
- Rupp, A. A. y Templin, J. L. (2008) Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement:*

*Interdisciplinary Research and Perspectives*, 6(4), 219-262.

<http://dx.doi.org/10.1080/15366360802490866>

Sánchez, E. (1996). Los textos divulgativos como una conversación encubierta: Análisis de los recursos comunicativos de un texto divulgativo. *Infancia y Aprendizaje*, 75, 85-96. <https://doi.org/10.1174/021037096762994276>

Santisteban, C. y Alvarado, J. (2001). *Modelos psicométricos*. Madrid: Universidad Nacional de Educación a Distancia.

Sapelli, C. y Vial, B. (2002). The performance of private and public school in the Chilean voucher system. *Cuadernos de Economía*, 39(118), 423-454. <http://dx.doi.org/10.4067/S0717-68212002011800006>

Sawaki, Y., Kim, H. y Gentile, C. (2009). Q-Matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessments. *Language Assessment Quarterly*, 6(3), 190-209. <http://dx.doi.org/10.1080/15434300902801917>

Schiefelbein, E. (1992). Análisis del SIMCE y sugerencias para mejorar su impacto en la calidad". En S. Gómez (Ed.), *La realidad en cifras. Estadísticas Sociales* (pp. 241-280). Santiago, CL: Flacso/INE.

Silva-Peña, I., Chelme-Bustos, A. y Salgado-Labra, I. (2003). Educación y conocimiento: ¿Ejes de la transformación productiva con equidad? El caso de Chile. *Relieve*, 9(1), 73-81. <https://doi.org/10.7203/relieve.9.1.4350>

Snow, R. E., y Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. En R. L. Linn (Ed.) *Educational Measurement*. (pp. 263–331). New York, NY: American Council on Education/Macmillan.

Somers, M., McEwan, P. J. y Willms, J. D. (2004). How effective are private schools in Latin America? *Comparative Education Review*, 48(1), 48-69. <https://doi.org/10.1086/379841>

Sotelo, L., Sotelo, N., Sotelo, C., Matalinares, M. L., Arenas, C. y Caycho, T. (2012) Efectos de un programa experimental interactivo para mejorar la comprensión

- lectora en una población rural de 5° grado de educación primaria. *Revista IIPSI*, 15(1), 93-107. <http://dx.doi.org/10.15381/rinvp.v15i1.3671>
- Svetina, D., Gorin, J. S. y Tatsuoka, K. K. (2011). Defining and comparing the reading comprehension construct: A cognitive-psychometric modeling approach. *International Journal of Testing*, 11, 1-23. <https://doi.org/10.1080/15305058.2010.518261>
- Sugate, S. P., Schaughency, E. A. y Reese, E. (2013) Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly*, 28, 33-48. <http://dx.doi.org/10.1016/j.ecresq.2012.04.004>
- Sun, J., Xin, T., Zhang, S. y de la Torre, J. (2013). Polytomous extension of the generalized distance discriminating method. *Applied Psychological Measurement*, 37(7), 503-521. <http://dx.doi.org/10.1177/0146621613487254>
- Tatsuoka, K. K., y Tatsuoka, M. M. (1983). Spotting erroneous rules of operation by the individual consistency index. *Journal of Educational Measurement*, 20(3), 221-230. <https://doi.org/10.1111/j.1745-3984.1983.tb00201.x>
- Tatsuoka, K. K. (1985). A probabilistic model for diagnosing misconceptions by the pattern classification approach. *Journal of Educational Statistics*, 10(1), 55-73. <https://doi.org/10.3102/10769986010001055>
- Tatsuoka, K. K., Corter, J. E. y Tatsuoka, C. (2004) Patterns of diagnosed mathematical content and process skills in TIMSS-R across a sample of 20 countries. *American Educational Research Journal*, 41(4), 901-926. Recuperado de <http://www.jstor.org/stable/3699467>
- Tatsuoka, K. K. (2009). *Cognitive assessment. An Introduction to the rule space method*. New York, NY: Routledge.
- Templin, J. L., Henson, R. A., Templin, S. E. y Roussos, L. (2008). Robustness of hierarchical modeling of skill association in cognitive diagnosis models. *Applied Psychological Measurement*, 32(7), 559-574. <http://dx.doi.org/10.1177/0146621607300286>

