



ALINEAMIENTO DE PRUEBAS SIMCE Y TIMSS DE CUARTO BASICO

Documento de Trabajo N°3

Diciembre 2013

Departamento de Estudios de la Calidad de la Educación

División de Estudios

AGENCIA DE CALIDAD DE LA EDUCACIÓN

Índice

Índice	2
1. Introducción	2
2. Marco Teórico y Metodologías Utilizadas	4
2.1. Análisis de los ítems en Marco de TCT	6
2.2. Teoría de Respuesta al Ítem	7
3. Análisis Descriptivos	9
4. Resultados	14
4.1. Análisis de Calibración Conjunta	17
4.2. Tabla de equivalencia	25
5. Conclusiones	30
Bibliografía	31
6. Anexos	33
6.1. Detalle de Análisis por cuadernillo	33
6.1.1. Simce - Forma C	33
6.1.2. Simce - Forma D	36
6.1.3. Simce - Forma E	39
6.1.4. Simce - Forma G	42
6.1.5. TIMSS - Forma 1	45
6.1.6. TIMSS - Forma 2	48
6.1.7. TIMSS - Forma 3	51
6.1.8. TIMSS - Forma 4	54
6.1.9. TIMSS - Forma 5	57
6.1.10. TIMSS - Forma 6	60
6.1.11. TIMSS - Forma 7	63
6.1.12. TIMSS - Forma 8	66
6.1.13. TIMSS - Forma 9	69
6.1.14. TIMSS - Forma 10	72

6.1.15. TIMSS - Forma 11	75
6.1.16. TIMSS - Forma 12	78
6.1.17. TIMSS - Forma 13	81
6.1.18. TIMSS - Forma 14	84
6.2. Tabla de Equivalencia	86

Esta es una publicación del departamento de Estudios de la Calidad de la Educación, División de Estudios, Agencia de la Calidad de la Educación.

Claudia Matus Correa¹

¹Ing. Civil Matemático. PhD Estadística University of Pittsburgh. Departamento de Estudios de la Calidad de la Educación.

1. Introducción

La participación de Chile en estudios internacionales de logro educativo tales como TIMSS y PISA, ha permitido contextualizar el rendimiento de los estudiantes chilenos, comparándolo con el obtenido por sus pares en otros países. La información recogida por dichas pruebas complementa la que proporciona la prueba nacional Simce. Si bien, tanto las pruebas internacionales como el Simce se construyen bajo similares modelos evaluativos y criterios de calidad, las poblaciones objetivo a las que son aplicadas son de distinta composición. Considerando que las buenas prácticas para el análisis y construcción de pruebas exigen que estas se adapten a la población objetivo (Hambleton, 1985), se plantean interrogantes acerca del grado de adecuación particular a Chile que tienen las pruebas internacionales. También, sobre la calidad técnica de la prueba nacional Simce en comparación con las pruebas internacionales. Adicionalmente, dado que las pruebas internacionales son aplicadas solo a una muestra representativa de estudiantes², surge la inquietud de determinar una equivalencia entre las escalas de puntaje de la prueba Simce y las pruebas internacionales. De esta manera, los distintos establecimientos podrían estimar cuál hubiera sido el rendimiento de sus alumnos en dichas pruebas internacionales a partir de su rendimiento promedio en la prueba Simce.

Este documento de trabajo tiene el propósito de contribuir a responder las interrogantes anteriores. Por un lado, se quiere determinar el grado de pertinencia psicométrica particular a Chile que presenta la prueba internacional TIMSS de Matemática y, simultáneamente, comparar la calidad métrica de las preguntas TIMSS y Simce. Asimismo, considerando que ambas evaluaciones son pruebas curriculares –Simce mide el currículo nacional vigente y TIMSS mide un currículo internacional consensuado– y que comparten un mismo modelo de medición en el marco de la Teoría de Respuesta al Ítem³, surge la posibilidad de establecer una equivalencia entre las escalas de medición de ambas pruebas mediante una tabla de correspondencia de puntajes.

El 19 de octubre del año 2010, los alumnos matriculados en cuarto básico de todo el país rindieron las pruebas Simce de Matemática y de Lenguaje y Comunicación: Comprensión de Lectura⁴. Ese mismo año, durante la segunda quincena de noviembre y la primera semana de diciembre, una muestra representativa de estudiantes a nivel nacional, adicionalmente participó en la aplicación de la prueba TIMSS⁵, contestando una prueba de Matemática y otra de Ciencias Naturales. Se dispone entonces de las respuestas de 5.148 estudiantes de cuarto básico tanto a las preguntas de la prueba Simce, como a las de la prueba TIMSS de Matemática. Dichas

²En grados o edades que también son medidos por el Simce.

³Es decir, son construidas bajo el supuesto que las preguntas siguen el mismo tipo de modelo en términos de la probabilidad de contestar correctamente.

⁴También rindieron la prueba de Ciencias Sociales.

⁵Corresponde a TIMSS 2011. Al estar en el hemisferio sur, Chile debe aplicar la prueba antes.

respuestas se analizarán desde dos ópticas, primero estudiando la calidad psicométrica de ambas pruebas y, posteriormente, construyendo la equivalencia entre ambas⁶.

El estudio psicométrico de ambas pruebas se realizará siguiendo las recomendaciones clásicas para el análisis de pruebas de logro estandarizadas (Hambleton y Swaminathan, 1985; Olson J. et al., 2009 y OECD, 2012). De esta manera, primeramente se estudiarán las pruebas en el marco de la Teoría Clásica de los Test (TCT) para luego examinarlas de manera separada y conjunta, en el marco de la Teoría de Respuesta al Ítem (TRI). Dentro de los análisis en el marco TCT se entregará la confiabilidad y el grado de consistencia interna de cada una de las formas o cuadernillos⁷ de cada prueba, como asimismo las dificultades relativas y el poder discriminante de las distintas preguntas. En el marco de la TRI se presentarán los parámetros bajo el mismo modelo probabilístico, lo que permitirá evaluar comparativamente el comportamiento de las preguntas de cada prueba en cuanto a sus características, y de esta manera se podrá concluir sobre la calidad estadística de ambas como instrumentos de medición para la población chilena.

Para establecer una equivalencia entre las escalas de medición de ambas pruebas, que se expresa en una tabla de equivalencia, se debe primero determinar la naturaleza y fuerza de la relación estadística existente entre ellas. Dependiendo de los resultados obtenidos y siguiendo las recomendaciones de la literatura especializada al respecto (Dorans, 2004), se procederá a ajustar un modelo matemático que permita alinear las escalas de ambas pruebas obteniéndose las equivalencias de sus puntajes. Asimismo, se ejecutará un procedimiento de re-muestreo a objeto de determinar las incertidumbres asociadas.

Este documento se estructura de la siguiente manera. En la siguiente sección se presenta el marco teórico y las técnicas estadísticas que se utilizarán tanto para el análisis psicométrico como para el alineamiento de las escalas de ambas pruebas. La tercera sección presentará un análisis descriptivo de los datos utilizados y en la cuarta sección se reportarán los resultados obtenidos. El documento finalizará discutiendo la implicancia de los resultados.

⁶Este documento presenta resultados obtenidos en Gaggero (2012) y Guzmán (2013), como asimismo en Agencia de Educación (2013a)

⁷La prueba Simce está dividida en 4 cuadernillos o formas, de 35 preguntas cada uno. La prueba TIMSS tiene 14 formas o cuadernillos con un número de preguntas que varía entre 22 y 29.

2. Marco Teórico y Metodologías Utilizadas

Actualmente las pruebas estandarizadas internacionales como PISA y TIMSS (OECD 2012, Olson et al., 2009) se analizan siguiendo los mismos principios generales⁸ :

- Evaluación de comportamiento psicométrico de preguntas usando indicadores descriptivos de la muestra⁹ o población objetivo.
- Evaluación y ajuste de modelo matemático con estimación de parámetros en el marco de la Teoría de Respuesta al Ítem.

A continuación se presentan brevemente los conceptos e indicadores que permiten el análisis psicométrico de las preguntas en el marco de la Teoría Clásica de los Test y para luego presentar lo propio con los indicadores psicométricos de calidad en el marco de la Teoría de Respuesta al Ítem. Se recuerda que en la Teoría Clásica de los Test el modelo subyacente relaciona de manera simple el rasgo o constructo¹⁰ que se desea medir en una persona, con su puntuación observada en una prueba, mediante:

$$X = V + \epsilon \quad (1)$$

Donde:

X : Puntuación observada de la persona.

V : Nivel de rasgo o puntuación “real” de la persona.

ϵ : Error de medición del rasgo en la prueba o test.

El error de medición se considera como una variable aleatoria de media cero y varianza finita, independiente del puntaje verdadero. Representa efectos que influyen en los resultados de la prueba, no relacionados con los conocimientos, contenidos o constructos que mide la prueba. Asimismo, se considera que si una persona rinde dos pruebas paralelas¹¹, los errores de medición asociados a ambas pruebas son independientes y por lo tanto, no correlacionados. Del modelo y supuestos anteriores se deducen las siguientes ecuaciones, donde X_1 y X_2 son los puntajes observados para dos pruebas que miden constructos distintos:

⁸En este caso se considera el análisis de la prueba definitiva cuyas preguntas se estudiaron previamente en una aplicación piloto.

⁹Usualmente obtenidos bajo TCT.

¹⁰Por ejemplo “logro en Matemática”.

¹¹Que tienen el mismo marco de referencia y mismas características estadísticas.

$$Cov(X, V) = Cov(V + \epsilon, V) \quad (2)$$

$$= Cov(V + V) \quad (3)$$

$$= Var(V) \quad (4)$$

$$Cov(X_1, X_2) = Cov(V_1 + \epsilon_1, V_2 + \epsilon_2) \quad (5)$$

$$= Cov(V_1 + V_2) \quad (6)$$

$$Var(X) = Cov(V + \epsilon, V + \epsilon) \quad (7)$$

$$= Var(V) \quad (8)$$

$$= Var(\epsilon) \quad (9)$$

Para la medición educacional se denomina **confiabilidad** de un test a la proporción de la varianza real que es explicada por la varianza observada (Thompson, 2003). Corresponde al λ descrito por la siguiente ecuación:

$$\lambda = \frac{Var(V)}{Var(X)} \quad (10)$$

Existen distintas maneras de estimarla confiabilidad según las características de los test:

Alfa de Cronbach: Este coeficiente estima el grado de alineamiento o consistencia que presentan los distintos ítems o preguntas del test¹². Se define de la siguiente forma:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_x^2} \right) \quad (11)$$

Donde:

n : Número de ítems del test.

$\sum_{i=1}^n \sigma_i^2$: Suma de las varianzas de los ítems.

σ_x^2 : Varianza de las puntuaciones en el test.

Observamos que el alfa de Cronbach depende del número de preguntas.

Coefficiente de Kuder-Richarson: Es el equivalente al alfa de Cronbach aplicado a test cuyos ítems son de carácter dicotómicos. En nuestro caso se utilizará el alfa de Cronbach para medir confiabilidad de las pruebas TIMSS y Simce.

¹²Es una cota inferior para confiabilidad.

Test-Retest: Considerando que dos test son paralelos cuando miden el mismo puntaje verdadero y solo difieren por el error de medición ¹³, se prueba que si dos test X_1 y X_2 son paralelos entonces:

$$\lambda = Corr(X_1, X_2) \quad (12)$$

Como un test es paralelo consigo mismo, se puede calcular la correlación entre los puntajes obtenidos en dos ocasiones diferentes. Sin embargo, es necesario resguardar que el test sea aplicado a dos muestras de individuos equivalentes. Este método apunta a la confiabilidad del test en el tiempo.

Dos mitades: Considerando que un mismo test puede dar nacimiento a dos test paralelos, se puede separar los diferentes ítems del test. Este criterio cobra mayor utilidad cuando la técnica de separación de ítems es capaz de extraer dos test con las mismas características estadísticas. Este método es indicativo de la confiabilidad interna del test.

2.1. Análisis de los ítems en Marco de TCT

En el marco de la TCT se calculan los siguientes estadísticos de las preguntas:

- **Índice de dificultad:** Es el porcentaje de respuesta correcta. Claramente depende del grado de competencia en la habilidad/contenido que mide la prueba de la muestra de alumnos que la rinde.
- **Índice de discriminación:** Es la correlación entre el resultado en el test y el resultado en la pregunta. Según la naturaleza de las preguntas (dicotómicas, politómicas, de respuesta abierta, etc.) y del resultado de la prueba, es la correlación que se usa (Pearson, Spearman, Punto biserial, Policórica, etc.).

Para cada una de las preguntas que conforman las pruebas Simce y TIMSS se obtendrá el índice de dificultad y el índice de discriminación, y para cada cuadernillo se entregará el alpha de Cronbach.

Para efectos de reporte, y siguiendo a Ebel (1972) y Ghiselli et al. (1981), se clasificarán¹⁴ los niveles de dificultad y discriminación según las Tablas 1 y 2.

¹³Que son independientes e idénticamente distribuidos

¹⁴Concordando la literatura, no existen en general estándares únicos en que las preguntas de un test debieran abarcar dificultad entre los valores 0,2 y 0,7, en tanto que las discriminaciones ser superiores a 0,2.

Tabla 1: Clasificación para el parámetro de Dificultad

Parámetro de Dificultad (D)	
Muy Díficil	$D < 0,3$
Díficil	$0,3 < D < 0,5$
Medio	$0,5 < D < 0,7$
Fácil	$0,7 < D < 0,8$
Muy Fácil	$D > 0,8$

Tabla 2: Clasificación para el parámetro de Discriminación

Parámetro de Discriminación (r)	
Negativa	$r < 0$
Pobre	$0 < r < 0,15$
Regular	$0,15 < r < 0,25$
Buena	$0,25 < r < 0,35$
Muy Buena	$r > 0,35$

2.2. Teoría de Respuesta al Ítem

Como se mencionó, en la TCT el análisis de las preguntas de la prueba se realiza condicional en la muestra que rinde dicha prueba. Esto ocasiona dificultades para la comparabilidad y seguimiento de los resultados de las pruebas estandarizadas (Hambleton y Swaminathan, 1985). Estas dificultades se sobrellevan al analizar las pruebas en el marco de la Teoría de Respuesta al Ítem. Recordamos que esta teoría se basa en los supuestos de:

- **Unidimensionalidad:** la prueba mide el nivel de habilidad en un solo rasgo o constructo.
- **Existencia de Curva Característica del Ítem (CCI):** la probabilidad de obtener una respuesta particular se relaciona funcionalmente con el nivel de habilidad en el rasgo o constructo medido.
- **Independencia local:** las preguntas se responden de manera independiente.

Para la verificación del supuesto de unidimensionalidad se realizará el ajuste de modelos factoriales a las matrices de correlación¹⁵ (San Martín y Manzi, 2003). Además se verificará la preponderancia del primer valor propio

¹⁵Tetracóricas cuando sea el caso.

sobre los otros, que se realizará analizando el gráfico de sedimentación y la razón entre el primer y segundo valor propio. También se analizará el porcentaje de la varianza explicada por todas las componentes, como asimismo la explicada solamente por la primera. Además, se verificará la pertinencia del análisis factorial usando el estadístico de Kaiser-Meyer-Olkin (KMO) y corroborando que esté dentro de rangos aceptables (Kaiser y Caffry, 1965):

$$KMO = \frac{\sum \sum_{j=k} r_{ij}^2}{\sum \sum_{j=k} r_{ij}^2 + \sum \sum_{j=k} s_{ij}^2} \quad (13)$$

Donde:

r_{ij}^2 : Correlación simple entre el ítem i y j .

s_{ij}^2 : Correlación parcial entre el ítem i y j .

Tabla 3: Clasificación estadístico de Kaiser-Meyer-Olkin (KMO)

KMO	Interpretación de los factores
Mayor que 0,7	Indica alta correlación entre las variables y, por lo tanto, es conveniente realizar el análisis factorial.
Entre 0,5 y 0,6	El grado de correlación es medio y el análisis factorial sería menos útil que en el caso anterior.
Menor que 0,5	Indica que el análisis factorial no resultaría una técnica útil.

Para la verificación de la existencia de la CCI se procede a ajustar un modelo logístico de tres parámetros¹⁶ :

$$P_i(\theta) = c_i + (1 + c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (14)$$

Donde θ representa la habilidad en el constructo que mide la pregunta i , D el parámetro de escala (habitualmente 1,7) y a_i , b_i y c_i son respectivamente los parámetros de discriminación, dificultad y pseudo-azar de la pregunta i ¹⁷.

Para la obtención de los parámetros se utiliza el *software* Parscale 4.1 que emplea el algoritmo EM¹⁸, con 41 puntos de cuadratura y valor crítico para la convergencia, obtenida a través del método de Newton-Raphson, igual al 0,005, para una cantidad de ciclos máximos igual a 100.

Considerando que los constructos o rasgo latente que evalúan ambas pruebas son similares (para TIMSS es un currículo internacional consensuado entre los países participantes y para Simce corresponde al currículo nacional vigente para 4.º año de educación básica) y que, además, comparten el mismo modelo de ajuste de CCI, se

¹⁶ Ambas pruebas fueron diseñadas bajo este mismo modelo.

¹⁷ Para mayor explicación sobre el modelo y la interpretación de los parámetros ver Hambleton y Swaminathan (1985) y en la página de la Agencia de Calidad, el documento Apunte N.º 6, junio 2013.

¹⁸ Expectation-Minimization Algorithm.

abrió la posibilidad de calibrar ambas pruebas de manera conjunta¹⁹. De esta forma, las preguntas se alinearon en una escala común, lo que permitió compararlas bajo un mismo referente de calidad psicométrica. Esto se realizó efectuando la estimación de los parámetros de los modelos probabilísticos de cada pregunta (CCI), como si la prueba TIMSS y Simce fuera una sola gran prueba. Cabe señalar que también se estimó el comportamiento psicométrico de las preguntas Simce dentro de la métrica de TIMSS, mediante la importación o anclaje de los valores originales de los parámetros de la prueba TIMSS (obtenidos de la población internacional). Los resultados de este último trabajo se presentaron en el documento Apunte N.º6. (Agencia de Calidad de la Educación, 2013b).

