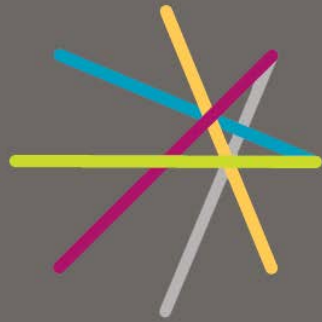




Informe Final Estudio de Profundización de Características Psicométricas del TADI Test de Aprendizaje y Desarrollo Infantil. Estudio TADI - CIAE 2015

Iván Armijo R., Héctor Galaz, Marta Edwards, Marcela Pardo



TADI

Test de Aprendizaje
y Desarrollo Infantil

INFORME FINAL

ESTUDIO DE PROFUNDIZACIÓN DE CARACTERÍSTICAS PSICOMÉTRICAS DEL TADI

(Test de Aprendizaje y Desarrollo Infantil)

ESTUDIO TADI – CIAE 2015

Iván Armijo R., Héctor Galaz, Marta Edwards, Marcela Pardo.

Santiago, Abril de 2016

INDICE

INTRODUCCIÓN	1
OBJETIVOS	1
METODOS	2
<i>Participantes</i>	2
<i>Depuración de la Muestra a Analizar</i>	3
<i>Instrumentos</i>	5
<i>Plan de Análisis</i>	11
RESULTADOS.....	12
<i>Comportamiento descriptivo de los puntajes TADI en la muestra estudiada</i>	12
<i>Validez de Constructo</i>	15
<i>Asociaciones entre las dimensiones TADI</i>	15
<i>Estructura interna de la Prueba</i>	17
<i>Categorías de interpretación de los puntaje TADI</i>	19
<i>Validez de Constructo</i>	25
<i>Análisis Preliminar de Sensibilidad y Especificidad del TADI</i>	56
<i>Análisis de Convergencia con otras variables medidas en ELPI, en las que se ha encontrado evidencia científica en relación a su influencia sobre los logros del desarrollo</i>	60
Síntesis de Resultados	68
Referencias	70

INTRODUCCIÓN

El estudio longitudinal de la primera infancia (ELPI), realizado por el Centro de Microdatos de la Facultad de Economía de la Universidad de Chile ha dejado disponible para fines de estudio la base de datos correspondiente a los registros de su primera fase de estudio, que incluye a un total de 14438 niños y niñas con edades que fluctúan entre los 7 y los 83 meses de vida (contados desde la fecha de nacimiento), incorporando datos tanto del niño o niña como de su familia.

En la ELPI se aplicó el TADI, junto a otros instrumentos que evalúan distintos aspectos del desarrollo psicológico de los niños y niñas. Además se aplicaron instrumentos a la madre o tutor del niño referidas a variables cognitivas y socioemocionales y una encuesta respecto al trabajo remunerado. Además se realizó una evaluación del ambiente del hogar y de las condiciones sociodemográficas.

OBJETIVOS

El presente Informe se refiere a los resultados obtenidos al analizar los resultados del TADI en la ELPI en relación con las otras pruebas aplicadas. Así, el objetivo general del presente estudio es:

Determinar indicadores de validez para la prueba TADI en base a los datos obtenidos en el estudio ELPI.

Para ello, se desarrollarán los siguientes objetivos específicos:

- a) Analizar la validez de constructo interno del TADI.
- b) Analizar la validez concurrente del TADI en relación a la información disponible en otras pruebas orientadas a logros de desarrollo.
- c) Determinar índices iniciales de Sensibilidad y Especificidad del TADI como herramienta para la detección de rezagos y riesgos en el logro de tareas del desarrollo temprano.

- d) Determinar la validez de constructo en relación a factores demográficos y de contexto familiar considerados como potenciales riesgos para el logro de tareas del desarrollo en infancia temprana.

METODOS

Participantes

La muestra disponible desde el estudio ELPI está compuesta por un total de 14438 niños y niñas con edades que fluctúan entre los 7 y los 83 meses de vida. Esto permite contar con una muestra de al menos 100 casos por tramo de edad, suficiente para asegurar la confiabilidad de las técnicas estadísticas utilizadas. La distribución muestral original se muestra en la tabla 1

Tabla 1.