- Thieme, C., Tortosa-Ausina, E., Prior, D., y Gempp, R. (2012). *Value-added and contextual factors in education: Evidence from Chilean schools*. Recuperado de [http://www.grupobbva.com/TLFU/dat/DT\\_05\\_2013.pdf](http://www.grupobbva.com/TLFU/dat/DT_05_2013.pdf)
- Thompson, N. A. y Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research y Evaluation*, 16(1), 1-9.
- Tijero, T. (2009) Representaciones mentales: Discusión crítica del modelo de situación de Kintsch. *Onomázein*, 19(1), 111-138
- Torche, F. (2005). Privatization reform and inequality of educational opportunity: The case of Chile. *Sociology in Education*, 78, 316-343. <https://doi.org/10.1177/003804070507800403>
- Tracey, D. H. y Morrow, L. M. (2017) *Lenses on reading. An Introduction to theories and models*. New York: Guilford Press
- Treviño, E., Salazar, F. y Donoso, F. (2011). ¿Segregar o incluir? esa no debería ser una pregunta en educación. *Docencia*, 45, 36-47.
- Ugarriza, N. (2006). La comprensión lectora inferencial de los textos especializados y el rendimiento académico de los estudiantes universitarios de primer ciclo. *Persona*, 9, 31-75.
- Usó-Juan, E. (2006). The compensatory nature of discipline-related knowledge and english-language proficiency in reading english for academic purposes. *The Modern Language Journal*, 90(2), 210-227. <https://doi.org/10.1111/j.1540-4781.2006.00393.x>
- Valenzuela, J. P., Bellei, C. y de los Ríos, D. (2008). *Evolución de la segregación socioeconómica de los estudiantes chilenos y su relación con el financiamiento compartido*. Santiago, CL: Fondo de Investigación y Desarrollo en Educación - FONIDE.
- Valenzuela, J. P., Bellei, C. y de los Ríos, D. (2013). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Education Policy*, 29(2), 217-241. <https://doi.org/10.1080/02680939.2013.806995>

- Van Dijk, T. A. y Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.
- Vergara, D., Strasser, K. y del Río, M. F. (2016). Más que palabras por minuto: Las otras habilidades que afectan la comprensión en 1° Básico. *Calidad en la Educación*, 44, 46-67.
- Vidal-Abarca, E. y Gilabert, R. (1990). Comprender para aprender: Un programa para mejorar la comprensión y el aprendizaje de textos Informativos. *Comunicación, Lenguaje y Educación*, 6, 113-124.
- Villalobos, C. y Valenzuela, J. P. (2012). Polarización y cohesión social del sistema escolar chileno. *Revista de Análisis Económico*, 27(2), 145-172. <http://dx.doi.org/10.4067/S0718-88702012000200005>
- Völkel, G., Seabi, J., Cockcroft, K. y Goldschagg, P. (2016) The impact of gender, socioeconomic status and home language on primary school children's reading comprehension in Kwazulu-Natal. *International Journal of Environmental Research and Public Health*, 13(3), 1-11. <http://dx.doi.org/10.3390/ijerph13030322>
- von Davier, M. (2005). *A general diagnostic model applied to language testing data*. Princeton, NJ: Educational Testing Service.
- von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, 61, 287-307. <https://doi.org/10.1348/000711007X193957>
- von Davier, M., DiBello, L. y Yamamoto, K. (2006) *Reporting test outcomes using models for cognitive diagnosis*. Princeton, NJ: Educational Testing Service
- Vygotsky, L. (1978). *Mind and society: The development of higher psychological process*. Cambridge, MA: Harvard University Press.
- Wang, C. y Gierl, M. J. (2011). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills in critical reading. *Journal of Educational Measurement*, 48(2), 165-187. <https://doi.org/10.1111/j.1745-3984.2011.00142.x>

- Wilkinson, I. (2002). Introduction: peer influences on learning: where are they?  
*International Journal of Educational research*, 37, 395-401.  
[https://doi.org/10.1016/S0883-0355\(03\)00012-0](https://doi.org/10.1016/S0883-0355(03)00012-0)
- Woolley, G. (2011). *Reading comprehension: Assisting children with learning difficulties*. New York, NY: Springer. [https://doi.org/10.1007/978-94-007-1174-7\\_2](https://doi.org/10.1007/978-94-007-1174-7_2)
- Wu, M., Tam, H. P. y Jen, T. (2016). *Educational measurement for applied researchers. Theory into practice*. Singapore: Springer.
- Yan, D., Mislevy, R. J. y Almond, R. G. (2003). *Design and analysis in a cognitive assessment*. Princeton, NJ: Educational Testing Service.
- Yang-Hansen, K. (2009). Ten-year trend in SES effects on reading achievement at school and individual levels: A cross-country comparison. *Educational Research and Evaluation*, 14(6), 521-537. <http://dx.doi.org/10.1080/13803610802576759>
- Yang, K., Rosén, M. y Gustafsson, J. (2011). Changes in the multi-level effects of socio-economic status on reading achievement in Sweden in 1991 and 2001. *Scandinavian Journal of Educational Research*, 55(2), 197-211.  
<http://dx.doi.org/10.1080/00313831.2011.554700>
- Yang, X. y Embretson, S. E. (2007) Construct validity and cognitive diagnosis assessment. En J. P. Leighton y M. J. Gierl (Eds.) *Cognitive diagnostic assessment for education* (pp. 119-145) New York, NY: Cambridge University Press.
- Yi, Y. (2017). Probing the relative importance of different attributes in L2 reading and listening comprehension items: An application of cognitive diagnostic models. *Language Testing*, 34(3), 337-355. <https://doi.org/10.1177/0265532216646141>
- Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa*, 20(2), 77-87.  
<http://dx.doi.org/10.1016/j.pse.2014.11.003>