Para obtener las tablas de equivalencia entre las pruebas Simce y TIMSS se debe primeramente establecer el tipo de relación con que se va a modelar la equivalencia. Para ello se debe determinar la fuerza de la relación existente entre ambas pruebas, como también estudiar el grado de similitud entre sus distribuciones empíricas de resultados, ya que los distintos modelos a utilizar exigen diferentes supuestos acerca de la similitud en las distribuciones de las pruebas (Dorans, 2004 y Gempp, 2010). El grado de asociación se puede estudiar mediante las correlaciones de Pearson y la similitud de las distribuciones se puede evaluar gráficamente o usando estadísticos no-paramétricos como Kolmogorov-Smirnov.

3. Análisis Descriptivos

A continuación se describen las características composicionales de la muestra de estudiantes que participó en la prueba TIMSS y que también rindieron la prueba Simce, se dispone del resultado de 5.148 estudiantes de 4.º básico. La muestra de estudiantes que rindió la prueba TIMSS es una muestra estratificada y, en consecuencia, para hacer inferencias a la población se debe aplicar ponderadores o pesos a los resultados. Sin embargo, para el análisis psicométrico no es necesario aplicar los ponderadores ya que no se pretende inferir a la población sino que solamente estudiar el comportamiento de las preguntas. Sin perjuicio de lo anterior, en el siguiente análisis descriptivo se aprecia que los principales grupos de interés de la población chilena están representados.

En la muestra, el 47% de los estudiantes pertenece a establecimientos particulares subvencionados, la segunda mayoría corresponde a estudiantes que asisten a establecimientos municipales, seguidos finalmente por los de colegios particulares pagados (ver Figura 1 y Tabla 4). En la muestra también hay estudiantes de establecimientos rurales, aunque con una presencia pequeña (Figura 2 y Tabla 5). La distribución según género muestra

¹⁹La calibración corresponde a la estimación del valor de los parámetros de las preguntas.

porcentajes similares para hombres y mujeres (Figura 3 y Tabla 6). En cuanto al GSE²⁰, es predominante el grupo medio, seguido por el medio bajo, medio alto, alto y, finalmente, el bajo (Figura 4 y Tabla 7). De los 5.148 estudiantes se aprecia que casi el 50% pertenece a la región Metropolitana de Santiago (ver Tabla 8). A continuación se grafican y describen las frecuencias en detalle:

Figura 1: Distribución de la muestra por Dependencia

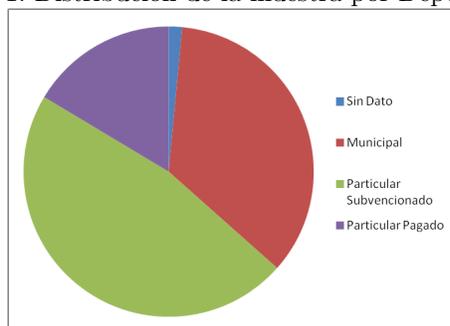
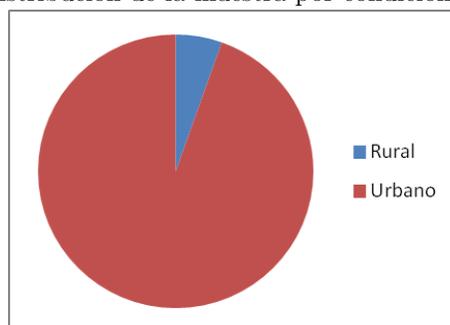


Tabla 4: Frecuencia por Dependencia

Dependencia educacional	N.º de alumnos	%
Municipal	1.806	35,1
Particular subvencionada	2.421	47
Particular pagado	844	16,4
Sin dato	77	1,5
Total	5.148	100,0

Figura 2: Distribución de la muestra por condición de ruralidad



²⁰Los GSE son los definidos por SIMCE 4.º básico 2010 y corresponden al establecimiento al que asisten los estudiantes.

Tabla 5: Frecuencia por condición de ruralidad

Ruralidad	N.º de alumnos	%
Rural	281	5,5
Urbano	4.867	94,5
Total	5.148	100,0

Figura 3: Distribución de la muestra por género

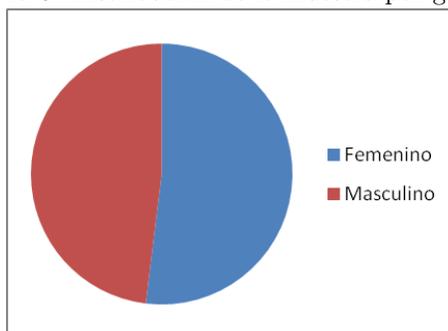


Tabla 6: Frecuencia por género

Género	N.º de alumnos	%
Femenino	2.674	51,9
Masculino	2.474	48,1
Total	5.148	100,0

Figura 4: Distribución de la muestra por GSE

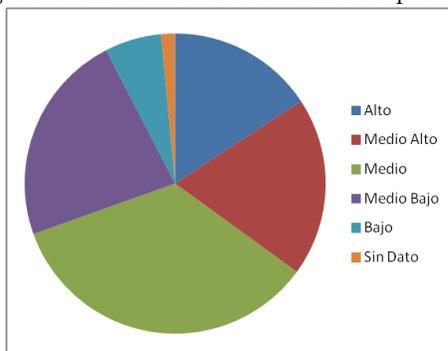


Tabla 7: Frecuencia por GSE

GSE	N.º de alumnos	%
Alto	812	15,8
Medio alto	991	19,3
Medio	1.777	34,5
Medio bajo	1.178	22,9
Bajo	313	6,1
Sin dato	77	1,5
Total	5.148	100,0

Tabla 8: Distribución porcentual de los estudiantes considerando regiones

Región	N.º de alumnos	%
Sin Dato	77	1,5
Región de Tarapacá	65	1,3
Región de Antofagasta	96	1,9
Región de Atacama	145	2,8
Región de Coquimbo	194	3,8
Región de Valparaíso	546	10,6
Región Metropolitana de Santiago	2.519	48,9
Región del Libertador General Bernardo O'Higgins	175	3,4
Región del Maule	215	4,2
Región del Biobío	618	12,0
Región de La Araucanía	175	3,4
Región de los Ríos	98	1,9
Región de los Lagos	163	3,2
Región de Aysén del General Carlos Ibáñez del Campo	25	0,5
Región de Magallanes y de la Antártica Chilena	37	0,7
Total General	5.148	100,0 %

Esta muestra de estudiantes es una muestra estratificada bi-etápica²¹. En la primera etapa se seleccionaron establecimientos educacionales dentro de estratos²² con una probabilidad proporcional al número de estudiantes. En la segunda etapa se escogió aleatoriamente un curso dentro del establecimiento.

La prueba Simce de Matemática 4.º básico del año 2010 consta de 4 formas o cuadernillos que contienen 35 preguntas cada uno, donde 20 preguntas son las mismas para cada cuadernillo. Estas preguntas sirven para la equiparación de la prueba 2010 con las pruebas de años anteriores y también para establecer la métrica común entre cuadernillos²³. Cada estudiante contesta un cuadernillo.

La prueba TIMSS de Matemática 4.º básico de 2011 consta de 14 cuadernillos con un número variable de preguntas que oscilan entre 22 y 29. Cada cuadernillo está conformado por dos grupos de preguntas llamados bloques. Los bloques se repiten solo en dos cuadernillos²⁴. Cada estudiante responde un solo cuadernillo.

²¹La muestra fue escogida en el Centro de Análisis de la prueba TIMSS.

²²Los estratos se conforman usando predeterminados por características que están asociadas al rendimiento educacional.

²³Este diseño de pruebas recibe el nombre de “Diseño con ítems comunes”.

²⁴Este diseño de pruebas recibe el nombre de “Diseño con bloques rotados”.

4. Resultados

En esta sección se presentan los resultados del análisis de las preguntas en el marco de la TCT, el resultados de la calibración de ambas pruebas conjuntamente y la construcción de la tabla de equivalencias entre las escalas de las pruebas TIMSS y Simce.

En el marco de la TCT, se presentan tablas con un resumen de la siguiente información para cada forma o cuadernillo de cada prueba: el número de preguntas, el número de estudiantes que contestan, el alpha de Cronbach, la distribución de dificultad y discriminación dentro de los rangos presentados (ver Tabla 9). También se presentan los resultados del Análisis Factorial Exploratorio: varianza explicada, número de componentes, varianza explicada por primera componente, magnitud relativa de dos primeros valores propios, porcentaje de ítems que cargan en primera componente y KMO (ver Tabla 10). En el Anexo se presenta el detalle para cada forma, junto con su gráfico de sedimentación.

Tabla 9: Estadísticos Bajo TCT para cada prueba y forma

Prueba/ Forma	N.º de ítems	N.º Alumnos	Alpha de Cronbach	Dificultad TCT					Discriminación TCT			
				MD	D	Med	F	MF	P	Reg	B	MB
SIMCE/C	35	1.419	0,896	3 %	23 %	43 %	20 %	11 %	0 %	0 %	14 %	86 %
SIMCE/D	35	1.383	0,917	0 %	23 %	37 %	31 %	9 %	0 %	0 %	3 %	97 %
SIMCE/E	35	1.354	0,927	0 %	20 %	49 %	43 %	9 %	0 %	0 %	0 %	100 %
SIMCE/G	35	1.295	0,945	0 %	20 %	40 %	26 %	14 %	0 %	0 %	3 %	97 %
TIMSS/1	22	399	0,836	32 %	36 %	23 %	5 %	5 %	5 %	9 %	18 %	68 %
TIMSS/2	26	396	0,845	23 %	38 %	27 %	4 %	8 %	4 %	8 %	15 %	73 %
TIMSS/3	26	390	0,811	23 %	42 %	19 %	8 %	8 %	8 %	12 %	12 %	69 %
TIMSS/4	24	396	0,830	29 %	33 %	17 %	21 %	0 %	8 %	17 %	8 %	67 %
TIMSS/5	26	402	0,860	15 %	27 %	23 %	27 %	8 %	8 %	8 %	0 %	85 %
TIMSS/6	27	405	0,860	11 %	33 %	37 %	11 %	7 %	7 %	4 %	4 %	85 %
TIMSS/7	25	399	0,821	20 %	40 %	36 %	4 %	0 %	8 %	12 %	16 %	64 %
TIMSS/8	26	399	0,802	23 %	42 %	19 %	8 %	8 %	4 %	8 %	42 %	46 %
TIMSS/9	29	390	0,778	28 %	38 %	21 %	3 %	10 %	3 %	21 %	55 %	21 %
TIMSS/10	29	401	0,809	31 %	38 %	24 %	0 %	7 %	0 %	17 %	55 %	28 %
TIMSS/11	26	391	0,801	31 %	46 %	19 %	0 %	4 %	8 %	12 %	27 %	54 %
TIMSS/12	26	396	0,837	23 %	42 %	23 %	4 %	8 %	8 %	0 %	27 %	65 %
TIMSS/13	26	406	0,853	19 %	42 %	19 %	12 %	8 %	8 %	0 %	23 %	69 %
TIMSS/14	22	396	0,825	27 %	45 %	9 %	14 %	5 %	14 %	5 %	18 %	64 %

Nota: MD: Muy Difícil; D: Difícil; Med: Medio; F: Fácil; MF: Muy Fácil; P: Pobre; Reg: Regular; B: Buena; MB: Muy Buena.

Tabla 10: Estadísticos del Análisis factorial para cada prueba y forma.

Prueba/Forma	% Varianza	N.º de	% Varianza	Razón entre	ítems cargan	KMO
	Total	Componentes	Explicada	1.er y 2.do	1.ra	
	Explicada		1.ra Componente	valor propio	Componente	
SIMCE/C	51	5	36	7,033	91	0,954
SIMCE/D	53	5	38	8,371	94	0,964
SIMCE/E	52	5	38	8,878	97	0,981
SIMCE/G	50	4	40	10,463	100	0,981
TIMSS/1	64	5	39	5,396	86	0,89
TIMSS/2	63	7	35	5,118	85	0,883
TIMSS/3	67	8	32	4,621	81	0,846
TIMSS/4	68	7	35	4,554	83	0,887
TIMSS/5	72	8	37	5,335	81	0,885
TIMSS/6	63	6	36	4,853	81	0,895
TIMSS/7	63	7	32	4,116	76	0,872
TIMSS/8	64	8	27	3,796	73	0,837
TIMSS/9	67	9	25	3,449	76	0,756
TIMSS/10	65	9	27	3,58	52	0,798
TIMSS/11	65	8	29	4,533	65	0,842
TIMSS/12	64	7	34	5,051	85	0,88
TIMSS/13	62	6	36	5,747	85	0,877
TIMSS/14	70	7	38	5,852	77	0,86

Observamos que la confiabilidad es alta para cada cuadernillo en ambas pruebas, siendo las formas de la prueba Simce las que presentan los más altos valores. Como se mencionó, este índice es sensible al número de preguntas y las formas Simce contienen en promedio 9 preguntas más que los cuadernillos TIMSS. Solo el cuadernillo 9 de la prueba TIMSS presenta una confiabilidad inferior a 0,8 aunque cercana a este valor (0,78). Concluimos entonces que ambas pruebas capturan gran parte de la variabilidad real subyacente en el resultado de los estudiantes.

Observamos que los cuadernillos de la prueba TIMSS se les hacen más difíciles a esta muestra de alumnos ya que contienen en promedio un 24% ²⁵ de preguntas “Muy Difíciles”, en tanto que dichas preguntas casi no existen en los cuadernillos de la prueba Simce. Algo similar sucede para las preguntas “Difíciles”, dado que los cuadernillos TIMSS contienen en promedio un 39% de preguntas de este tipo versus un 22% de la prueba Simce. Los cuadernillos Simce tienen mayor porcentaje de preguntas de dificultad “Medio” y “Fácil” que los cuadernillos TIMSS, y en ambas pruebas el porcentaje de preguntas “Muy Fácil” es similar.

Considerando los propósitos de la prueba Simce²⁶, la distribución de preguntas en las distintas categorías de dificultad es consistente con las recomendaciones. Aunque es necesario disponer de preguntas en todo el espectro de dificultad, se requiere concentrar en el rango de “Difícil” a “Fácil”, para evitar problemas de desaliento o desmotivación por parte de los estudiantes.

La discriminación medida en el enfoque clásico se mueve en muy buenos rangos para ambas pruebas. El 100% de las preguntas de los cuadernillos de la prueba Simce se mueve en los rangos “Buena” o “Muy buena”. En la prueba TIMSS, en promedio, el 84% de las preguntas de los distintos cuadernillos se mueve en ese rango.

Podemos concluir que, desde un punto de vista de la Teoría Clásica de los Test, ambas pruebas satisfacen criterios de calidad, observándose que la prueba Simce estaría mejor adaptada para la población chilena que la prueba TIMSS. Esto no es sorprendente ya que esta última debe adaptarse a una población objetivo con mayores rangos de heterogeneidad en cuanto a dominio curricular que la chilena, pues es aplicada a 53 países de todo el mundo (ver Agencia de Calidad, 2012b).

Los resultados de los análisis factoriales para ambas pruebas son muy satisfactorios. Los índices de Kaiser-Meyer-Olkin (KMO) son en promedio 0,97 para los cuadernillos de Simce y 0,86 para los cuadernillos de TIMSS, ambos

²⁵Alrededor de seis preguntas.

²⁶Que mide apropiación curricular.

muy superiores a lo definido como “Bueno”. Esto indica que las preguntas de cada cuadernillo comparten una estructura de factores subyacentes fuerte.

La varianza total de la dispersión de las respuestas explicada por el análisis factorial está dentro de rangos aceptables, siendo el valor promedio de 52 % para los cuadernillos Simce y un poco más alta (66 %) para los cuadernillos de la prueba TIMSS. El número de componentes o factores que se distinguen parece un poco alto, con una moda de 5 para las formas de la prueba Simce y de 7 para las formas de la prueba TIMSS. Esto, que podría parecer indicativo del no cumplimiento de unidimensionalidad se ve contrarrestado por el análisis de la razón entre el primer y segundo valor propio. Se observa que todas estas razones son superiores a 3,0, que es el punto de corte aceptado para la preponderancia de una componente por sobre las otras. De hecho, la razón promedio para las formas de la prueba Simce es 8,7 y para las formas de la prueba TIMSS, 4,7. Esto se ratifica por el porcentaje de preguntas que están alineadas o cargan en la primera componente. Para los cuadernillos de la prueba Simce dicho porcentaje promedio es 96 % y para la prueba TIMSS es de 78 %.

Los resultados anteriores permiten concluir que tanto las preguntas de la prueba Simce como las de la prueba TIMSS están midiendo un constructo único y que se justifica sobremanera proceder a un análisis bajo la Teoría de Respuesta al Ítem (TRI).

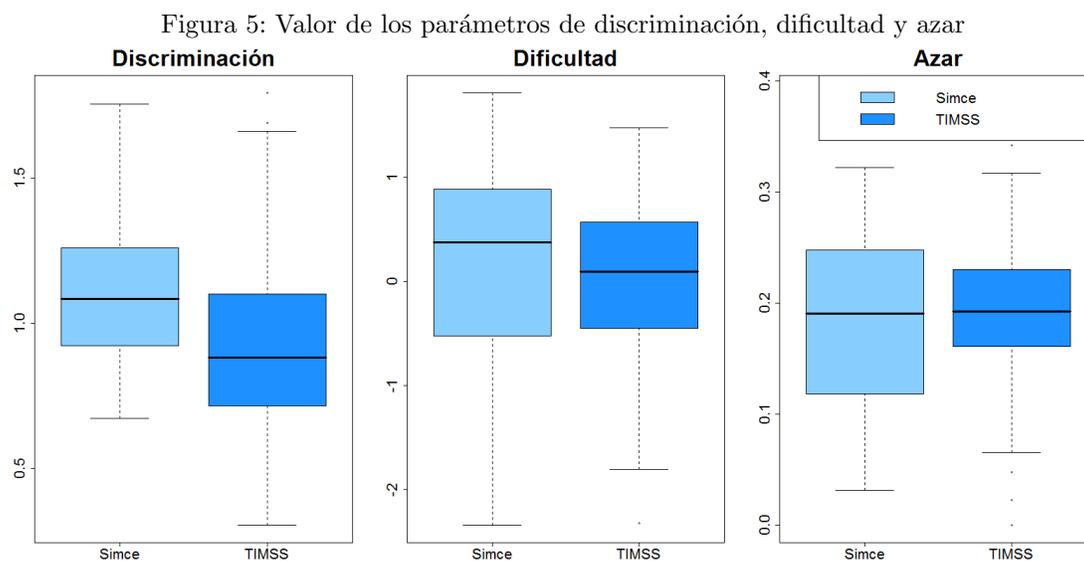
En el marco de la TRI se presentarán conjuntamente los valores²⁷ de los parámetros de discriminación, dificultad y azar del modelo logístico de tres parámetros basado en la calibración conjunta. Finalmente se presentará el procedimiento utilizado para la construcción de la tabla de equivalencia entre ambas pruebas.