Distribución muestral por tramo de edad y sexo del niño(a) evaluado(a)

TRAMO DE EDAD	SEXO			
	MASCULINO		FEMENINO	
	n	%	n	%
2. 6 meses,1 día a 9 meses, 0 días	66	0.46	51	0.35
3. 9 meses,1 día a 12 meses, 0 días	158	1.09	128	0.89
4. 12 meses,1 día a 18 meses, 0 días	288	1.99	296	2.05
5. 18 meses,1 día a 24 meses, 0 días	306	2.12	286	1.98
6. 2 años,0 meses,1 día a 2 años,6 meses,0 días	300	2.08	292	2.02
7. 2 años,6 meses,1 día a 3 años,0 meses,0 días	531	3.68	476	3.30
8. 3 años,0 meses,1 día a 3 años,6 meses,0 días	853	5.91	817	5.66
9. 3 años,6 meses,1 día a 4 años,0 meses,0 días	818	5.67	827	5.73
10. 4 años,0 meses,1 día a 4 años,6 meses,0 días	774	5.36	762	5.28
11. 4 años,6 meses,1 día a 5 años,0 meses,0 días	839	5.81	832	5.76
12. 5 años,0 meses,1 día a 6 años,0 meses,0 días	1572	10.89	1545	10.70
13. 6 años,0 meses,1 día a 7 años,0 meses,0 días	794	5.50	827	5.73
TOTAL	7299	50.56	7139	49.45

Depuración de la Muestra a Analizar

Los puntajes TADI fueron estandarizados por tramo de edad en base a puntajes T, con media de 50 puntos y desviación estándar de 10 puntos. Como estrategia de depuración muestral, se decidió excluir de los análisis los casos de puntuación extrema, definidos como aquellas evaluaciones con puntaje estandarizado de 81 puntos, esto es, a más de 3 desviaciones estándar sobre el promedio registrado en su tramo de edad. Se tomó esta decisión porque representan casos con rendimiento de muy baja probabilidad de acuerdo a lo esperado y que podrían estar indicando posibles fallas en la aplicación del test, o, en caso de que no existiera una falla en la aplicación, un comportamiento suficientemente anómalo como para considerarlos parte de una subpoblación de características especiales en cuanto a su logros de desarrollo. Además, se excluyó el tramo con niños sobre 6 años debido a que dicho tramo de edad se encuentra fuera del rango de aplicación oficial del TADI. Finalmente, el grupo menor a 6 meses, si bien cuenta con normas TADI, no fue incluido en el estudio ELPI del año 2012 en ninguna de las otras pruebas, de modo que fue excluido también del estudio.

La tabla 2 muestra distribución muestral final considerando estos cambios, separada por tramo de edad y sexo, y que está compuesta por un total de 11283 casos.

Tabla 2.

Distribución muestral final por tramo de edad y sexo del niño(a) evaluado(a), indicando el porcentaje de casos excluidos por obtener puntajes extremos (T=81)

TRAMO DE EDAD	% casos excluidos	HOMBRES		MUJERES	
		n	%	n	%
2. 6 meses,1 día a 9 meses, 0 días	5.13	64	1.07	47	0.81
3. 9 meses,1 día a 12 meses, 0 días	4.90	154	2.56	118	2.04
4. 12 meses,1 día a 18 meses, 0 días	2.40	280	4.66	290	5.02
5. 18 meses,1 día a 24 meses, 0 días	3.21	297	4.95	276	4.77

6. 2 años,0 meses,1 día a 2 años,6 meses,0 días	6.93	288	4.80	263	4.55
7. 2 años,6 meses,1 día a 3 años,0 meses,0 días	4.47	505	8.41	457	7.90
8. 3 años,0 meses,1 día a 3 años,6 meses,0 días	4.43	806	13.42	790	13.66
9. 3 años,6 meses,1 día a 4 años,0 meses,0 días	5.29	772	12.86	786	13.59
10. 4 años,0 meses,1 día a 4 años,6 meses,0 días	7.36	724	12.06	699	12.09
11. 4 años,6 meses,1 día a 5 años,0 meses,0 días	9.75	759	12.64	749	12.95
12. 5 años,0 meses,1 día a 6 años,0 meses,0 días	30.73	1096	18.25	1063	18.38
TOTAL		5745	50.92	5538	49.08

La pérdida de casos por exclusión de sujetos con rendimiento extremo está en torno al 5% del total de los registros, a excepción de los niños y niñas de 5 años o más, donde se eleva hasta un 30%. En el contexto general de la aplicación, sin embargo, el impacto de la exclusión de puntajes extremos se distribuye homogéneamente en los distintos tramos de edad y afecta por igual a niñas y niños, tal como se detalla en la tabla 3.

Tabla 3

Incidencia de puntajes marginales altos (puntaje T=81) en los distintos tramos de edad.

Tramo	Cognición				Lenguaje				Motricidad				Socioemocional			
	n		(%)		n		(%)		n		(%)		n		(%)	
	H	M	H	M	H	M	H	M	H	M	H	M	H	M	H	M
2	2	0	3.0	0.0	0	2	0.0	3.9	0	1	0.0	2.0	1	0	1.5	0.0
3	0	2	0.0	1.6	1	2	0.6	1.6	0	3	0.0	2.4	1	1	0.6	0.8
4	1	0	0.3	0.0	2	0	0.7	0.0	2	1	0.7	0.3	2	2	0.7	0.7
5	2	1	0.7	0.4	0	1	0.0	0.4	4	2	1.3	0.7	3	2	1.0	0.7
6	1	2	0.3	0.7	1	2	0.3	0.7	1	2	0.3	0.7	7	19	2.3	6.6
7	1	1	0.2	0.2	0	3	0.0	0.6	1	0	0.2	0.0	4	7	0.8	1.5
8	3	3	0.4	0.4	6	3	0.7	0.4	0	5	0.0	0.6	2	2	0.2	0.2
9	1	0	0.1	0.0	2	3	0.3	0.4	8	8	1.0	1.0	12	14	1.5	1.7
10	5	3	0.7	0.4	1	3	0.1	0.4	10	18	1.3	2.4	14	31	1.8	4.1