4.1. Análisis de Calibración Conjunta

Para analizar el comportamiento de las preguntas de ambas pruebas bajo un mismo referente y comparar sus calidades psicométricas se procede a realizar una calibración conjunta de ambas pruebas, utilizando el *software* Parscale 4.1.

En La Figura 5 los siguientes gráficos de caja se muestra el valor de los parámetros de discriminación, dificultad y azar obtenidos de esta manera.

²⁷Con la estimación de su incerteza.



Observamos que los valores de los tres parámetros cubren rangos similares en ambas pruebas. Se nota que para los parámetros de discriminación y azar, los valores de la prueba Simce están contenidos en el rango de la prueba TIMSS. Para el parámetro de dificultad, los valores de la prueba Simce están un poco por debajo de los de la prueba TIMSS, aunque las preguntas más fáciles pertenecen a la prueba TIMSS.

A continuación se examinan los parámetros de manera gráfica. Para cada uno se presenta el histograma de sus valores, el gráfico de sus errores y la participación de cada prueba en según quintiles. Además se muestra un gráfico donde se identifican la preguntas con sus pruebas de origen. Los puntos de corte para los quintiles²⁸ de discriminación se presentan en la Tabla 11; para los quintiles de dificultad se presentan en la Tabla 12; y para el parámetro de azar se usarán los puntos de corte que se muestran en la Tabla 13.

A continuación se examina comparativamente el valor de cada tipo de parámetro según prueba, comenzando con el de **discriminación**, siguiendo con el de **dificultad** para finalizar con el de **azar**. Si las dos pruebas tienen similares características psicométricas, se esperaría que los valores de los parámetros para ambas fueran similares, lo mismo para sus errores de estimación y la distribución de preguntas en quintiles.

²⁸Los puntos de corte para quintiles para los distintos parámetros, se obtuvieron de determinar los número que dejaban al 20% de los preguntas de las pruebas Simce y TIMSS conjuntamente.

Tabla 11: Puntos de corte para quintiles de Discriminación

Quintil	Puntos de Corte
1	$\leq 0,7310$
2	0,7310 - 0,8822
3	0,8822 - 1,0223
4	1,0223 - 1,1985
5	$> 1,1985$

Tabla 12: Puntos de corte para quintiles de Dificultad

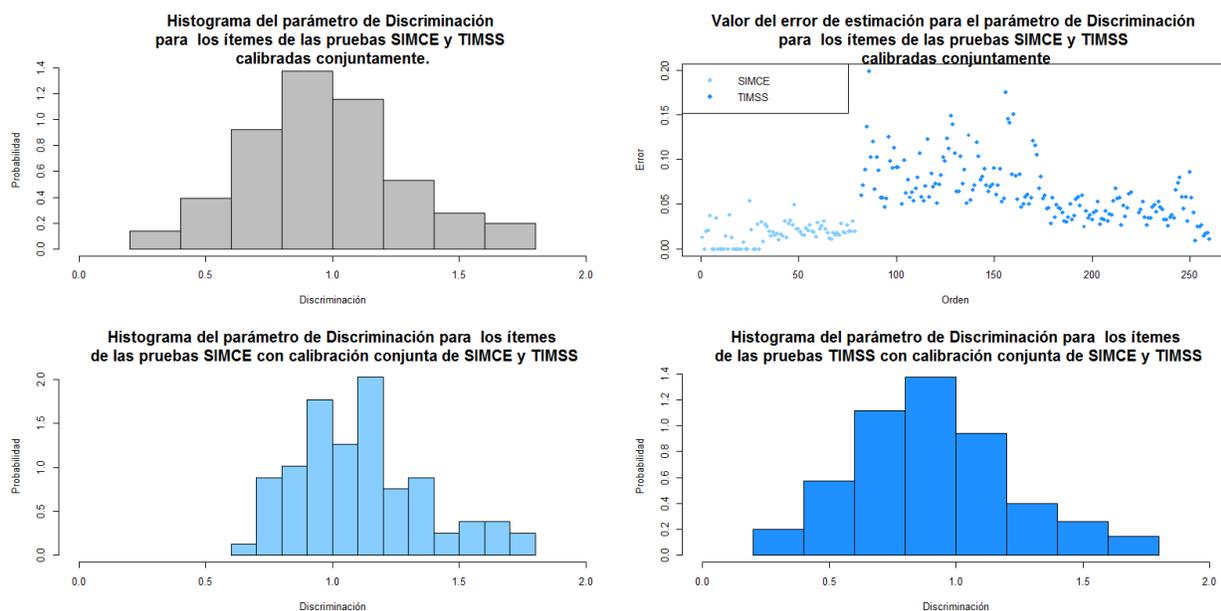
Quintil	Puntos de Corte
1	$\leq -0,5893$
2	-0,5893 - -0,0644
3	-0,0644 - 0,4134
4	-0,4134 - 0,7199
5	$> -0,7199$

Tabla 13: Puntos de corte para quintiles de Azar

Quintil	Puntos de Corte
1	$\leq 0,1322$
2	0,1322 - 0,1754
3	0,1754 - 0,2115
4	0,2115 - 0,35
5	$> 0,35$

Parámetro de discriminación

Figura 6: Error parámetro de discriminación pruebas Simce y TIMSS, en conjunto y por separado



En la Figura 6 observamos que tanto las preguntas Simce como TIMSS tienen discriminaciones en rangos similares. La distribución del parámetro de discriminación toma valores que en la gran mayoría son superiores a 0,5²⁹ (las preguntas con discriminación inferior a 0,5 pertenecen a la prueba TIMSS). Los errores de estimación de este parámetro se concentran en rangos aceptables, siendo en la mayoría de los casos inferior a 0,3. Asimismo, al examinar la distribución de las preguntas en quintiles de discriminación (ver gráfico en la Figura 7), no se observa que una u otra prueba concentre los valores altos o bajos (salvo para el primer quintil de la prueba TIMSS).

En el gráfico de la Figura 8 se ratifica el hecho que ambas pruebas contienen preguntas en donde el parámetro de discriminación varía en el mismo rango. Es decir, ambas pruebas comparten similares características de discriminación³⁰.

²⁹Considerando que la discriminación representa una pendiente, son deseables valores altos superiores a 0,5.

³⁰Que en el marco de TRI indica la capacidad de distinguir sujetos de diferentes habilidades.

Figura 7: Distribución parámetro de discriminación según quintiles

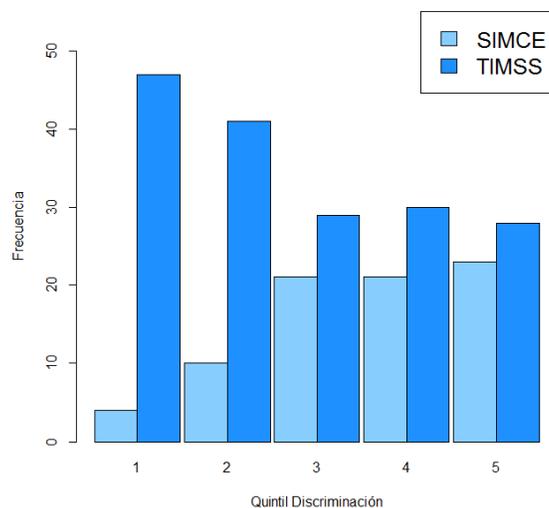
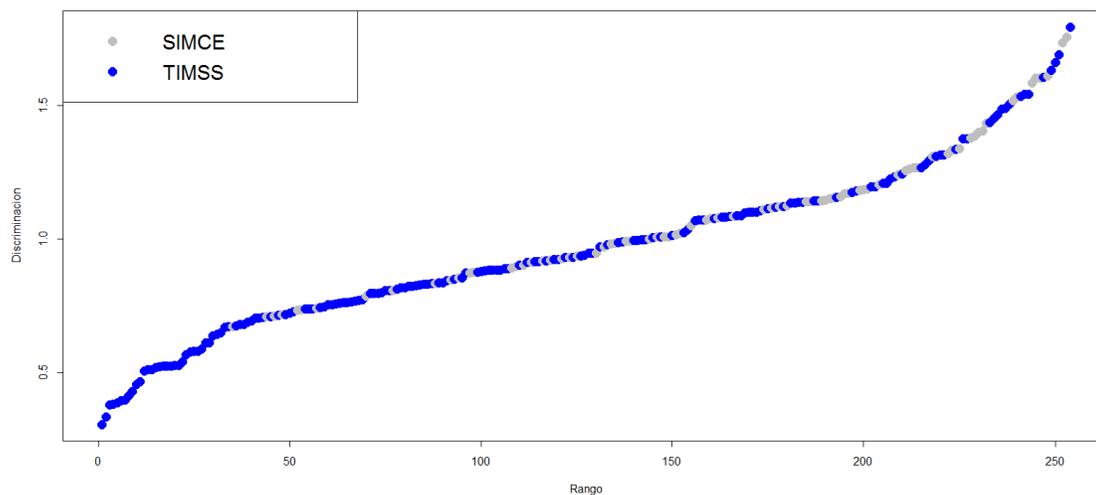


Figura 8: Orden parámetro de discriminación Simce-TIMSS

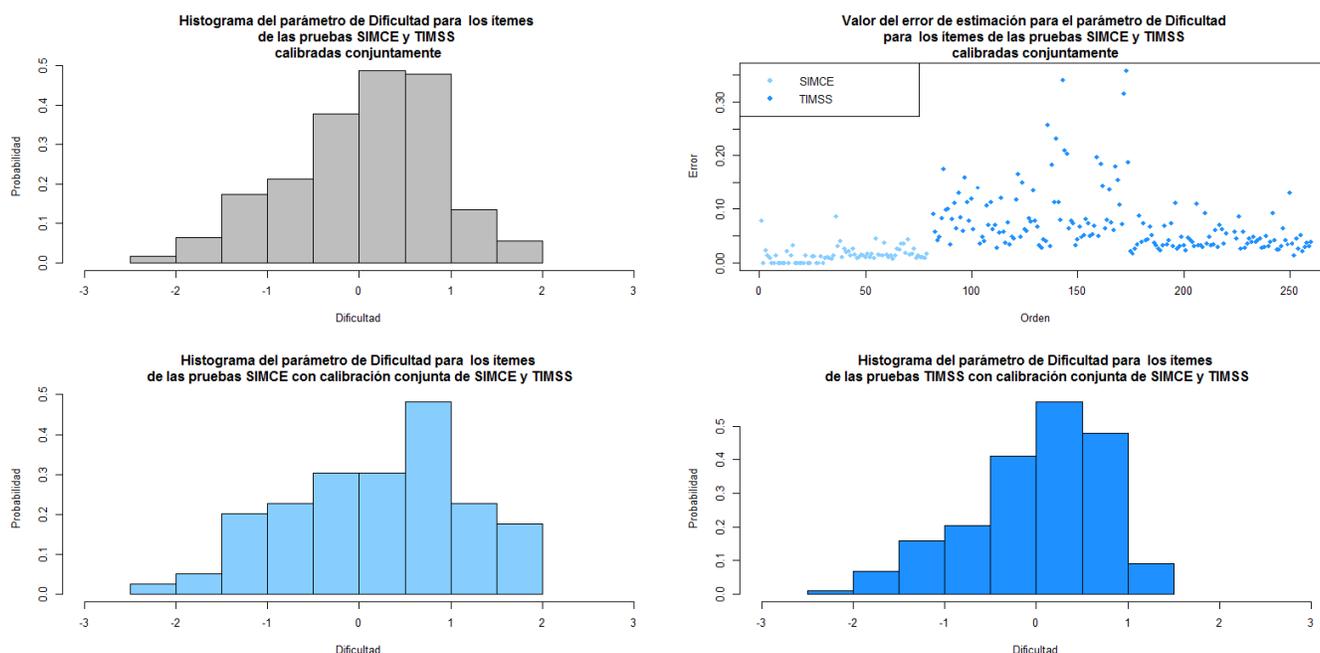


Parámetro de dificultad

En ambas pruebas el parámetro de dificultad se mueve en rangos similares (Figura 9). En este caso, los errores de estimación son notoriamente menores para las preguntas Simce que para las preguntas TIMSS.

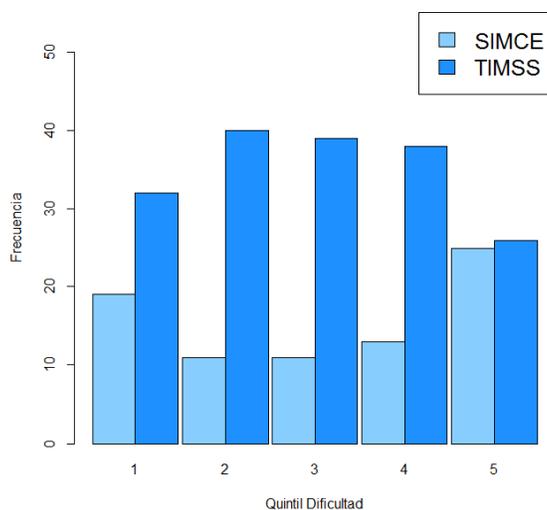
Del gráfico de la Figura 10 se observa que las preguntas Simce tienden a concentrarse en los primeros quintiles,

Figura 9: Error parámetro de dificultad pruebas Simce y TIMSS, en conjunto y por separado



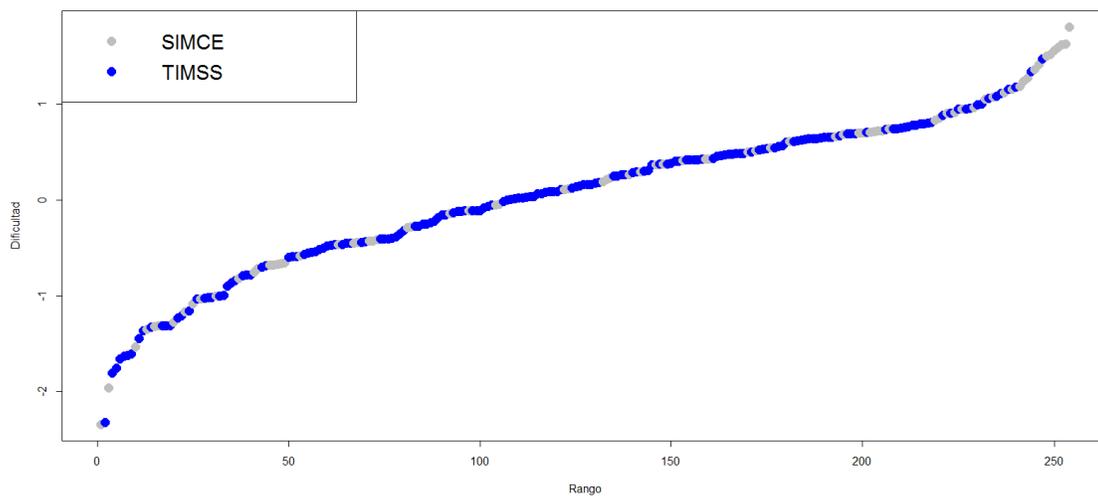
por lo que a los estudiantes les resultarían más fáciles. Se debe recordar que es deseable que en una prueba analizada en el marco de TRI existan preguntas en todo el rango de dificultad, lo que ocurre para ambas pruebas.

Figura 10: Distribución parámetro de dificultad según quintiles



Pese a lo anterior, se observa si bien las preguntas Simce tienden a concentrarse en la parte más fácil de la escala también existen preguntas Simce difíciles.

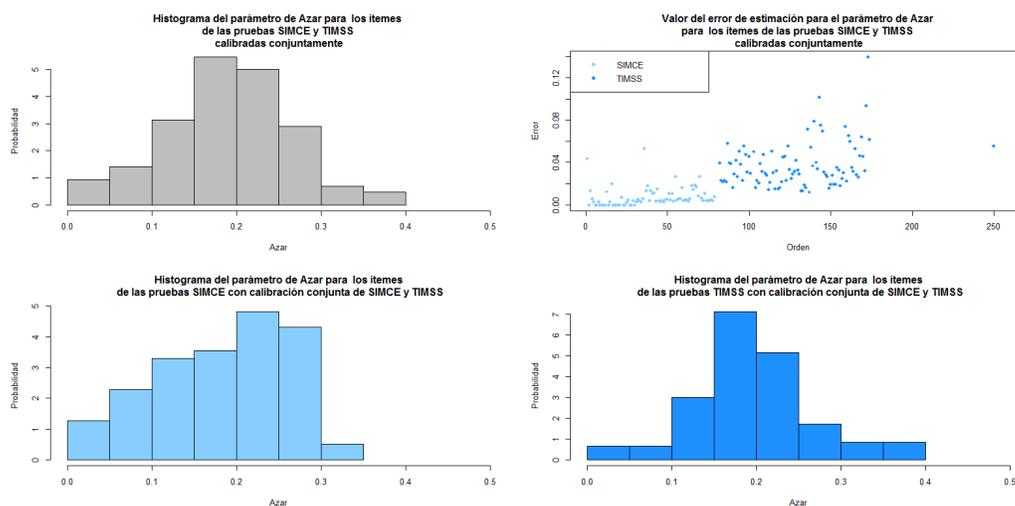
Figura 11: Orden parámetro de dificultad Simce-TIMSS



Parámetro de azar

Aquí también se observa que la distribución de los valores no difiere mayormente entre ambas pruebas. Sin embargo los valores de los errores de estimación son bastante más bajos para las preguntas de la prueba Simce que para las preguntas de la prueba TIMSS. Examinando el gráfico de quintiles de la Figura 14 y la Figura 13 se aprecia que, si bien el valor de los parámetros de azar de las pruebas Simce tiende a concentrarse en la parte media, ambas pruebas contienen preguntas con valor del parámetro de azar alto y bajos³¹.

Figura 12: Error parámetro de azar pruebas Simce y TIMSS, en conjunto y por separado



³¹En este caso es deseable valores cercanos a cero del parámetro de azar.