11	10	14	1.2	1.7	10	10	1.2	1.2	33	30	4.0	3.7	23	23	2.8	2.8
12	139	138	9.0	9.1	50	50	3.2	3.3	266	280	17.2	18.4	134	136	8.7	8.9

NOTA: H=HOMBRES, M= MUJERES, n=nº de casos, %= porcentaje dentro del tramo de edad

De este modo podría asumirse que los casos excluidos efectivamente reflejan un comportamiento estadístico anómalo residual de los puntajes más que un efecto significativo de alguna variable externa, de modo que se puede considerar a la muestra depurada como la mejor representante disponible del comportamiento poblacional.

Instrumentos

El estudio ELPI contó con un set amplio de pruebas y cuestionarios orientados al levantamiento de datos asociados al estudio del desarrollo en infancia temprana. Los instrumentos disponibles en ELPI usados en este informe son los siguientes

a) Pruebas Orientadas a la evaluación del Desarrollo Psicomotor General:

a. **TADI (Test de Aprendizaje y Desarrollo Infantil)**
(Edwards y Pardo, 2013)

El TADI es una escala estandarizada que permite evaluar de manera continua el desarrollo y el aprendizaje de niñas y niños entre 3 meses y 6 años de edad.

El TADI está integrado por cuatro dimensiones: Cognición, Motricidad, Lenguaje y Socioemocionalidad, cada una de las cuales constituye una escala independiente, donde los ítems están ordenados por dificultad creciente. La dimensión **Cognitiva** evalúa atención, memoria, resolución de problemas, razonamiento lógico matemático, conocimiento del mundo, interés por aprender. La dimensión **Motricidad** evalúa motricidad gruesa y motricidad fina. La dimensión **Lenguaje** evalúa comprensión del lenguaje oral, expresión oral, iniciación a la escritura e iniciación a la lectura. La dimensión **Socioemocionalidad** evalúa independencia, cuidado de sí mismo, conocimiento y valoración de sí mismo, reconocimiento y expresión de sentimientos, interacción social, formación valórica, autorregulación y vínculo afectivo cercano.

El TADI se organiza en trece tramos de edad. Ofrece puntajes estandarizados hasta los 6 años de edad, si bien cuenta con ítems para niños y niñas de hasta 7 años de edad; éstos tienen como función asegurar techo a niños y niñas que avancen más allá de su edad.

El instrumento se puntúa con puntajes T con un promedio de 50 y una desviación estándar de 10.

b. **BATTELLE DEVELOPMENTAL INVENTORY SCREENING TEST 2**

(De la Cruz & González, 1998)

Es una prueba de tamizaje, cuyos ítems son extraídos de la versión completa del Battelle Developmental Inventory, que evalúa el desarrollo infantil de 0 a 8 años de edad. El BDI fue estandarizado en Ohio, EEUU. Las etapas están caracterizadas por rangos de meses.

El BDI- ST2 evalúa 5 áreas de desarrollo:

- **Área de Destrezas de Adaptación:** capacidad para utilizar la información y las habilidades evaluadas en las otras áreas.
- **Área de Destrezas Cognitivas:** habilidades y capacidades de tipo conceptual.
- **Área de Destrezas Motrices:** desarrollo motor grueso y fino.
- **Área de Destrezas de Comunicación:** recepción y expresión de información, pensamientos e ideas por medios verbales y no verbales.
- **Área de Destrezas Socio-Personales:** capacidades y características que permiten establecer interacciones sociales significativas.

Entrega puntaje T, con promedio 50 y desviación estándar 10.

b) Pruebas Asociadas a habilidades cognitivas

a. TVIP TEST DE VOCABULARIO EN IMÁGENES DE PEABODY

(Dunn, Padilla, Lugo & Dunn, 1986)

El Test de Vocabulario en Imágenes Peabody (TVIP) es la adaptación al español del test Peabody Picture Vocabulary Test, de uso masivo en los países de habla inglesa.

Es una prueba psicométrica que mide el vocabulario receptivo o auditivo de un individuo. Se aplica a niños/as desde los 30 meses (2,5 años) de edad en adelante.

El TVIP consiste en un cuadernillo que tiene diferentes láminas con imágenes entre las cuales el niño o la niña evaluado(a) debe elegir una, según el concepto que se le señale. La puntuación del test, corresponde a 1 por cada ítem correcto.

b. BDST BACKWARD DIGIT SPAN TASK/ BACKWARD WORD SPAN TASK

(Davis & Pratt, 1996) (Zelazo, Müller, Frye, & Marcovitch, 2003)

Mide memoria de corto plazo. Esta prueba se basa en las escalas de dígitos del WISC-III (Weschler, 1992). Las tareas de memoria de trabajo incluyen la capacidad de guardar información en la memoria, así como la capacidad de procesamiento de esa información.

La tarea consiste, en que al(la) niño(a) se le enuncia oralmente un set de números que debe repetir, verbalmente, en sentido inverso. El puntaje a reportar es el número de dígitos (2, 3, 4 ó 5) del último nivel en que el (la) niño(a) tuvo al menos un éxito. Luego de la estandarización realizada con la muestra ELPI, se obtiene un puntaje T con promedio 50 y desviación estándar 10; es posible clasificar el desarrollo de los niños y niñas evaluados, en 3 niveles: bajo el promedio, promedio y sobre el promedio.

c) *Pruebas asociadas a función ejecutiva*

a. **PTT PENCIL TAPPING TASK**

(Diamond, & Taylor, 1995):

Mide la habilidad del(de la) niño(a) de mantener dos porciones de información en su mente (usar su memoria de trabajo) y la capacidad de inhibir la respuesta impulsiva. El test incluye 16 ejercicios que se puntúan 0=incorrecto y 1=correcto. Luego de la estandarización realizada con la muestra ELPI, se entrega un puntaje T con promedio 50 y desviación estándar 10 y es posible clasificar el desarrollo de los niños y niñas evaluados, en 3 niveles: bajo el promedio, promedio y sobre el promedio.

b. **HTKS HEAD TOES KNEES SHOULDERS TASK**

(McClelland, Cameron, Connor, Farris, Jewkes & Morrison, 2007).

La Tarea de “Cabeza, Pies, Rodillas y Hombros” es un test que evalúa los 3 componentes de la función ejecutiva: control inhibitorio, memoria de trabajo y flexibilidad cognitiva. Se evalúa la capacidad del niño de operar con más de una instrucción en su memoria, controlar el impulso a dar una determinada respuesta y flexibilizar las respuestas que entrega alternando el uso de distintas instrucciones o reglas. La puntuación de la prueba refiere a las dos secciones que contiene. Cada ejercicio de la prueba se puntúa 2, 1 ó 0. Luego de la estandarización realizada con la muestra ELPI, se obtiene un puntaje T con promedio 50 y desviación estándar 10; es posible clasificar el desarrollo de los niños y niñas evaluados, en 3 niveles: bajo el promedio, promedio y sobre el promedio

d) *Pruebas Asociadas a habilidades socioemocionales*

a. **SDT SNACK DELAY TASK**

(Kochanska, Murray, Jacques, Koenig, & Vandergeest, 1996):

Esta prueba mide el grado de control inhibitorio de niños en edad preescolar. La hipótesis bajo la construcción de esta tarea es que el control inhibitorio es una cualidad del temperamento que contribuye a la capacidad de internalización. La tarea de demora de la colación testea la posibilidad de los niños de demorar su gratificación. En este test se puntúa la conducta observada en el niño, obteniendo puntajes entre 0 a 4. Luego de la estandarización realizada con la muestra ELPI, se entrega un puntaje T con promedio 50 y desviación estándar 10; es posible clasificar el desarrollo de los niños y niñas evaluados, en 3 niveles: bajo el promedio, promedio y sobre el promedio.

b. **CHILD BEHAVIOR CHECK LIST**

(Achenbach & Rescorla, 2000):

El Child Behavior Checklist tiene una versión para niños entre 18 y 72 meses (CBCL 1) y otra para mayores de 72 meses (CBCL 2). El CBCL 1 tiene 99 afirmaciones y el CBCL 2, 112 afirmaciones, referentes al comportamiento que despliega el niño(a), en relación a las cuales la madre o cuidador principal debe responder en qué medida éstas se constatan en el caso de ella/él y su hijo(a). Las afirmaciones se agrupan en distintos síndromes que se agrupan en una escala de externalización y otra de internalización.

Para cada afirmación presentada a la madre existen tres alternativas posibles: “No es cierto” (0 punto), “Es cierto en cierta manera” (1 punto), o “Es muy cierto” (2 puntos). El puntaje total para cada síndrome se calcula sumando los puntajes 1 y 2 de todos los ítems de la escala. El número resultante se transforma a puntajes T y a percentil, por lo que los puntajes son equivalentes entre escalas, y percentiles.

Para cada escala de síndromes, el puntaje T obtenido se clasifica en las categorías normal, riesgo y rango clínico. Los puntajes T indican cuán elevado está la Internalización y Externalización del niño, obteniendo una imagen global de la

tendencia hacia la problemas relacionados con Internalización, Externalización, ambos o ninguna de éstas áreas.

e) *Registro de variables asociadas a entorno familiar/social*

a. **ADAPTACIÓN ESCALA HOME**

(HOME OBSERVATION FOR THE MEASUREMENT OF THE ENVIRONMENT)

(Caldwell, S.)