Figura 13: Distribución parámetro de azar según quintiles

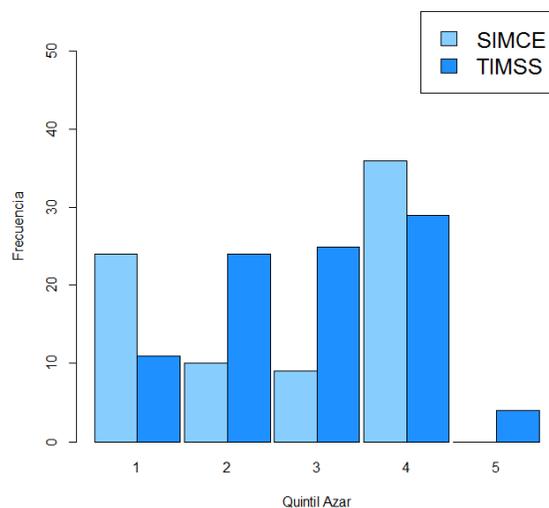
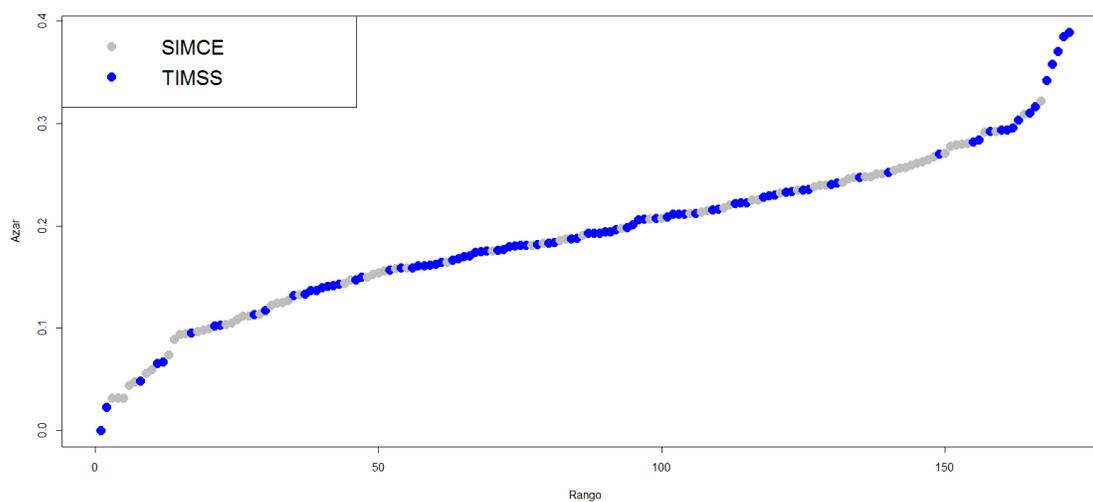


Figura 14: Orden parámetro de azar Simce-TIMSS



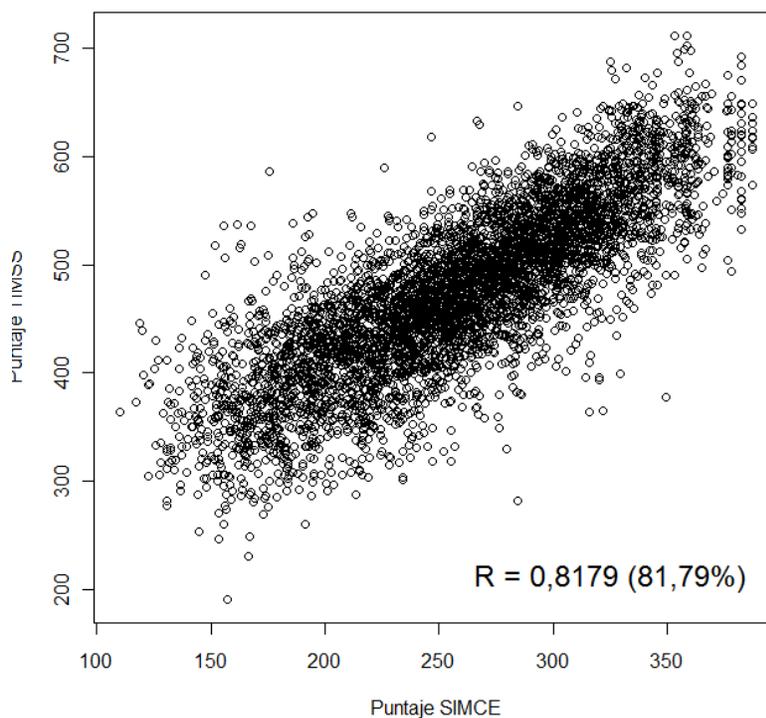
4.2. Tabla de equivalencia

Para la construcción de la tabla de equivalencia, primeramente se analizó la correlación existente entre los puntajes de ambas pruebas. En el gráfico de la Figura 15 se observa el resultado de cada alumno de la muestra en ambas pruebas³².

³²Se consideró el promedio de los “plausibles value” como el resultado de la prueba TIMSS.

En la Tabla 14 se presentan las correlaciones entre los Plausible Values (PV) de la prueba TIMSS y el puntaje Simce para la muestra de estudiantes que contestaron ambas pruebas.

Figura 15: Asociación Puntajes Simce-TIMSS



El rango de estas correlaciones se da con valores bastante altos por lo que justificaría realizar un alineamiento entre ambas pruebas (Dorans, 2004; Gempp, 2010). La distribución de los puntajes se describe en el gráfico de percentiles (Figura 16) y los histogramas de la Figura 17.

A partir de los histogramas se aprecia nítidamente que las distribuciones no provienen de la misma familia. Esto es ratificado por el Test de Kolmogorov-Smirnov donde se rechaza la hipótesis nula de que los puntajes de ambas pruebas provengan de la misma distribución (p-value =0,00001).

Corresponde entonces utilizar una técnica equipercantil para construir la tabla de correspondencia. En esta técnica se construye una función de correspondencia entre los puntajes de ambas pruebas, aplicando la ecuación:

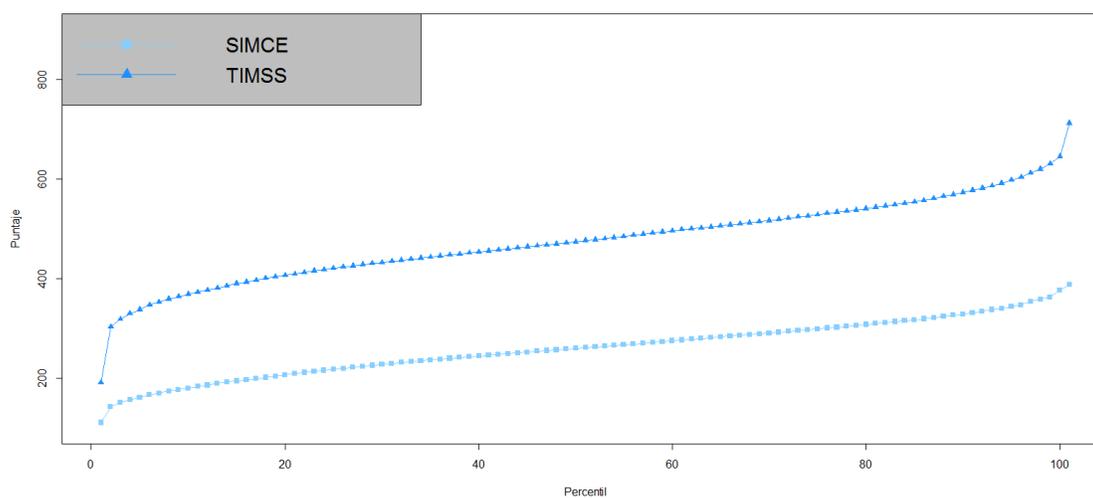
$$e_Y(x) = G^{-1}[F(x)] \quad (15)$$

Donde F es la función de distribución de los puntajes de la prueba TIMSS y G es la función de distribución de

Tabla 14: Correlación PV-TIMSS-Puntaje Simce

Plausible Value	Correlación
PV1	0,7731
PV2	0,7851
PV3	0,7781
PV4	0,7821
PV5	0,7771
Promedio PV	0,8179

Figura 16: Percentiles de puntajes Simce y TIMSS



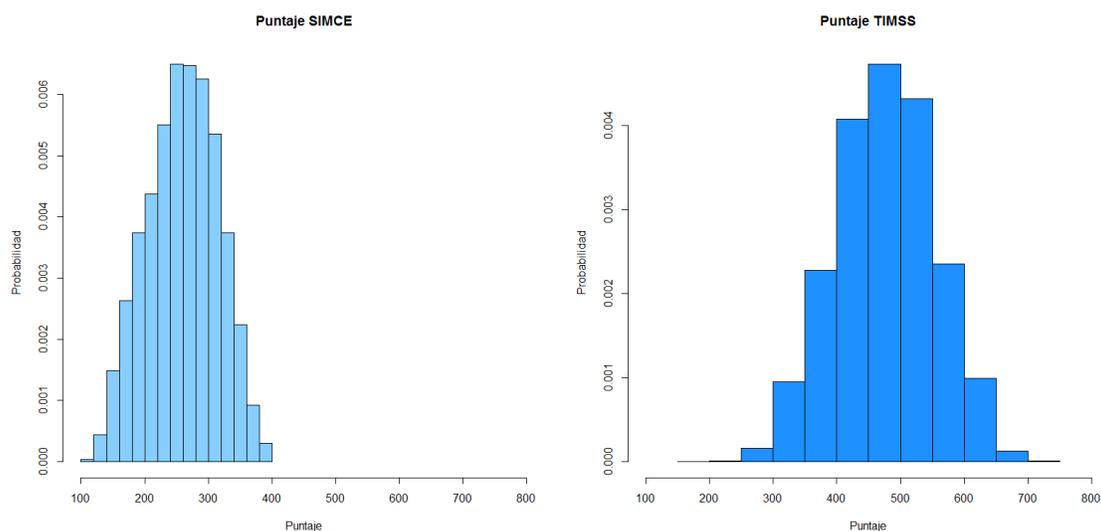
la prueba Simce. Se emplean para este efecto los *software* SPSS y LEGS³³.

Para examinar los errores de incertidumbre asociados a la tabla de concordancia, se desarrolla un procedimiento de re-muestreo de la muestra de estudiantes de TIMSS 2011, basado en técnicas de Bootstrap. Si bien las muestras de TIMSS están muy bien adaptadas para estimar variabilidades usando BRR³⁴, que es un procedimiento tipo Jackknife, se puede demostrar que los métodos basados en técnicas de Bootstrap se comportan mejor para la estimación de parámetros que provienen de funciones no-suaves de los datos (Efron, 1982). Este es el caso de los percentiles, por lo que en esta situación es la técnica más apropiada.

³³Obtenido del sitio web de Center for Advanced Studies in Measurement and Assessment (CASMA)

³⁴Balanced Replicated Resampling

Figura 17: Histogramas de puntajes Simce y TIMSS



En el Anexo se entrega para cada percentil la correspondencia entre los puntajes TIMSS y Simce. En la Figura 18 se plasma dicha correspondencia.

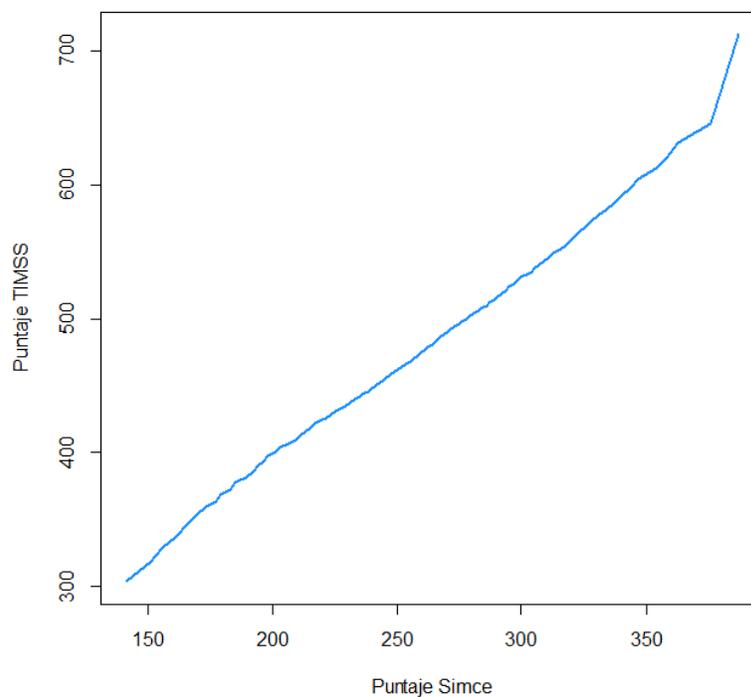
Se puede construir una tabla con los puntos de referencia (o *benchmark*), que entregan la prueba TIMSS y su puntaje Simce asociado. La Tabla 15 muestra esa relación.

Tabla 15: Correspondencia entre puntajes Simce y TIMSS con su error o incertidumbre para los Puntos de referencia de la prueba TIMSS

Puntos de Referencia	TIMSS	Simce	Error
Avanzado (Advanced)	625	361	2,1
Alto (High)	550	314	1,5
Intermedio (Intermediate)	475	260	1,4
Bajo (Low)	400	201	1,7

En la prueba Simce de cuarto básico 2012 de Matemática, el promedio nacional fue 261, lo que correspondería a un puntaje TIMSS de 476, que lo ubica en el segundo punto de referencia. Los actuales estándares de aprendizaje para Simce se pueden relacionar con los puntajes TIMSS utilizando la tabla de equivalencias. En la Tabla 16 se

Figura 18: Correspondencia entre puntaje Simce y puntaje TIMSS



presentan los puntos de corte que definen los Estándares de Aprendizaje para Cuarto Básico en Matemática y su correspondiente puntaje en la prueba TIMSS³⁵.

Tabla 16: Puntos de Corte para los Estándares de Aprendizaje de prueba Simce de Matemática de cuarto básico.

Nivel de Aprendizaje	Simce	TIMSS
Adecuado	295	524
Elemental	245	455

Observamos que el Nivel Elemental de los Estándares, se ubicaría entre los puntos de referencia bajo e intermedio, y el Nivel de aprendizaje Adecuado, estaría entre los puntos de referencia intermedio y alto.

³⁵El puntaje describe la cota inferior del rango. Se alcanza el nivel Elemental cuando se obtiene 245 puntos o más, y el nivel Adecuado con 295 puntos o más.

5. Conclusiones

El análisis en el marco de la TCT indicó que cada cuadernillo o forma de las pruebas Simce o TIMSS tiene una confiabilidad adecuada, ya que los valores del **alpha de Cronbach** son altos, siendo los de la prueba Simce algo superiores. Asimismo, se concluyó que cada forma o cuadernillo de ambas pruebas contiene preguntas de variado rango de **dificultad**, siendo las preguntas TIMSS un poco más difíciles para los estudiantes chilenos que las de Simce. El grado de **discriminación** de las preguntas en los cuadernillos³⁶ es adecuado en ambas pruebas, sin embargo, es mejor para la prueba Simce.

En el marco TRI se observó que se puede verificar la unidimensionalidad³⁷, medida tanto por los cuadernillos Simce como TIMSS, presentándose en ambas pruebas una fuerte componente única. Esta es particularmente importante en la prueba Simce. La calibración conjunta de las dos pruebas arrojó que el ajuste del modelo para ambas pruebas era bueno, siendo los errores de estimación menores para la prueba Simce que para la prueba TIMSS.

De lo anterior se puede concluir que ambas pruebas se comportan similarmente desde un visto psicométrico y que la prueba Simce está mejor adaptada a la población chilena, dado que se obtienen mejores ajustes y menores errores de estimación. Lo anterior no es sorprendente ya que la prueba TIMSS pretende medir un currículo consensuado y por lo tanto, debe adaptarse a una población más heterogénea.

El análisis psicométrico justifica la construcción de una tabla de equivalencia entre las escalas de ambas pruebas. Con la tabla de equivalencia generada, cada establecimiento educacional chileno puede determinar cómo sería su puntaje promedio TIMSS y de esta manera, tiene un referente internacional con el cual compararse.

³⁶Que en este caso es una correlación.

³⁷Que significa que cada prueba mide un constructo o rasgo latente único.

Bibliografía

- [1] AGENCIA DE CALIDAD DE LA EDUCACIÓN. (2012). *Resultados TIMMS Chile*. Santiago: IEA TIMMS & PIRLS. International Study Center.
- [2] AGENCIA DE CALIDAD DE LA EDUCACIÓN. (2013a). Apunte N.º 4 Alineamiento de las Pruebas de Matemática SIMCE 2010 y Matemática TIMSS 2011 para Cuarto Básico. Santiago
- [3] AGENCIA DE CALIDAD DE LA EDUCACIÓN. (2013b). Apunte N.º 6 Relación TIMSS-SIMCE: Calidad psicométrica e invarianza de puntajes y parámetros. Santiago.
- [4] DORANS, N.J. (2004). Equating, concordance and expectation. En *Applied Psychological Measurement*, 28, 227-246.
- [5] EBEL, R. (1972). *Essentials of educational measurement*. Prentice Hall 622 páginas.
- [6] EFRON, B. (1980). *The Jackknife, The Bootstrap and Other Resampling Plans, Division of Biostatistics*, Stanford University, California. Recuperado de: <http://statistics.stanford.edu/~ckirby/techreports/BIO/BIO%2063.pdf>
- [7] GAGGERO, M. A. (2012). Análisis psicométrico en la comparación de test curriculares en el ámbito nacional e internacional. [Informe de Tesis Ingeniería Estadística, Departamento de Matemática y Ciencia de la Computación, Facultad de Ciencia, Universidad de Santiago de Chile.
- [8] GHISELLI, E. Y ZEDECK, S. (1981). *Measurement Theory for the Behavioral Sciences*, Freeman and Company 494 páginas.
- [9] GEMPP, R. (2010). Equiparación, Alineamiento y Predicción de Puntuaciones en Medición Educativa. En *Revista Iberoamericana de Evaluación Educativa* 2010 Volumen 3, N.º 2, 103-126.
- [10] GUZMÁN, V. (2013). Rendimiento Educativo Chileno Bajo Criterios Internacionales: Alineación SIMCE-TIMSS4º Básico Matemática. [Informe de Tesis] Ingeniería Estadística, Departamento de Matemática y Ciencia de la Computación, Facultad de Ciencia, Universidad de Santiago de Chile.
- [11] HAMBLETON, R. Y SWAMINATHAN H. (1985). *Item Response Theory: Principles and Applications* Kluwer-Nijhoff Publishing Boston/Dordrecht/Lancaster.
- [12] KAISER, H Y CAFFRY, J., (1965). Alpha factor Analysis. *Psychometrika*. 30: 1-14
- [13] OECD (2012). *PISA 2009 Technical Report*. PISA. OECD Publishing Paris.