El objetivo de esta escala es evaluar el ambiente educativo familiar. Con este objetivo se diseñó una escala de observación basada en el Inventario HOME. A las preguntas del HOME se le añadió un par de preguntas sobre sensibilidad materna. De este modo, se incluyen las siguientes áreas:

Respuesta emocional y verbal de la madre o tutor(a); Anulación de restricción y castigo; Compromiso maternal con el niño(a) y Hogar.

Descripción de la tarea: El cuestionario consta de 21 preguntas para el evaluador, las cuales debe responder en base a la observación directa de conductas y elementos del hogar.

La puntuación se realiza con 1=es cierto y 0=no es cierto, según si se observa o no la conducta o elemento descrito.

b. **Registro Sociodemográfico**

Se consideró las características del hogar, el nivel de ingreso, edad, situación ocupacional, escolaridad y género del jefe de hogar; relación con el niño/a, estado civil, situación ocupacional, edad y escolaridad del cuidador principal.

Plan de Análisis

Dado que el desarrollo psicomotor en primera infancia es un constructo complejo, que requiere la consideración de diversas dimensiones, se decidió la realización de una serie de análisis más que la evaluación del ajuste simple a algún *Gold* estándar predefinido (el cual además, no ha sido definido en forma clara para la población chilena).

El plan de análisis definitivo incluyó las siguientes etapas:

1. Análisis del comportamiento estadístico del test en cuanto a distribución de puntaje y comportamiento de la estandarización original.
2. Análisis del constructo interno del test
3. Análisis convergente de los resultados TADI respecto de cada una de las pruebas específicas usadas para evaluar áreas de desarrollo asociadas a las dimensiones del test o al desarrollo en general.
4. Estimación de índices de sensibilidad y especificidad del test, al momento de usarlo como una herramienta para la detección de retraso en el desarrollo en primera infancia
5. Análisis del comportamiento del test en relación a predictores de posibles retrasos en el desarrollo, de acuerdo a lo definido por la evidencia científica disponible, en cuanto a factores demográficos y de entorno familiar.

RESULTADOS

Comportamiento descriptivo de los puntajes TADI en la muestra estudiada

Una vez depurada la base de datos, vemos que las distintas escalas del TADI se comportan de acuerdo a lo esperado, en cuanto a las medias y desviaciones estándar registradas (tabla 4) y a la forma de distribución de puntajes de las distintas escalas (figura 1, en la página siguiente), en las que se respeta una aproximación a la normal tanto a nivel general como en cada dimensión

Tabla 4

Resultados observados en las distintas dimensiones del TADI, en base a la estandarización original generada para el test

Escala	N	M (DE)	Min-Max	Mediana	Sesgo
Cognición	11283	49.41 (9.97)	23 - 77	50	-.18
Motricidad	11283	50.53 (10.13)	23 - 77	51	-.23
Lenguaje	11283	49.15 (10.33)	23 - 77	49	-.19
Socioemocional	11283	49.36 (9.96)	23 - 77	50	-.22
TADI Total	11283	49.61 (7.73)	23 - 75.5	50	-.26

Nota: M = Media, DS = Desviación estándar

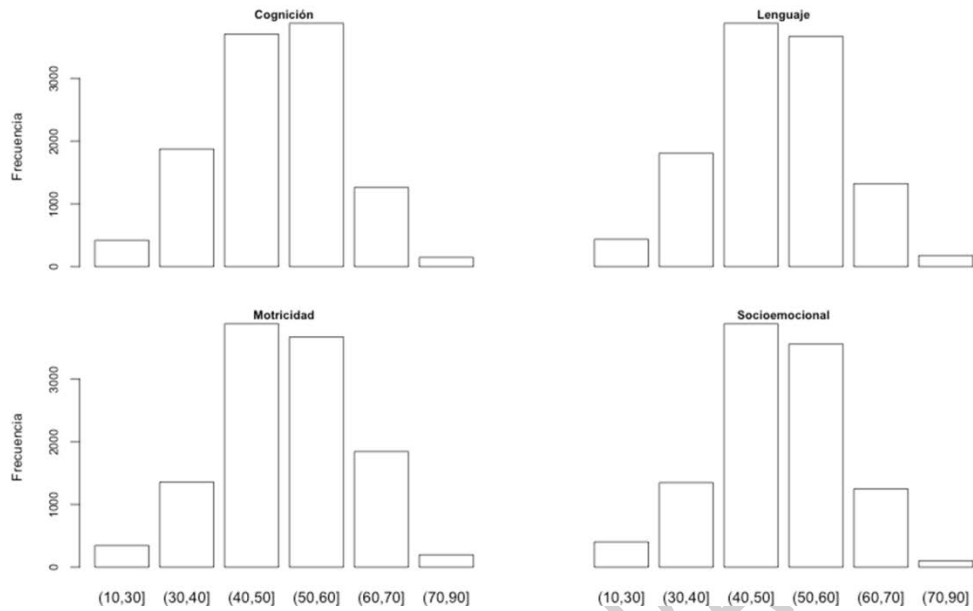


fig. 1. Distribución de puntajes T de las distintas dimensiones TADI en la muestra analizada

El puntaje total del TADI (figura 2), también se distribuye de una forma aproximadamente normal, en un comportamiento que se ajusta a lo esperado de acuerdo al procedimiento de estandarización expuesta en el manual de aplicación.