- [14] OLSON J., MARTIN M. ANDMULLIS I. (2009). *TIMSS 2007 Technical Report*. TIMSS & PIRLS International Study Center Boston.
- [15] SAN MARTÍN, E. Y MANZI, J.(2003). *La necesaria complementariedad entre teoría clásica de la medición (TCT) y la Teoría de Respuesta al Ítem (IRT), aspectos conceptuales y aplicaciones*. Santiago: Estudios Públicos.
- [16] THOMPSON, B. (EDITOR).(2003). *Score Reliability*, Sage Publications.
- [17] Las bases de datos se obtienen de:
<http://www.agenciaeducacion.cl/simce/bases-de-datos-nacionales/>

6. Anexos

6.1. Detalle de Análisis por cuadernillo

6.1.1. Simce - Forma C

Esta forma, al igual que todas las relacionadas a la prueba SIMCE 2010, posee un total de 35 ítems y fue presentada a 1.419 estudiantes que también fueron sometidos a la prueba TIMSS; además, posee un coeficiente de confiabilidad de 0,90 lo cual es bastante adecuado dado el número moderado de preguntas.

Tabla 17: Simce - Forma C

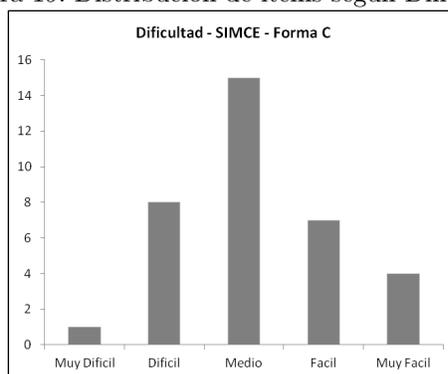
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
C	35	1.419	0,896

En cuanto a los parámetros clásicos, la dificultad de los ítems es de carácter transversal, puesto que considera ítems desde “Muy Fácil” a “Muy Difícil”, concentrando la mayor frecuencia en el nivel “Medio” de esta característica.

Tabla 18: Dificultad Simce - Forma C

Grado de Dificultad	N.º de ítems	%
Muy Difícil	D<0,3	1 3 %
Difícil	0,3<D<0,5	8 23 %
Medio	0,5<D<0,7	15 43 %
Fácil	0,7<D<0,8	7 20 %
Muy Fácil	D>0,8	4 11 %
Total		35 100 %

Figura 19: Distribución de ítems según Dificultad



En cuanto al parámetro de discriminación, mientras mayor sea este, mejor será la calidad del ítem. Todos los ítems poseen un valor superior al 0,35, lo que puede interpretarse que cada ítem influye de manera positiva en el resultado final o total del estudiante sometido a este test.

Tabla 19: Discriminación Simce - Forma C

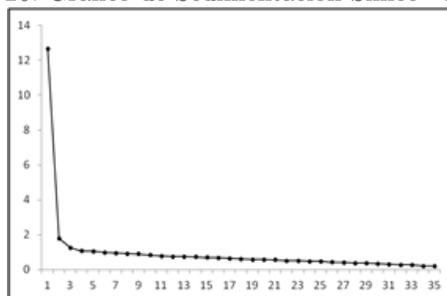
Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	0	0 %
Regular $0,15 < r < 0,25$	0	0 %
Buena $0,25 < r < 0,35$	5	14 %
Muy Buena $r > 0,35$	30	86 %
Total	35	100 %

En cuanto al análisis factorial, se analiza para corroborar la unidimensionalidad del constructo que miden los ítems. El análisis indica que, utilizando un total de 5 componentes, se explica casi un 51 % de la variabilidad total. Además, la primera componente por sí sola explica el 36 %. Este es un buen indicador para demostrar que la gran mayoría de los ítems está enfocado en analizar el mismo constructo.

Tabla 20: Varianza Explicada Simce - Forma C

Componente	Autovalor	% de la Varianza	% Acumulado
1	12,66	36,16	36,16
2	1,8	5,15	41,31
3	1,26	3,6	44,91
4	1,07	3,06	47,97
5	1,05	3	50,97

Figura 20: Gráfico de Sedimentación Simce - Forma C



En cuanto a la cantidad de ítems que está relacionado con cada componente³⁸, se desprende que el 91 % de ellos presentan un valor máximo relacionado con la primera componente, mientras que solo un 3 % presenta la relación más alejada.

³⁸Esta influencia se calcula mediante el análisis de las cargas factoriales de cada pregunta.

Tabla 21: Simce - Forma C

Factor	N.º de ítems	% de ítems
1	32	91 %
2	1	3 %
3	1	3 %
4	0	0 %
5	1	3 %

Además, este análisis factorial posee un KMO igual al 0,95, lo que hace válida y útil la interpretación del análisis factorial.

Tabla 22: Simce - Forma C

KMO	0,954
-----	-------

6.1.2. Simce - Forma D

Este cuestionario, al igual que todos los relacionados a la prueba Simce 2010, posee un total de 35 ítems y fue presentado a 1.383 estudiantes que también fueron sometidos a la prueba TIMSS, además posee un coeficiente de confiabilidad cercano al 0,92.

Tabla 23: Simce - Forma D

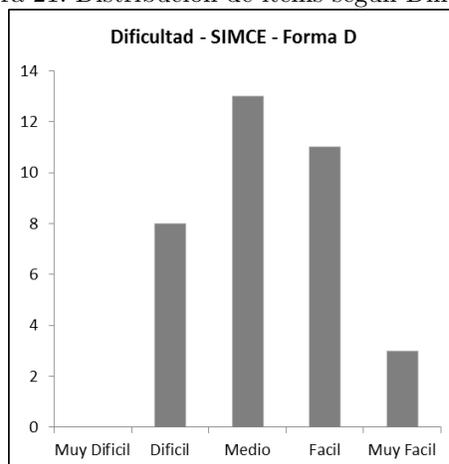
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
D	35	1.383	0,917

En cuanto a los parámetros clásicos, la dificultad de los ítems, considera ítems desde “Muy Fácil” a “Difícil”, concentrando la mayor frecuencia en el nivel “Medio” de esta característica.

Tabla 24: Dificultad Simce - Forma D

Grado de Dificultad	N.º de ítems	%
Muy Difícil	$D < 0,3$	0 0 %
Difícil	$0,3 < D < 0,5$	8 23 %
Medio	$0,5 < D < 0,7$	13 37 %
Fácil	$0,7 < D < 0,8$	11 31 %
Muy Fácil	$D > 0,8$	3 9 %
Total	35	100 %

Figura 21: Distribución de ítems según Dificultad



En cuanto al parámetro de discriminación, todos los ítems poseen un valor superior al 0,25 (exceptuando uno), lo que puede interpretarse que cada ítem influye de manera positiva en el resultado final o total del estudiante.

sometido a este test.

Tabla 25: Discriminación Simce - Forma D

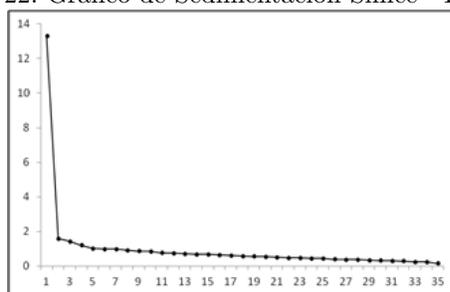
Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	0	0 %
Regular $0,15 < r < 0,25$	0	0 %
Buena $0,25 < r < 0,35$	1	3 %
Muy Buena $r > 0,35$	34	97 %
Total	35	100 %

El análisis factorial indica que utilizando un total de 5 componentes se explica casi un 53 % de la variabilidad total, además, la primera componente por sí sola explica el 38 %, buen indicador para demostrar que la gran mayoría de los ítems está enfocada en analizar el mismo constructo.

Tabla 26: Varianza Explicada Simce - Forma D

Componente	Autovalor	% de la Varianza	% Acumulado
1	13,31	38,02	38,02
2	1,59	4,55	42,56
3	1,42	4,05	46,61
4	1,20	3,43	50,04
5	1,01	2,88	52,93

Figura 22: Gráfico de Sedimentación Simce - Forma D



En cuanto a la cantidad de ítems que está relacionado con cada componente, se desprende que el 94 % de ellos presenta un valor máximo relacionado con la primera componente, mientras que la segunda, solo un 3 %. Además, este análisis factorial posee un KMO igual al 0,96, lo que hace válida y útil la interpretación del análisis factorial.

Tabla 27: Simce - Forma D

Factor	N.º de ítems	% de ítems
1	33	94%
2	1	3%
3	0	0%
4	1	3%
5	0	0%

Tabla 28: Simce - Forma D

KMO	0,964
-----	-------

6.1.3. Simce - Forma E

Al igual que todos los cuestionarios Simce, este posee 35 ítems y lo contestaron 1.354 alumnos que además contestaron la prueba TIMSS, en cuanto al coeficiente de confiabilidad, este es similar al obtenido en la “Forma D”, con un indicador equivalente a 0,92.

Tabla 29: Simce - Forma E

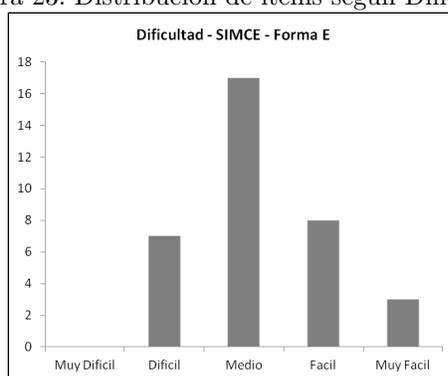
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
E	35	1.354	0,927

Los parámetros de dificultad de esta forma no presentan indicadores que puedan ser interpretados como “Muy Difíciles”, sin embargo, no se observa una pronunciada tendencia hacia las preguntas de menor dificultad.

Tabla 30: Dificultad Simce - Forma E

Grado de Dificultad	N.º de ítems	%
Muy Difícil	D < 0,3	0 0 %
Difícil	0,3 < D < 0,5	7 20 %
Medio	0,5 < D < 0,7	17 49 %
Fácil	0,7 < D < 0,8	8 23 %
Muy Fácil	D > 0,8	3 9 %
Total	35	100 %

Figura 23: Distribución de ítems según Dificultad



La discriminación es el punto a destacar en este formato de presentar la prueba Simce, puesto que los 35 ítems (el total) están por sobre el 0,35, lo cual los hace considerar como muy buenos, por lo tanto se puede asumir que todas las preguntas de esta forma están vinculadas fuertemente con el puntaje total obtenido por los estudiantes

sometidos a esta.

Tabla 31: Discriminación Simce - Forma E

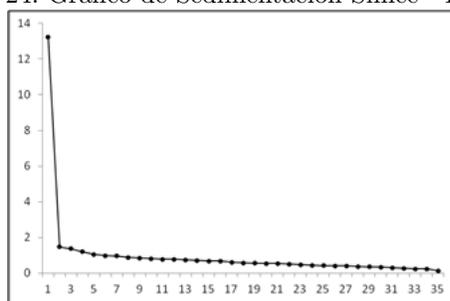
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	0	0 %
Regular	$0,15 < r < 0,25$	0	0 %
Buena	$0,25 < r < 0,35$	0	0 %
Muy Buena	$r > 0,35$	35	100 %
Total		35	100 %

El respectivo análisis factorial queda determinado por 5 componentes, los que explican un 52 % de la variabilidad total, donde solo el primer componente explica poco menos del 37 %, lo que indica, de nuevo, que efectivamente existe un fuerte indicio para asumir la unidimensionalidad de las preguntas de la forma.

Tabla 32: Varianza Explicada Simce - Forma E

Componente	Autovalor	% de la Varianza	% Acumulado
1	13,24	37,82	37,82
2	1,48	4,22	42,04
3	1,38	3,93	45,97
4	1,20	3,43	49,40
5	1,05	3,01	52,42

Figura 24: Gráfico de Sedimentación Simce - Forma E



La unidimensionalidad queda confirmada luego de corroborar que el 97 % de los 35 ítems a evaluar expresan una mayor vínculo al primer componente, mientras que solo dos preguntas se apartan de esta conclusión.

En cuanto al índice de Kaiser-Meyer-Olkin, este es equivalente al 0,98, fuerte indicador para concluir que efectivamente cobra sentido y utilidad el realizar esta reducción de variables.

Tabla 33: Simce - Forma E

Factor	N.º de ítems	% de ítems
1	34	97%
2	0	0%
3	0	0%
4	0	0%
5	1	3%

Tabla 34: Simce - Forma E

KMO	0,981
-----	-------

6.1.4. Simce - Forma G

Este formato de prueba Simce, fue contestado por 1.295 alumnos que también fueron sometidos a la prueba TIMSS, consta de 35 ítems y posee el mayor coeficiente de confiabilidad, instalándose con un cifra equivalente al 0,95, indicador de altísima categoría, lo que en la práctica no deja duda alguna que el instrumento evaluador, en repetidas ocasiones, tenderá a conseguir los mismos resultados.

Tabla 35: Simce - Forma G

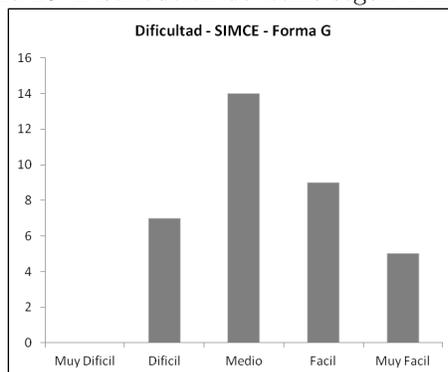
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
G	35	1.295	0,945

Los parámetros de dificultad de esta forma no presentan indicadores que puedan ser interpretados como “Muy Dificiles”, sin embargo no se observa una pronunciada tendencia hacia las preguntas de menor dificultad.

Tabla 36: Dificultad Simce - Forma G

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	0 0%
Díficil	0,3<D<0,5	7 20%
Medio	0,5<D<0,7	14 40%
Fácil	0,7<D<0,8	9 26%
Muy Fácil	D>0,8	5 14%
Total	35	100%

Figura 25: Distribución de ítems según Dificultad



En cuanto a la discriminación del test, este es de alto nivel, puesto que solo un ítem posee un nivel discriminativo inferior al 0,35, sin embargo, este mismo es mayor al 0,25, lo que continúa siendo de buen nivel. A continuación se aprecia en detalle:

Al aplicar el análisis factorial, que está formado con 4 componentes, explica un 50 % de la variabilidad total,

Tabla 37: Discriminación Simce - Forma G

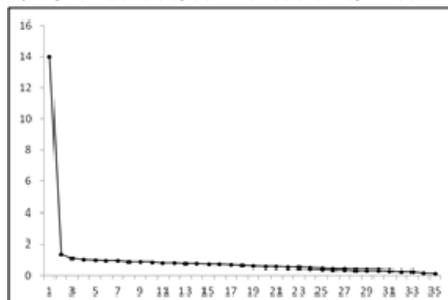
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	0	0 %
Regular	$0,15 < r < 0,25$	0	0 %
Buena	$0,25 < r < 0,35$	1	3 %
Muy Buena	$r > 0,35$	34	97 %
	Total	35	100 %

desde el cual solo el primero considera el 40 % del total, lo que ya es un indicio potente de que efectivamente existe unidimensionalidad del test.

Tabla 38: Varianza Explicada Simce - Forma G

Componente	Autovalor	% de la Varianza	% Acumulado
1	14,02	40,07	40,07
2	1,34	3,83	43,9
3	1,09	3,11	47,01
4	1,01	2,87	49,88

Figura 26: Gráfico de Sedimentación Simce - Forma G



Lo comentado queda completamente corroborado al visualizar que todos los ítems están orientados a solo un componente o factor de este análisis.

Además, el KMO es de altísimo nivel (valor equivalente al 0,98) considerando que es aceptado que si se presenta un valor de este indicador superior al 0,7, ya es aceptable y coherente considerar este análisis vinculado a la estadística multivariada.

Tabla 39: Simce - Forma G

Factor	N.º de ítems	% de ítems
1	35	100 %
2	0	0 %
3	0	0 %
4	0	0 %

Tabla 40: Simce - Forma G

KMO	0,981
-----	-------

6.1.5. TIMSS - Forma 1

Tal como se comentó anteriormente, la prueba TIMSS, presenta cuestionarios de diferentes tamaños, al menos este consta de 22 ítems, que fueron respondidos por 399 alumnos, además este formato de prueba presenta un valor del coeficiente del alfa de Cronbach igual a 0,84, lo que en la práctica es inferior a las formas del Simce. Al ser un instrumento medidor internacional y no estar particularmente adaptado necesariamente a las características naturales de nuestro país, es adecuado, dado el bajo número de preguntas.

Tabla 41: TIMSS- Forma 1

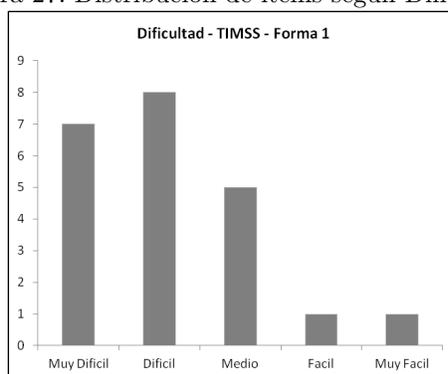
Forma	N.º de Ítems	N.º Contesta	Alfa de Cronbach
1	22	399	0,836

Los parámetros de clásicos de dificultad en este formato tienden a estar desplazados hacia aquellos de alto nivel, puesto que el 68 % posee un valor menor al 0,5.