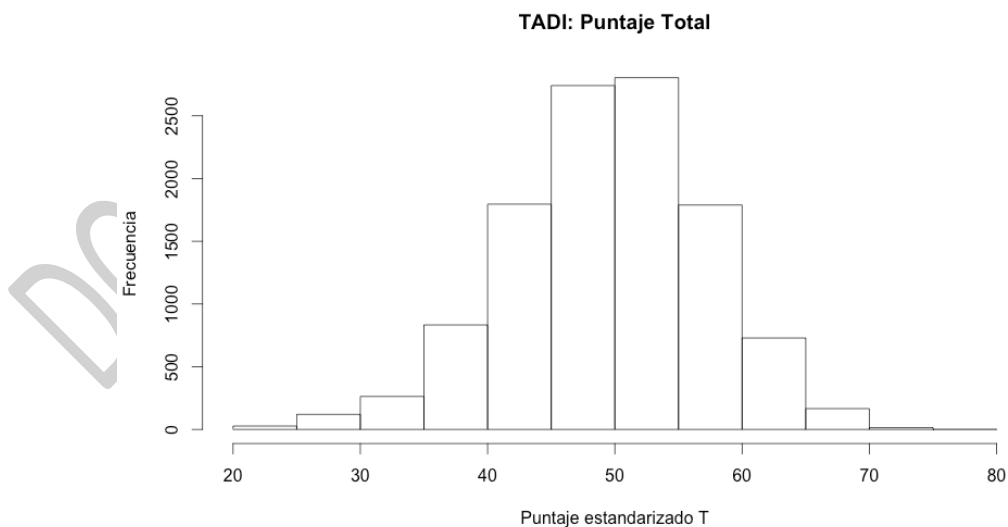


fig. 2. Histograma de Puntajes estándar TADI

Al desglosar el análisis por tramo de Edad, se mantiene el comportamiento registrado a nivel general (ver tabla 5), lo que indica que la estandarización propuesta en el manual mantiene su estabilidad en las distintas muestras (con media en 50 puntos y desviación estándar de aproximadamente 10 puntos).

Tabla 5

Tamaño de la muestra, media y desviación estándar para cada escala TADI, según tramo de edad, excluyendo casos con puntajes extremos.

Tramo	Cognición			Lenguaje			Motricidad			Socioemocional			Total		
	N	M	DE	N	M	DE	N	M	DE	N	M	DE	N	M	DE
2	111	48.7	9.3	111	52.0	10.2	111	50.6	9.7	111	50.4	11.7	111	50.4	6.9
3	272	51.1	10.1	272	51.8	10.3	272	50.8	10.4	272	51.2	9.9	272	51.2	6.9
4	570	51.7	9.6	570	51.2	9.7	570	52.5	9.5	570	51.9	9.4	570	51.8	6.3
5	573	49.7	10.8	573	51.3	10.0	573	51.0	10.3	573	50.8	10.1	573	50.7	7.6
6	551	49.4	10.3	551	50.2	10.4	551	49.7	10.6	551	51.2	8.8	551	50.1	7.4
7	962	50.2	9.6	962	51.1	9.8	962	50.8	9.8	962	50.6	9.8	962	50.7	7.6
8	1596	50.6	10.2	1596	50.5	9.5	1596	52.0	9.6	1596	50.5	9.8	1596	50.9	8.0
9	1558	48.9	10.0	1558	47.4	10.2	1558	50.5	9.4	1558	48.8	9.9	1558	48.9	7.7
10	1423	49.1	10.2	1423	48.6	10.5	1423	49.9	9.9	1423	48.9	9.4	1423	49.1	7.8
11	1508	48.6	10.0	1508	47.6	10.3	1508	50.5	11.3	1508	48.3	9.8	1508	48.8	8.0
12	2159	48.4	9.3	2159	48.1	10.8	2159	49.3	10.1	2159	47.6	10.4	2159	48.4	7.6