Tabla 42: Dificultad TIMSS- Forma 1

Grado de Dificultad	N.º de ítems	%
Muy Díficil	$D < 0,3$	7 32 %
Díficil	$0,3 < D < 0,5$	8 36 %
Medio	$0,5 < D < 0,7$	5 23 %
Fácil	$0,7 < D < 0,8$	1 5 %
Muy Fácil	$D > 0,8$	1 5 %
Total	22	100 %

Figura 27: Distribución de ítems según Dificultad



La discriminación del test, es de buen nivel, sin embargo, existe un ítem con valor discriminativo inferior al 0,15. Esto vuelve cuestionable si efectivamente ese ítem está vinculado con la estimación de la habilidad en esta

manera de presentar la prueba.

Tabla 43: Discriminación TIMSS- Forma 1

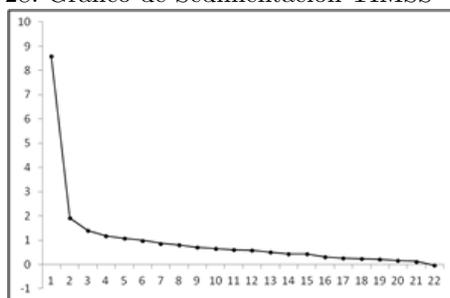
Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	1	5 %
Regular $0,15 < r < 0,25$	2	9 %
Buena $0,25 < r < 0,35$	4	18 %
Muy Buena $r > 0,35$	15	68 %
Total	22	100 %

El análisis factorial de esta prueba arroja 5 factores de influencia, los cuales en conjunto explican el 64 % de la variabilidad total, donde casi el 39 % es manifestado solo por la primera componente.

Tabla 44: Varianza Explicada TIMSS- Forma 1

Componente	Autovalor	% de la Varianza	% Acumulado
1	8,58	39,02	39,02
2	1,59	8,66	47,68
3	1,25	6,40	54,08
4	1,11	5,37	59,45
5	1,04	4,85	64,30

Figura 28: Gráfico de Sedimentación TIMSS- Forma 1



De las cargas factoriales de los ítems, el 86 % están vinculadas al primer factor, lo que en frecuencia absoluta son 19 preguntas, mientras que las cuatro restantes están repartidas en los 5 componentes restantes.

El índice KMO arroja un valor superior al 0,7, lo que hace que este análisis cobre sentido y utilidad realizarlo.

Tabla 45: TIMSS- Forma 1

Factor	N.º de ítems	% de ítems
1	19	86 %
2	0	0 %
3	2	9 %
4	1	5 %
5	0	0 %
6	0	0 %

Tabla 46: TIMSS- Forma 1

KMO	0,890
-----	-------

6.1.6. TIMSS - Forma 2

La segunda manera de presentar esta prueba consta de 26 preguntas, las cuales fueron respondidas por 396 alumnos, lo que arroja un coeficiente del alfa de Cronbach igual al 0,85, coeficiente de buen nivel, el cual permite concluir que en esta prueba, en la medida que su repetición se ejecute, los resultados serán similares.

Tabla 47: TIMSS- Forma 2

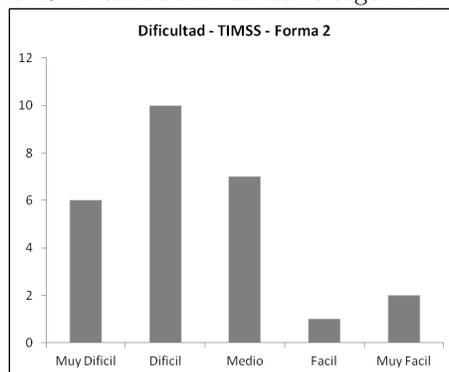
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
2	26	396	0,845

Los parámetros de dificultad de esta forma, presentan valores en todos los niveles, sin embargo, aquellos de categoría “Fáciles” no son predominantes y se hace evidente además que el 40 % de los ítems son categorizados en “Difícil”:

Tabla 48: Dificultad TIMSS- Forma 2

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	6 23 %
Díficil	0,3<D<0,5	10 38 %
Medio	0,5<D<0,7	7 27 %
Fácil	0,7<D<0,8	1 4 %
Muy Fácil	D>0,8	2 8 %
Total	26	100 %

Figura 29: Distribución de ítems según Dificultad



La discriminación en esta forma es de buen nivel, sin embargo, solo uno ítem posee un indicador inferior a 0,15, mientras que el 73 % posee valores superiores al 0,35.

Tabla 49: Discriminación TIMSS- Forma 2

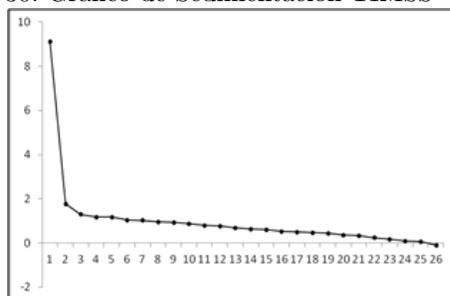
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	1	4 %
Regular	$0,15 < r < 0,25$	2	8 %
Buena	$0,25 < r < 0,35$	4	15 %
Muy Buena	$r > 0,35$	19	73 %
	Total	26	100 %

El análisis factorial arroja 7 factores capaces de explicar el 63 % de la variabilidad total del conjunto de todos los ítems, además, solo la primera componente por sí sola explica el 35 % de la variabilidad.

Tabla 50: Varianza Explicada TIMSS- Forma 2

Componente	Autovalor	% de la Varianza	% Acumulado
1	9,11	35,05	35,05
2	1,78	6,86	41,91
3	1,30	4,99	46,90
4	1,19	4,57	51,46
5	1,17	4,49	55,95
6	1,04	4,02	59,97
7	1,03	3,95	63,92

Figura 30: Gráfico de Sedimentación TIMSS- Forma 2



Se aprecia unidimensionalidad de los ítems, donde el 85 % de ellos presenta una máxima carga factorial asociada a la primera componente, que explica la mayor variabilidad del test.

Finalmente, el índice KMO arroja un valor del 0,88, índice de alto valor, lo que en definitiva indica que es coherente aplicar este análisis factorial y que efectivamente las cargas de correlación parcial son de bajo valor.

Tabla 51: TIMSS- Forma 2

Factor	N.º de ítems	% de ítems
1	22	85 %
2	1	4 %
3	1	4 %
4	0	0 %
5	1	4 %
6	0	0 %
7	1	4 %

Tabla 52: TIMSS- Forma 2

KMO	0,883
-----	-------

6.1.7. TIMSS - Forma 3

Este cuestionario de la prueba TIMSS posee un total de 26 ítems y fue presentado a 390 estudiantes, además posee un coeficiente de confiabilidad del 0,81, valor más que suficiente para aceptar su eficacia.

Tabla 53: TIMSS- Forma 3

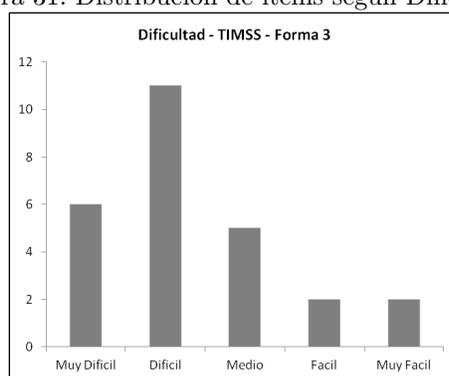
Forma	N.º de Ítems	N.º Contesta	Alfa de Cronbach
3	26	390	0,811

En cuanto a los parámetros clásicos, la dificultad de los ítems es de carácter transversal, puesto que considera ítems desde “Muy Fácil” a “Muy Difícil”, concentrando la mayor frecuencia en el nivel “Difícil” de esta característica. A continuación se presentan los indicadores:

Tabla 54: Dificultad TIMSS- Forma 3

Grado de Dificultad	N.º de ítems	%
Muy Difícil	D<0,3	6 23 %
Difícil	0,3<D<0,5	11 42 %
Medio	0,5<D<0,7	5 19 %
Fácil	0,7<D<0,8	2 8 %
Muy Fácil	D>0,8	2 8 %
Total		26 100 %

Figura 31: Distribución de ítems según Dificultad



En cuanto al parámetro de discriminación, mientras mayor sea este, mejor será el ítem al que está relacionado el valor. A continuación se especifican los valores en esta forma:

Tabla 55: Discriminación TIMSS- Forma 3

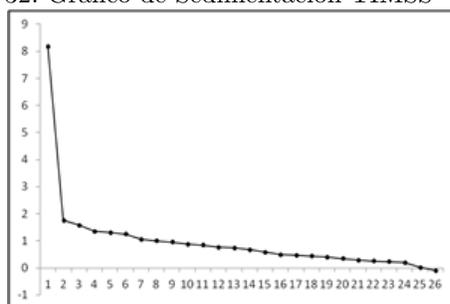
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	2	8 %
Regular	$0,15 < r < 0,25$	3	12 %
Buena	$0,25 < r < 0,35$	3	12 %
Muy Buena	$r > 0,35$	18	69 %
	Total	26	100 %

En cuanto al análisis factorial, este se analiza, para corroborar la unidimensionalidad de los ítems El análisis factorial indica que utilizando un total de 8 componentes se explica un 67 % de la variabilidad total, además, la primera componente por sí sola explica el 31,48, buen indicador para demostrar que todos los ítems (o la gran mayoría) están enfocados en analizar el mismo constructo.

Tabla 56: Varianza Explicada TIMSS- Forma 3

Componente	Autovalor	% de la Varianza	% Acumulado
1	8,18	31,48	31,48
2	1,77	6,80	38,28
3	1,58	6,06	44,35
4	1,35	5,21	49,55
5	1,31	5,02	54,58
6	1,25	4,81	59,38
7	1,06	4,07	63,46
8	1,00	3,85	63,30

Figura 32: Gráfico de Sedimentación TIMSS- Forma 3



En cuanto a la cantidad de ítems que está relacionado con cada componente, se desprende que el 81 % de ellos presenta un valor predominante relacionado con la primera componente, mientras que solo un 4 % están relacionados con las siguientes.

Tabla 57: TIMSS- Forma 3

Factor	N.º de ítems	% de ítems
1	21	81 %
2	0	0 %
3	0	0 %
4	1	4 %
5	1	4 %
6	1	4 %
7	2	8 %
8	0	0 %

Además, este análisis factorial posee un KMO igual al 0,85, lo que hace válida y útil la interpretación relacionada a este.

Tabla 58: TIMSS- Forma 3

KMO	0,846
-----	-------

6.1.8. TIMSS - Forma 4

Como todos los cuestionarios de TIMSS, la cantidad de ítems que poseen varían según el cuestionario, este presenta 24 ítems y lo contestaron 396 estudiantes que además respondieron la prueba Simce. En cuanto al coeficiente de confiabilidad, este es levemente superior al anterior, con un indicador equivalente a 0,83.

Tabla 59: TIMSS- Forma 4

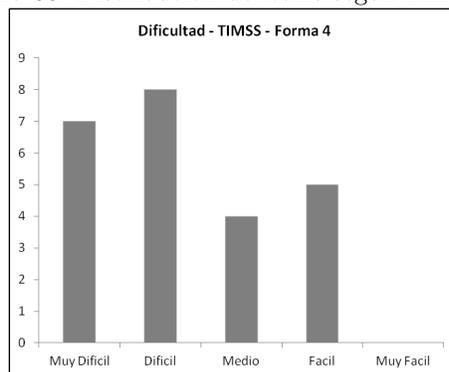
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
4	24	396	0,830

En relación con los parámetros de dificultad, esta forma no presenta indicadores que puedan ser interpretados como “Muy Fáciles”, además se observa una pronunciada tendencia hacia las preguntas de mayor dificultad, valores que están presentados a continuación:

Tabla 60: Dificultad TIMSS- Forma 4

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	7 29 %
Díficil	0,3<D<0,5	8 33 %
Medio	0,5<D<0,7	4 17 %
Fácil	0,7<D<0,8	5 21 %
Muy Fácil	D>0,8	0 0 %
Total	24	100 %

Figura 33: Distribución de ítems según Dificultad



En cuanto a la discriminación del cuestionario, este, al igual que el anterior, presenta ítems cuyos indicadores son de baja categoría, pero solamente dos. A continuación se especifican las proporciones comentadas:

Tabla 61: Discriminación TIMSS- Forma 4

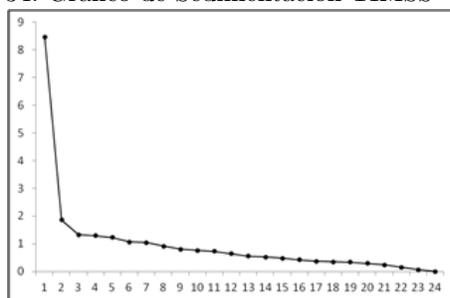
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	2	8 %
Regular	$0,15 < r < 0,25$	4	17 %
Buena	$0,25 < r < 0,35$	2	8 %
Muy Buena	$r > 0,35$	16	67 %
	Total	26	100 %

Al aplicar el análisis factorial, formado con 7 componentes, explica un 68% de la variabilidad total, desde el cual solo el primero considera el 35% del total, lo que ya es un buen indicio de que efectivamente existe unidimensionalidad del test.

Tabla 62: Varianza Explicada TIMSS- Forma 4

Componente	Autovalor	% de la Varianza	% Acumulado
1	8,47	35,30	35,30
2	1,86	7,77	43,06
3	1,33	5,54	48,61
4	1,30	5,41	54,02
5	1,23	5,14	59,15
6	1,07	4,46	63,62
7	1,05	4,39	68,01

Figura 34: Gráfico de Sedimentación TIMSS- Forma 4



Anteriormente se intuía la corroboración de unidimensionalidad, sin embargo, esto se confirma al verificar que del total de los ítems del análisis factorial, el 83% muestra una mayor inclinación hacia el primer factor.

Finalmente, el índice KMO es ampliamente superior al 0,7, por lo tanto se puede concluir que esta reducción de variables explicativas realizadas a través de la herramienta multivariada del análisis factorial, sí tiene sentido y utilidad en su aplicación. A continuación se presenta dicho valor:

Tabla 63: TIMSS- Forma 4

Factor	N.º de ítems	% de ítems
1	20	83 %
2	1	4 %
3	0	0 %
4	2	8 %
5	0	0 %
6	0	0 %
7	1	4 %

Tabla 64: TIMSS- Forma 4

KMO	0,887
-----	-------

6.1.9. TIMSS - Forma 5

Este cuestionario de la prueba TIMSS posee un total de 26 ítems y fue presentado a 402 estudiantes, además posee un coeficiente de confiabilidad levemente superior, cercano al 0,86, valor más que aceptable.

Tabla 65: TIMSS- Forma 5

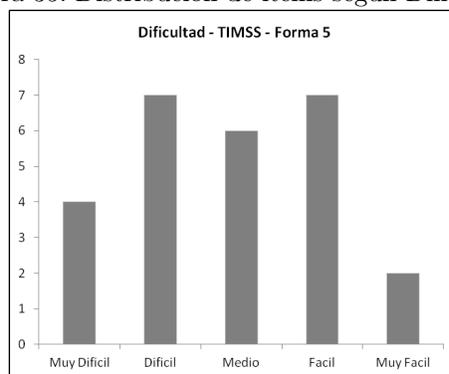
Forma	N.º de Ítems	N.º Contesta	Alfa de Cronbach
5	26	402	0,860

Los parámetros de dificultad de esta forma, presentan indicadores que pueden ser interpretados desde “Muy Difícil” a “Muy Fácil”.

Tabla 66: Dificultad TIMSS- Forma 5

Grado de Dificultad	N.º de ítems	%
Muy Difícil	$D < 0,3$	4 15 %
Difícil	$0,3 < D < 0,5$	7 27 %
Medio	$0,5 < D < 0,7$	6 23 %
Fácil	$0,7 < D < 0,8$	7 27 %
Muy Fácil	$D > 0,8$	2 8 %
Total	26	100 %

Figura 35: Distribución de ítems según Dificultad



En cuanto a la discriminación del cuestionario, al igual que el anterior, este presenta ítems cuyos indicadores son bajos.

El respectivo análisis factorial queda determinado por 8 componentes, los que explican un 72 % de la variabilidad total, donde solo el primer componente explica el 38 %, lo que indica de nuevo que efectivamente existe un

Tabla 67: Discriminación TIMSS- Forma 5

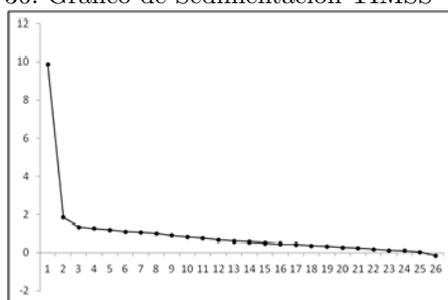
Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	2	8 %
Regular $0,15 < r < 0,25$	2	8 %
Buena $0,25 < r < 0,35$	0	0 %
Muy Buena $r > 0,35$	22	85 %
Total	26	100 %

fuerte indicio para asumir desde ya la unidimensionalidad de los test.

Tabla 68: Varianza Explicada TIMSS- Forma 5

Componente	Autovalor	% de la Varianza	% Acumulado
1	9,87	37,96	37,96
2	1,85	7,11	45,06
3	1,36	5,23	50,30
4	1,28	4,93	55,22
5	1,20	4,63	59,85
6	1,12	4,29	64,14
7	1,08	4,15	68,29
8	1,01	3,90	72,19

Figura 36: Gráfico de Sedimentación TIMSS- Forma 5



La unidimensionalidad queda confirmada luego de corroborar que el 81 % de los 26 ítems a evaluar expresa un mayor vínculo al primer componente, mientras que solo 2 preguntas se apartan de esta conclusión.

En cuanto al índice de Kaiser-Meyer-Olkin, este es equivalente al 0,89, fuerte indicador para concluir que efectivamente cobra sentido y utilidad el realizar esta reducción de variables.

Tabla 69: TIMSS- Forma 5

Factor	N.º de ítems	% de ítems
1	21	81 %
2	0	0 %
3	0	0 %
4	1	4 %
5	1	4 %
6	1	4 %
7	1	4 %
8	1	4 %

Tabla 70: TIMSS- Forma 5

KMO	0,885
-----	-------

6.1.10. TIMSS - Forma 6

Este formato de prueba TIMSS, fue contestado por 405 alumnos que también fueron sometidos a la prueba Simce, consta de 27 ítems y posee el mayor coeficiente de confiabilidad, instalándose con una cifra equivalente al 0,86, indicador de alta categoría, lo que en la práctica significa que el instrumento evaluador, en repetidas ocasiones, tenderá a conseguir los mismos resultados.

Tabla 71: TIMSS- Forma 6

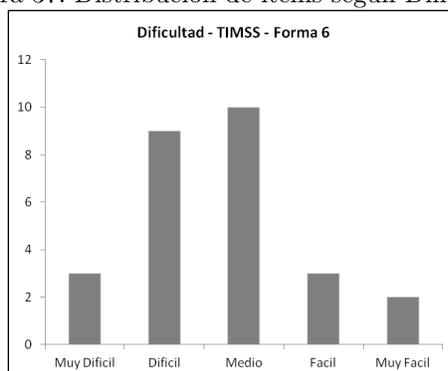
Forma	N.º de Ítems	N.º Contesta	Alfa de Cronbach
6	27	405	0,860

Los parámetros de dificultad de esta forma presentan indicadores que pueden ser interpretados desde “Muy Difícil” a “Muy Fácil”, sin embargo, no se observa una pronunciada tendencia.

Tabla 72: Dificultad TIMSS- Forma 6

Grado de Dificultad	N.º de ítems	%
Muy Difícil $D < 0,3$	3	11 %
Difícil $0,3 < D < 0,5$	9	33 %
Medio $0,5 < D < 0,7$	10	37 %
Fácil $0,7 < D < 0,8$	3	11 %
Muy Fácil $D > 0,8$	2	7 %
Total	27	100 %

Figura 37: Distribución de ítems según Dificultad



La discriminación del cuestionario, al igual que el anterior, presenta 2 ítems cuyos indicadores son bajos.

Tabla 73: Discriminación TIMSS- Forma 6

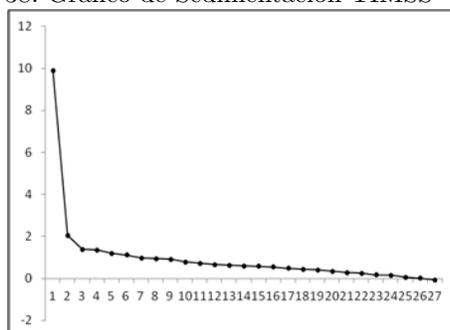
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	2	7 %
Regular	$0,15 < r < 0,25$	1	4 %
Buena	$0,25 < r < 0,35$	1	4 %
Muy Buena	$r > 0,35$	23	85 %
	Total	27	100 %

Al aplicar el análisis factorial, formado con 7 componentes, explica un 63 % de la variabilidad total, desde el cual solo el primero considera el 36 % del total, lo que ya es un indicio potente de que efectivamente existe unidimensionalidad del test.

Tabla 74: Varianza Explicada TIMSS- Forma 6

Componente	Autovalor	% de la Varianza	% Acumulado
1	9,90	36,65	36,65
2	2,04	7,57	44,22
3	1,40	5,19	49,41
4	1,36	5,05	54,46
5	1,21	4,47	58,93
6	1,12	4,15	63,09

Figura 38: Gráfico de Sedimentación TIMSS- Forma 6



Lo comentado queda completamente corroborado, el 81 % de los ítems están orientados a solo un componente o factor de este análisis.

Además, el KMO es de altísimo nivel (valor equivalente al 0,9).

Tabla 75: TIMSS- Forma 6

Factor	N.º de ítems	% de ítems
1	22	81 %
2	2	7 %
3	0	0 %
4	1	4 %
5	0	0 %
6	2	7 %

Tabla 76: TIMSS- Forma 6

KMO	0,895
-----	-------

6.1.11. TIMSS - Forma 7

Tal como se dijo anteriormente, la prueba TIMSS, presenta cuestionarios de diferentes tamaños, al menos este consta de 25 ítems, fue respondido por 399 alumnos, además este formato de prueba presenta un valor del coeficiente del alfa de Cronbach igual a 0,82, lo que en la práctica es inferior a los de Simce, pero que al ser TIMSS un instrumento medidor internacional (no adaptado necesariamente a las características naturales de nuestro país), es de buen nivel.

Tabla 77: TIMSS- Forma 7

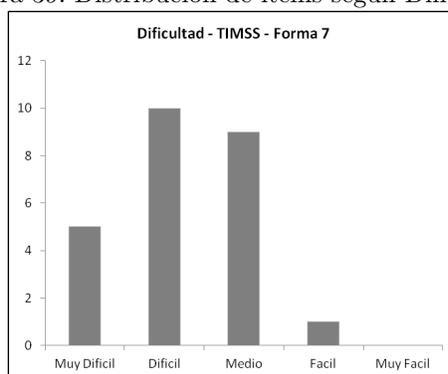
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
7	25	399	0,821

Los parámetros de clásicos de dificultad en este formato tienden a estar desplazados hacia aquellos de alto nivel, puesto que el 60% posee un valor menor al 0,5.

Tabla 78: Dificultad TIMSS- Forma 7

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	5 20%
Díficil	0,3<D<0,5	10 40%
Medio	0,5<D<0,7	9 36%
Fácil	0,7<D<0,8	1 4%
Muy Fácil	D>0,8	0 0%
Total	25	100%

Figura 39: Distribución de ítems según Dificultad



La discriminación del cuestionario, al igual que el anterior, tiene ítems con bajo valor. A continuación se especifican las proporciones comentadas:

Tabla 79: Discriminación TIMSS- Forma 7

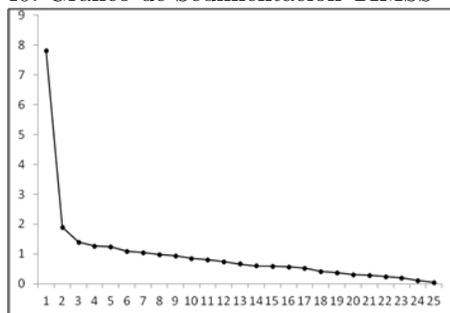
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	2	8 %
Regular	$0,15 < r < 0,25$	3	12 %
Buena	$0,25 < r < 0,35$	4	16 %
Muy Buena	$r > 0,35$	16	64 %
	Total	25	100 %

El análisis factorial de esta prueba arroja 7 factores de influencia, los cuales en conjunto explican el 63 % de la variabilidad total, donde casi el 32 % es manifestado solo por la primera componente.

Tabla 80: Varianza Explicada TIMSS- Forma 7

Componente	Autovalor	% de la Varianza	% Acumulado
1	7,82	31,29	31,29
2	1,90	7,60	38,89
3	1,39	5,56	44,45
4	1,27	5,06	49,52
5	1,24	4,96	54,48
6	1,08	4,33	58,82
7	1,05	4,18	63,00

Figura 40: Gráfico de Sedimentación TIMSS- Forma 7



De las cargas factoriales de los ítems, el 76 % están vinculadas al primer factor, lo que en frecuencia absoluta corresponde a 19 preguntas, mientras que las 4 restantes están repartidas en los demás 6 componentes.

El índice KMO arroja un valor superior al 0,7, lo que hace que este análisis cobre sentido y tenga utilidad realizarlo.

Tabla 81: TIMSS- Forma 7

Factor	N.º de ítems	% de ítems
1	19	76 %
2	1	4 %
3	1	4 %
4	2	8 %
5	0	0 %
6	1	4 %
7	1	4 %

Tabla 82: TIMSS- Forma 7

KMO 0,872

6.1.12. TIMSS - Forma 8

La segunda manera de presentar esta prueba consta de 26 preguntas, las cuales fueron respondidas por 399 alumnos, lo que arroja un coeficiente del alfa de Cronbach igual al 0,80, coeficiente de buen nivel, el cual permite concluir que en esta prueba, en la medida que su repetición se ejecute, los resultados serán similares.

Tabla 83: TIMSS- Forma 8

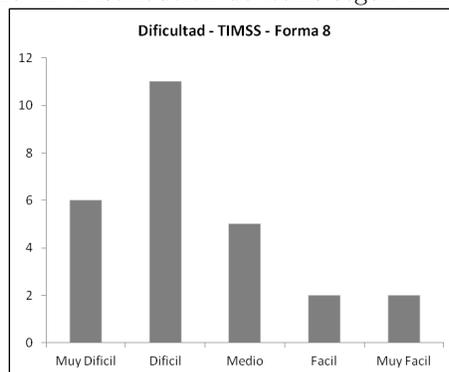
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
8	26	399	0,802

Los parámetros de dificultad de esta forma, presentan valores de todos los niveles, sin embargo, aquellos de categoría fáciles no son predominantes y se hace evidente además que el 40 % de los ítems son categorizados en “Difícil”.

Tabla 84: Dificultad TIMSS- Forma 8

Grado de Dificultad	N.º de ítems	%
Muy Difícil $D < 0,3$	6	23 %
Difícil $0,3 < D < 0,5$	11	42 %
Medio $0,5 < D < 0,7$	5	19 %
Fácil $0,7 < D < 0,8$	2	8 %
Muy Fácil $D > 0,8$	2	8 %
Total	26	100 %

Figura 41: Distribución de ítems según Dificultad



La discriminación en esta forma es de buen nivel, sin embargo solo dos ítems poseen un indicador inferior a 0,15, mientras que el 80 % posee valores superiores al 0.35.

Tabla 85: Discriminación TIMSS- Forma 8

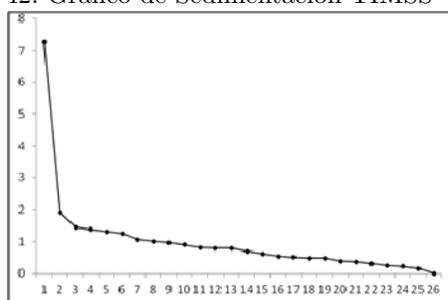
Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	1	4 %
Regular $0,15 < r < 0,25$	2	8 %
Buena $0,25 < r < 0,35$	11	42 %
Muy Buena $r > 0,35$	12	46 %
Total	26	100 %

El análisis factorial arroja 8 factores capaces de explicar el 64 % de la variabilidad total del conjunto de todos los ítems, además, solo la primera componente por sí sola explica el 27 % de la variabilidad.

Tabla 86: Varianza Explicada TIMSS- Forma 8

Componente	Autovalor	% de la Varianza	% Acumulado
1	7,25	27,90	27,90
2	1,91	7,36	35,26
3	1,45	5,57	40,84
4	1,38	5,32	46,16
5	1,32	5,09	51,25
6	1,26	4,84	56,09
7	1,07	4,11	60,21
8	1,01	3,90	64,11

Figura 42: Gráfico de Sedimentación TIMSS- Forma 8



Se aprecia unidimensionalidad de los ítems, donde el 73 % de ellos presenta una máxima carga factorial asociada a la primera componente, la que explica la mayor variabilidad del test.

Finalmente, el índice KMO arroja un valor del 0,83, índice de alto valor, lo que en definitiva indica que es coherente aplicar este análisis multivariado y que efectivamente las cargas de correlación parcial son de bajo valor.

Tabla 87: TIMSS- Forma 8

Factor	N.º de ítems	% de ítems
1	19	73 %
2	0	0 %
3	2	8 %
4	0	0 %
5	2	8 %
6	1	4 %
7	1	4 %
8	1	4 %

Tabla 88: TIMSS- Forma 8

KMO 0,837

6.1.13. TIMSS - Forma 9

La novena manera de presentar esta prueba consta de 29 preguntas, las cuales fueron respondidas por 390 alumnos, lo que arroja un coeficiente del alfa de Cronbach igual al 0,78, coeficiente aceptable, el cual permite concluir que en esta prueba, en la medida que su repetición se ejecute los resultados serán similares.

Tabla 89: TIMSS- Forma 9

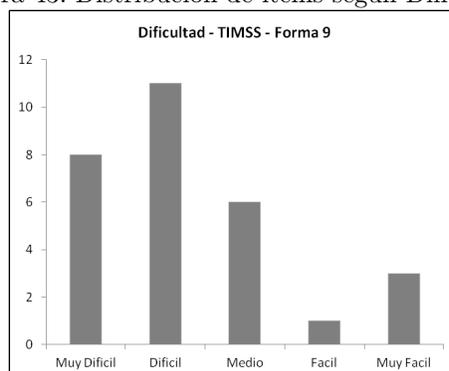
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
9	29	390	0,778

Los parámetros de dificultad de esta forma presentan valores de todos los niveles, sin embargo, aquellos de categoría fáciles no son predominantes y se hace evidente además que el 40 % de los ítems son categorizados en “Difícil”:

Tabla 90: Dificultad TIMSS- Forma 9

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	8 28 %
Díficil	0,3<D<0,5	11 38 %
Medio	0,5<D<0,7	6 21 %
Fácil	0,7<D<0,8	1 3 %
Muy Fácil	D>0,8	3 10 %
Total	29	100 %

Figura 43: Distribución de ítems según Dificultad



La discriminación del cuestionario presenta ítems cuyos indicadores son bajos.

Tabla 91: Discriminación TIMSS- Forma 9

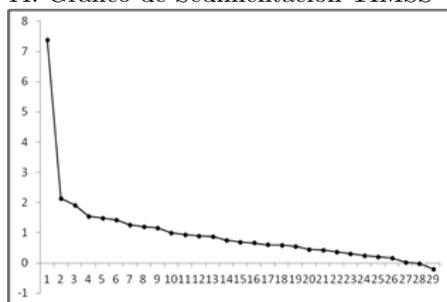
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	1	3 %
Regular	$0,15 < r < 0,25$	6	21 %
Buena	$0,25 < r < 0,35$	16	55 %
Muy Buena	$r > 0,35$	6	21 %
	Total	26	100 %

El análisis factorial arroja 9 factores capaces de explicar el 67 % de la variabilidad total del conjunto de todos los ítems, además, solo la primera componente por sí sola explica el 25 % de la variabilidad.

Tabla 92: Varianza Explicada TIMSS- Forma 9

Componente	Autovalor	% de la Varianza	% Acumulado
1	7,38	25,46	25,46
2	2,14	7,37	32,83
3	1,91	6,60	39,43
4	1,54	5,32	44,75
5	1,49	5,14	49,89
6	1,42	4,89	54,78
7	1,27	4,36	59,15
8	1,20	4,13	63,28
9	1,16	3,99	67,27

Figura 44: Gráfico de Sedimentación TIMSS- Forma 9



Se aprecia unidimensionalidad de los ítems donde el 76 % de ellos presenta una máxima carga factorial asociada a la primera componente que explica la mayor variabilidad del test. Finalmente, el índice KMO arroja un valor del 0,76, índice de alto valor, lo que en definitiva indica que es coherente aplicar este análisis multivariado y que efectivamente las cargas de correlación parcial son de bajo valor.

Tabla 93: TIMSS- Forma 9

Factor	N.º de ítems	% de ítems
1	22	76%
2	0	0%
3	0	0%
4	1	3%
5	1	3%
6	2	7%
7	2	7%
8	0	0%
9	1	3%

Tabla 94: TIMSS- Forma 9

KMO	0,756
-----	-------

6.1.14. TIMSS - Forma 10

Este cuestionario de la prueba TIMSS posee un total de 29 ítems y fue presentado a 401 alumnos, además posee un coeficiente de confiabilidad levemente superior cercano al 0,80, valor más que suficiente para aceptar su eficacia.

Tabla 95: TIMSS- Forma 10

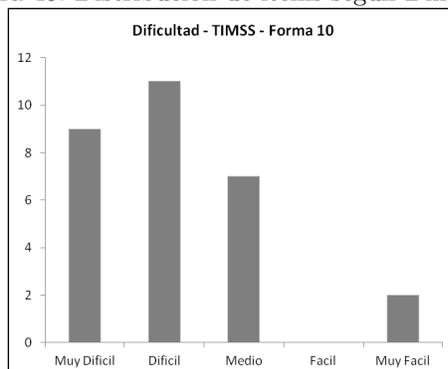
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
10	29	401	0,809

En cuanto a los parámetros clásicos, la dificultad de los ítems es de carácter transversal, puesto que considera ítems desde “Muy fácil” a “Muy Difícil”, concentrando la mayor frecuencia en el nivel “Difícil” de esta característica. A continuación se presentan los indicadores:

Tabla 96: Dificultad TIMSS- Forma 10

Grado de Dificultad	N.º de ítems	%
Muy Difícil $D < 0,3$	9	31 %
Difícil $0,3 < D < 0,5$	11	38 %
Medio $0,5 < D < 0,7$	7	24 %
Fácil $0,7 < D < 0,8$	0	0 %
Muy Fácil $D > 0,8$	2	7 %
Total	29	100 %

Figura 45: Distribución de ítems según Dificultad



En cuanto al parámetro de discriminación, a continuación se especifican los valores:

Tabla 97: Discriminación TIMSS- Forma 10

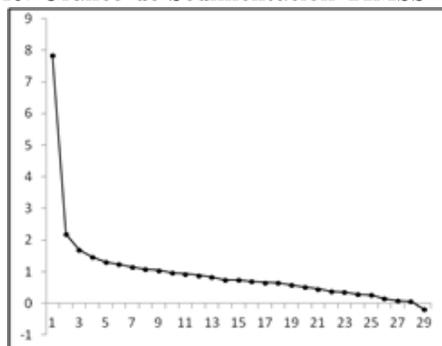
Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	0	0 %
Regular $0,15 < r < 0,25$	5	17 %
Buena $0,25 < r < 0,35$	16	55 %
Muy Buena $r > 0,35$	8	28 %
Total	29	100 %

En cuanto al análisis factorial, este se analiza para corroborar la unidimensionalidad de los ítems. Este análisis factorial indica que utilizando un total de 9 componentes, se explica un 65 % de la variabilidad total, además, la primera componente por sí sola explica el 27 %, buen indicador para demostrar que todos los ítems (o la gran mayoría) están enfocados en analizar el mismo constructo.

Tabla 98: Varianza Explicada TIMSS- Forma 10

Componente	Autovalor	% de la Varianza	% Acumulado
1	7,84	27,02	27,02
2	2,19	7,54	34,57
3	1,68	5,80	40,36
4	1,47	5,06	45,42
5	1,30	4,47	49,89
6	1,23	4,24	54,13
7	1,15	3,97	58,10
8	1,08	3,74	61,84
9	1,05	3,60	65,44

Figura 46: Gráfico de Sedimentación TIMSS- Forma 10



En cuanto a la cantidad de ítems que está relacionado con cada componente, se desprende que el 52 % de los ítems presenta una mayor carga en la primera componente, mientras que la componente menos importante,

representa sólo el 3% de los ítems.

Tabla 99: TIMSS- Forma 10

Factor	N.º de ítems	% de ítems
1	15	52%
2	3	10%
3	4	14%
4	2	7%
5	1	3%
6	0	0%
7	1	3%
8	2	7%
9	1	3%

Además, este análisis factorial posee un KMO igual al 0,8, lo que hace válida y útil la interpretación relacionada.

Tabla 100: TIMSS- Forma 10

KMO	0,798
-----	-------

6.1.15. TIMSS - Forma 11

Como todos los cuestionarios de TIMSS, la cantidad de ítems que poseen varían según el cuestionario, este presenta 26 ítems y lo contestaron 391 alumnos que además respondieron la prueba Simce; en cuanto al coeficiente de confiabilidad, este es levemente superior a 0,80.

Tabla 101: TIMSS- Forma 11

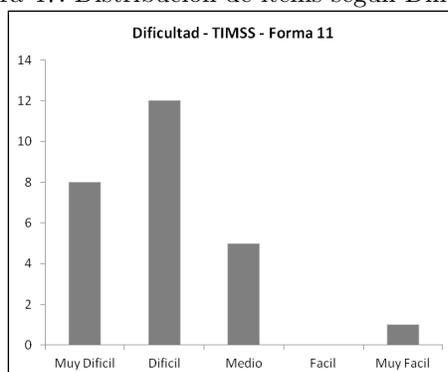
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
11	26	391	0,801

En relación con los parámetros de dificultad, este no presenta indicadores que puedan ser interpretados como fáciles, además se observa una pronunciada tendencia hacia las preguntas de mayor dificultad.

Tabla 102: Dificultad TIMSS- Forma 11

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	8 31 %
Díficil	0,3<D<0,5	12 46 %
Medio	0,5<D<0,7	5 19 %
Fácil	0,7<D<0,8	0 0 %
Muy Fácil	D>0,8	1 4 %
Total	26	100 %

Figura 47: Distribución de ítems según Dificultad



Al aplicar el análisis factorial, formado con 8 componentes, explica un 65 % de la variabilidad total, desde el que solo el primer factor considera el 29 % del total, lo que ya es un buen indicio de que efectivamente existe unidimensionalidad del test.

Tabla 103: Discriminación TIMSS- Forma 11

Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	2	8 %
Regular	$0,15 < r < 0,25$	3	12 %
Buena	$0,25 < r < 0,35$	7	27 %
Muy Buena	$r > 0,35$	14	54 %
	Total	22	100 %

Tabla 104: Varianza Explicada TIMSS- Forma 11

Componente	Autovalor	% de la Varianza	% Acumulado
1	7,66	29,47	29,47
2	1,69	6,51	35,98
3	1,55	5,95	41,93
4	1,36	5,24	47,17
5	1,27	4,90	52,07
6	1,18	4,53	56,60
7	1,09	4,19	60,79
8	1,05	4,04	64,83

Anteriormente se intuía la corroboración de unidimensionalidad, esto se confirma al verificar que del total de los ítems de esta forma, el 65 % muestra una mayor inclinación hacia el primer factor.

Tabla 105: TIMSS- Forma 11

Factor	N.º de ítems	% de ítems
1	17	65 %
2	2	8 %
3	2	8 %
4	2	8 %
5	0	0 %
6	1	4 %
7	1	4 %
8	1	4 %

Finalmente, el índice KMO es ampliamente superior al 0,7, por lo tanto se puede concluir que esta reducción de variables explicativas realizadas a través de la herramienta multivariada del análisis factorial, sí tiene sentido y hay utilidad en su aplicación.

Figura 48: Gráfico de Sedimentación TIMSS- Forma 11

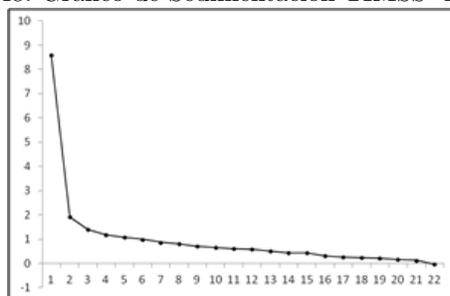


Tabla 106: TIMSS- Forma 11

KMO 0,842

6.1.16. TIMSS - Forma 12

Este cuestionario de la prueba TIMSS posee un total de 26 ítems y fue presentado a 396 estudiantes, además posee un coeficiente de confiabilidad levemente superior a 0,83, valor más que suficiente para aceptar su eficacia.

Tabla 107: TIMSS- Forma 12

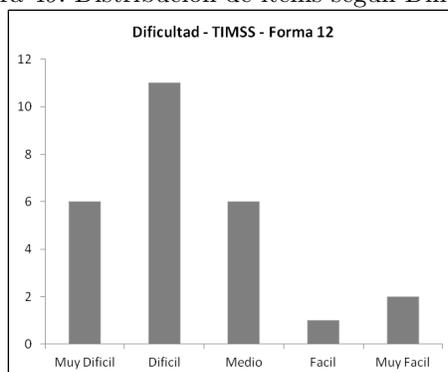
Forma	N.º de Ítems	N.º Contesta	Alfa de Cronbach
12	26	396	0,837

Los parámetros de dificultad de esta forma presentan indicadores que puedan ser interpretados desde “Muy Difícil” a “Muy Fácil”, valores que están presentados a continuación:

Tabla 108: Dificultad TIMSS- Forma 12

Grado de Dificultad	N.º de ítems	%
Muy Difícil	D<0,3	6 23 %
Difícil	0,3<D<0,5	11 42 %
Medio	0,5<D<0,7	6 23 %
Fácil	0,7<D<0,8	1 4 %
Muy Fácil	D>0,8	2 8 %
Total	26	100 %

Figura 49: Distribución de ítems según Dificultad



El respectivo análisis factorial queda determinado por 7 componentes, los que explican un 64 % de la variabilidad total, donde solo el primer componente explica poco menos del 34 %, lo que indica de nuevo que efectivamente existe un fuerte indicio para asumir desde ya la unidimensionalidad de los test.

Tabla 109: Discriminación TIMSS- Forma 12

Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	2	8 %
Regular $0,15 < r < 0,25$	0	0 %
Buena $0,25 < r < 0,35$	7	27 %
Muy Buena $r > 0,35$	17	65 %
Total	26	100 %

Tabla 110: Varianza Explicada TIMSS- Forma 12

Componente	Autovalor	% de la Varianza	% Acumulado
1	8,89	34,19	34,19
2	1,76	6,77	40,95
3	1,47	5,64	46,59
4	1,28	4,91	51,50
5	1,15	4,42	55,92
6	1,14	4,37	60,29
7	1,10	4,22	64,51

La unidimensionalidad queda confirmada luego de corroborar que el 85 % de los 26 ítems a evaluar expresan una mayor vínculo al primer componente, mientras que solo dos preguntas se apartan de esta conclusión.

Tabla 111: TIMSS- Forma 12

Factor	N.º de ítems	% de ítems
1	22	85 %
2	0	0 %
3	2	8 %
4	0	0 %
5	1	4 %
6	1	4 %
7	0	0 %

En cuanto al índice de Kaiser-Meyer-Olkin, este es equivalente al 0,88, fuerte indicador para concluir que efectivamente cobra sentido y utilidad el realizar esta reducción de variables.

Figura 50: Gráfico de Sedimentación TIMSS- Forma 12

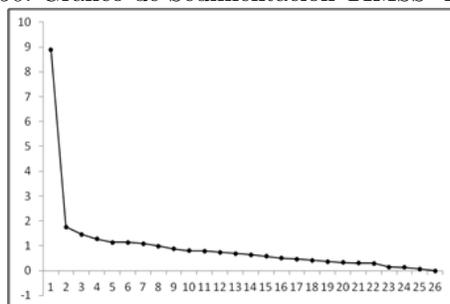


Tabla 112: TIMSS- Forma 12

KMO 0,880

6.1.17. TIMSS - Forma 13

Este formato de prueba TIMSS fue contestado por 406 alumnos que también fueron sometidos a la prueba Simce, consta de 26 ítems y posee el coeficiente de confiabilidad de 0,85, lo que en la práctica significa que el instrumento evaluador, en repetidas ocasiones, tenderá a conseguir los mismos resultados.

Tabla 113: TIMSS- Forma 13

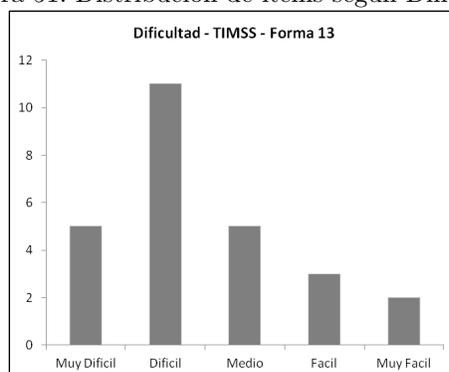
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
13	26	406	0,853

Los parámetros de dificultad de esta forma, presentan indicadores que pueden ser interpretados desde “Muy Difícil” a “Muy Fácil”, sin embargo, no se observa una pronunciada tendencia.

Tabla 114: Dificultad TIMSS- Forma 13

Grado de Dificultad	N.º de ítems	%
Muy Difícil	D<0,3	5 19 %
Difícil	0,3<D<0,5	11 42 %
Medio	0,5<D<0,7	5 19 %
Fácil	0,7<D<0,8	3 12 %
Muy Fácil	D>0,8	2 8 %
Total	26	100 %

Figura 51: Distribución de ítems según Dificultad



Al aplicar el análisis factorial, formado con 6 componentes, explica un 62% de la variabilidad total, desde la que solo el primero considera el 36% del total, lo que ya es un indicio potente de que efectivamente existe unidimensionalidad del test.

Tabla 115: Discriminación TIMSS- Forma 13

Grado de Discriminación	N.º de ítems	%
Negativa $r < 0$	0	0 %
Pobre $0 < r < 0,15$	2	8 %
Regular $0,15 < r < 0,25$	0	0 %
Buena $0,25 < r < 0,35$	6	23 %
Muy Buena $r > 0,35$	18	69 %
Total	26	100 %

Tabla 116: Varianza Explicada TIMSS- Forma 13

Componente	Autovalor	% de la Varianza	% Acumulado
1	9,54	36,69	36,69
2	1,66	6,40	43,09
3	1,57	6,03	49,12
4	1,29	4,97	54,09
5	1,12	4,30	58,39
6	1,08	4,16	62,55

Lo comentado queda completamente corroborado por el 85 % de los ítems que están orientados a un solo componente o factor de este análisis.

Tabla 117: TIMSS- Forma 13

Factor	N.º de ítems	% de ítems
1	22	85 %
2	2	8 %
3	2	8 %
4	0	0 %
5	0	0 %
6	0	0 %

Además, el KMO es de altísimo nivel (valor equivalente al 0,8).

Figura 52: Gráfico de Sedimentación TIMSS- Forma 13

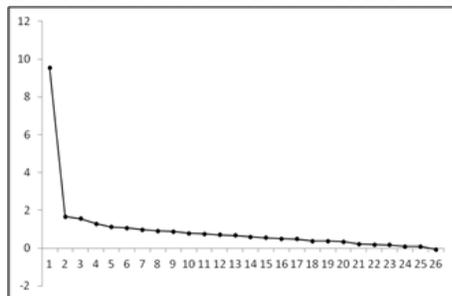


Tabla 118: TIMSS- Forma 13

KMO 0,877

6.1.18. TIMSS - Forma 14

Tal como se dijo anteriormente, la prueba TIMSS presenta cuestionarios de diferentes tamaños, y esta forma consta al menos de 22 ítems, la que fue respondida por 396 alumnos. Además, este formato de prueba presenta un valor del coeficiente del alfa de Cronbach igual a 0,825, lo que es aceptable.

Tabla 119: TIMSS- Forma 14

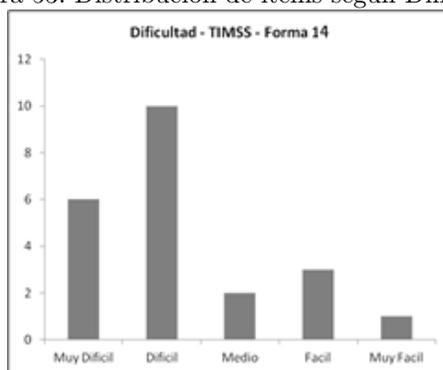
Forma	N.º de Items	N.º Contesta	Alfa de Cronbach
14	22	396	0,825

Los parámetros de clásicos de dificultad en este formato tienden a estar desplazados hacia aquellos de alto nivel, puesto que el 60% posee un valor menor al 0,5.

Tabla 120: Dificultad TIMSS- Forma 14

Grado de Dificultad	N.º de ítems	%
Muy Díficil	D<0,3	6 27%
Díficil	0,3<D<0,5	10 45%
Medio	0,5<D<0,7	2 9%
Fácil	0,7<D<0,8	3 14%
Muy Fácil	D>0,8	1 5%
Total	22	100%

Figura 53: Distribución de ítems según Dificultad



La discriminación del cuestionario, al igual que el anterior, presenta ítems cuyos indicadores son de baja categoría.

Tabla 121: Discriminación TIMSS- Forma 14

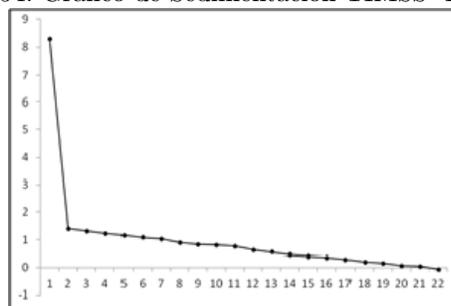
Grado de Discriminación	N.º de ítems	%	
Negativa	$r < 0$	0	0 %
Pobre	$0 < r < 0,15$	3	14 %
Regular	$0,15 < r < 0,25$	1	5 %
Buena	$0,25 < r < 0,35$	4	18 %
Muy Buena	$r > 0,35$	14	64 %
	Total	26	100 %

El análisis factorial de esta prueba arroja 7 factores de influencia, los cuales en conjunto explican el 70 % de la variabilidad total, donde casi el 38 % es manifestado solo por la primera componente.

Tabla 122: Varianza Explicada TIMSS- Forma 14

Componente	Autovalor	% de la Varianza	% Acumulado
1	8,31	37,75	37,75
2	1,42	6,46	44,21
3	1,33	6,04	50,25
4	1,24	5,62	55,87
5	1,18	5,35	61,22
6	1,09	4,96	66,18
7	1,04	4,73	70,91

Figura 54: Gráfico de Sedimentación TIMSS- Forma 14



De las cargas factoriales de los ítems, el 77 % están vinculadas al primer factor, lo que en frecuencia absoluta corresponde a 18 preguntas, mientras que las 4 restantes están repartidas en los 6 componentes restantes.

El índice KMO arroja un valor superior al 0,7, lo que hace que este análisis cobre sentido y sea de utilidad realizarlo.

Tabla 123: TIMSS- Forma 14

Factor	N.º de ítems	% de ítems
1	17	77 %
2	0	0 %
3	1	5 %
4	0	0 %
5	3	14 %
6	1	0 %
7	0	5 %

Tabla 124: TIMSS- Forma 14

KMO 0,860

6.2. Tabla de Equivalencia

Tabla 125: Tabla de Equivalencia

Percentil	Puntaje SIMCE	Puntaje TIMSS	Error Equating
1	141	303	2,5785
2	151	318	2,1038
3	156	329	2,2792
4	161	337	2,4601
5	166	347	2,2238
6	169	352	2,1087
7	173	359	1,4996
8	177	363	1,8439
9	179	368	1,9498
10	183	372	1,7306
11	185	377	1,5283
12	189	381	1,7128
13	192	385	2,1948
14	194	390	1,7679
15	196	393	1,5411
16	198	397	1,9813
17	201	400	1,7139
18	203	403	1,6453
19	206	406	1,7690
20	209	409	1,6520
21	211	412	1,7206
22	213	415	1,7216
23	215	418	1,6902
24	217	421	1,7146
25	219	423	1,4670
26	221	425	1,4176
27	223	428	1,2648
28	225	430	1,2978
29	227	432	1,4237
30	229	434	1,4197
31	231	437	1,3188
32	233	439	1,4257
33	235	441	1,5479

Percentil	Puntaje SIMCE	Puntaje TIMSS	Error Equating
34	236	443	1,2586
35	238	445	1,2867
36	239	447	1,1610
37	241	449	1,2578
38	243	452	1,2774
39	244	453	1,2552
40	245	455	1,2887
41	247	457	1,2099
42	249	460	1,2207
43	250	461	1,1384
44	252	464	1,3739
45	254	466	1,1551
46	255	467	1,1230
47	256	469	1,1133
48	258	472	1,3091
49	259	474	1,4565
50	261	476	1,3854
51	262	478	1,1417
52	264	480	1,2081
53	265	482	1,3159
54	267	485	1,4131
55	268	487	1,4897
56	270	490	1,4905
57	271	492	1,3441
58	273	494	1,2108
59	275	496	1,4947
60	276	498	1,5375
61	278	500	1,1065
62	279	502	1,1132
63	281	504	1,2600
64	283	506	1,1739
65	284	508	1,3199
66	286	510	1,3403

Nota: Esta es una continuación de la Tabla 125: Tabla de Equivalencia

Percentil	Puntaje SIMCE	Puntaje TIMSS	Error Equating
67	287	512	1,3577
68	289	514	1,3308
69	290	516	1,2914
70	292	519	1,3725
71	294	521	1,4046
72	295	524	1,3200
73	297	526	1,3668
74	298	528	1,3575
75	300	531	1,4466
76	302	533	1,2795
77	304	535	1,4095
78	305	538	1,2643
79	307	540	1,3258
80	309	543	1,3702
81	311	546	1,2780
82	313	549	1,4354
83	315	551	1,5226
84	317	554	1,6378
85	319	557	1,6872
86	321	561	1,5756
87	323	565	1,8877
88	326	569	1,6167
89	328	573	1,8824
90	331	577	1,8397
91	334	582	1,8441
92	337	586	1,7115
93	340	592	1,7504
94	344	598	1,9958
95	347	604	2,0025
96	354	612	2,2317
97	358	620	1,6102
98	363	631	2,5207
99	376	646	1,9016
100	387	712	4,0920

Nota: Esta es una continuación de la Tabla 125: Tabla de Equivalencia