NOTA: N= número de casos evaluados, M=media, DE= Desviación estándar,

Validez de Constructo

Asociaciones entre las dimensiones TADI

Entendiendo que el Desarrollo Infantil es un constructo complejo, formado por distintas dimensiones interrelacionadas entre sí, se espera que los puntajes obtenidos en las distintas escalas que componen el test se asocien entre sí de manera positiva (en la medida en que fueron construidas pensando en que mayores puntajes apuntan a mayores logros de aprendizaje y desarrollo), y que todas muestren además un comportamiento coherente con el puntaje total de la prueba. Como puede verse en la figura 3, las escalas muestran asociaciones estadísticamente significativas entre ellas, registrándose la mayor asociación entre las dimensiones de Cognición y Lenguaje ($r=.52$, $p<.01$), y luego entre las escalas Socioemocional y Lenguaje ($r=.49$, $p<.01$). Si bien esta correlación podría ser interpretada como elevada y, por lo mismo, indicativa de una posible falla en la diferenciación de las dimensiones, debe tomarse en cuenta en primer lugar, la consideración estadística de que correlaciones de ese monto indican un 25% de varianza compartida entre las dimensiones asociadas, lo que deja un 75% de variabilidad independiente y, en segundo lugar, la consideración teórica respecto del rol que el lenguaje cumple como medio de operacionalización de la conducta cognitiva y socioemocional de las personas, lo cual necesariamente nos lleva a esperar que exista una unión en el comportamiento observado de las dimensiones, que se refleja en las correlaciones obtenidas.

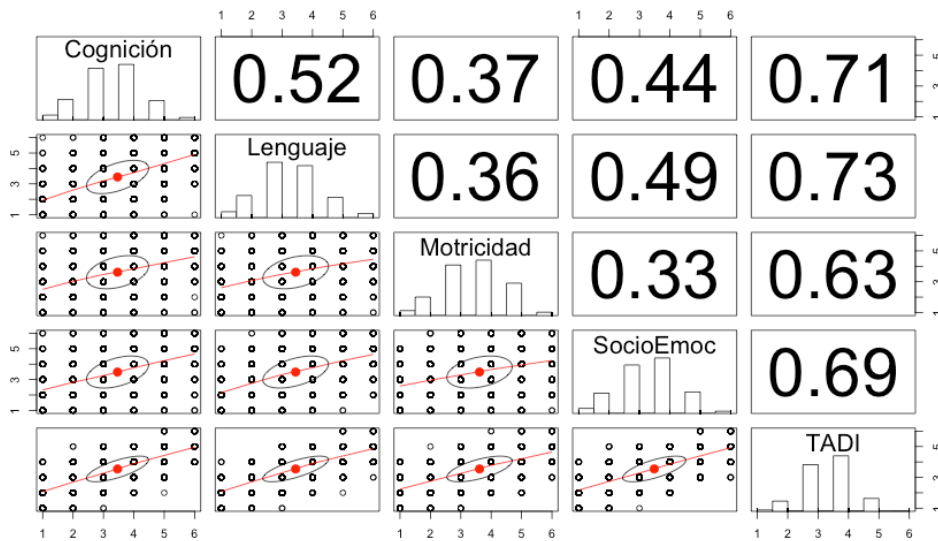


fig. 3. Correlaciones interescalas de las dimensiones de Evaluación TADI.

Nota: Los números indican el rango de puntaje: 1=20-29, 2=30-29, 3=40-49, 4=50-59, 5=60-69, 6=70-81

Vemos además que el puntaje total de la prueba muestra altas asociaciones con cada una de sus dimensiones, apoyando el supuesto propio de las pruebas de evaluación en relación a la alineación esperada entre las evaluaciones específicas (a nivel de escala) y las evaluaciones generales (TADI Total).

Estructura interna de la Prueba

La formulación teórica de la prueba TADI visualiza el constructo de desarrollo psicológico como una variable latente a la que contribuyen distintas dimensiones en las que el desarrollo se hace observable, tal como se muestra en la figura 4.

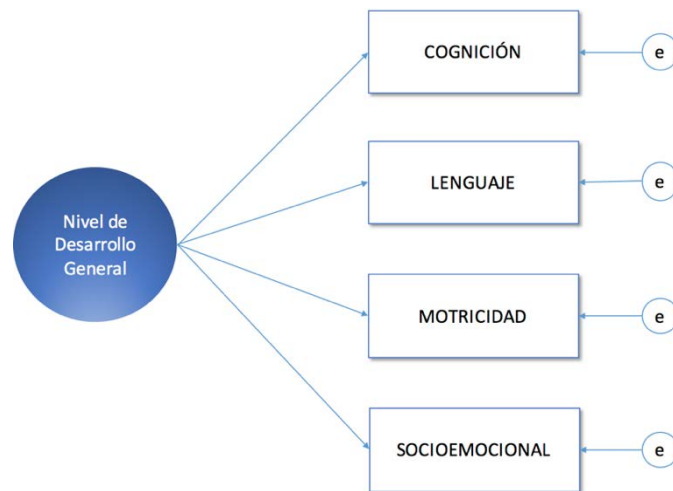


fig. 4. Modelo estructural de Evaluación de desarrollo del TADI

Se realizó un análisis factorial confirmatorio basado en estimadores de máxima verosimilitud sobre una muestra total de 11283 casos. Para el cálculo se usó el modelo Lavaan (Rosseel, 2012) implementado sobre la plataforma estadística R (R Core Team, 2015). El análisis arrojó buenos niveles de ajuste del modelo propuesto a los datos ($\chi^2(gf=2)=27,895$, $p<.01$; CFI=.998, TLI=.993, RMSEA=0.034, IC90% para RMSEA=[.023;.046], SRMR=.009). las cargas observadas entre los componentes del modelo se muestran a continuación:

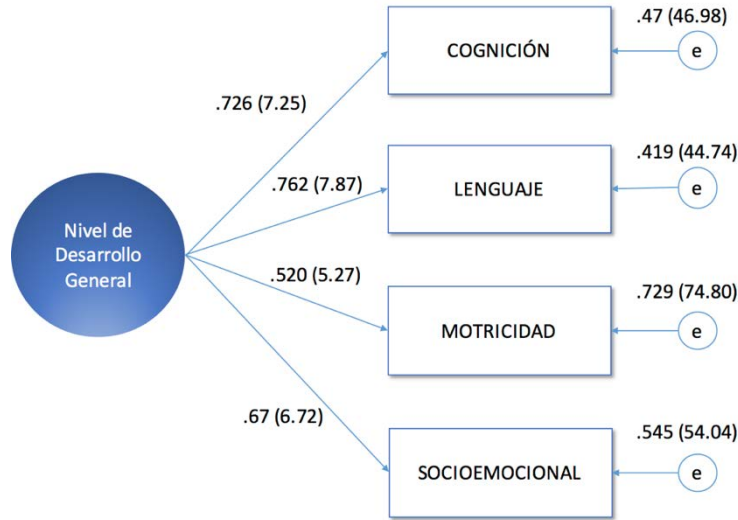


fig. 5 Modelo factorial confirmatorio para las dimensiones TADI.

NOTA: se muestran los parámetros estandarizados y en paréntesis los no estandarizados.

La estructura se mantiene en forma relativamente estable en los distintos grupos de edad, tanto en niños y niñas de 4 años o más ($n=3667$, $\chi^2(gf=2)=3.358$, $p=.187$; CFI=1.00, TLI=.99, RMSEA=0.014, IC90% para RMSEA=[.00;.04], SRMR=.006), como en niños y niñas entre 2 y 4 años ($n=5539$, $\chi^2(gf=2)=19,296$, $p<.01$; CFI=.997, TLI=.992, RMSEA=0.040, IC90% para RMSEA=[.02;.05], SRMR=.009), aunque pierde ajuste en niños menores a dos años ($n=2077$, $\chi^2(gf=2)=40.623$, $p<.01$; CFI=.967, TLI=.902, RMSEA=0.096, IC90% para RMSEA=[.07;.12], SRMR=.030). En este último caso, al agregar al modelo las covarianzas entre las dimensiones, particularmente entre las dimensiones de lenguaje y socioemocional, se logra un ajuste aceptable ($\chi^2(gf=2)=9.699$, $p=.002$; CFI=.993, TLI=.956, RMSEA=0.065, IC90% para RMSEA=[.032;.104], SRMR=.014).

El comportamiento observado, con un ajuste del modelo que va mejorando su ajuste a medida que aumenta la edad, y con una menor diferenciación entre las dimensiones en edades más tempranas es coherente con el proceso de diferenciación propia del desarrollo en la primera infancia, de modo que se puede considerar que el comportamiento observado a nivel estadístico es coherente con lo esperado desde el punto de vista teórico, lo cual puede ser considerado como una nueva evidencia que apoya la validez de constructo del TADI.

Categorías de interpretación de los puntaje TADI

La aplicación del TADI, en la medida que se utiliza para determinar los niveles de logro de metas del desarrollo en distintos tramos de edad, debe permitir la identificación de casos en los que se pueden presentar retrasos en el desarrollo y permitir, de esa manera, incluir esta información como dato en la toma de decisiones en cuanto a las intervenciones a realizar en la población medida y en la evaluación de la efectividad de dichas intervenciones. La capacidad de la prueba de detectar en forma confiable estos grupos de rezago es por tanto otro índice de validez del mismo. El uso de puntajes T como estrategia de estandarización en el TADI permite definir niveles de logro de las tareas del desarrollo en cada una de las áreas evaluadas. Tomando como referencia que el puntaje estadísticamente esperable se encuentra a una desviación estándar del promedio, se pueden definir 4 niveles de logro en torno a dicho puntaje esperado

- a. **Categoría 1: Retraso.** Corresponde a puntajes bajo 30 puntos T, esto es, bajo dos desviaciones estándar de la media esperada para su grupo de referencia, lo que implicaría muy bajos niveles de logro relativos a la edad del niño(a), y que correspondería a casos en los que se estaría mostrando un claro retraso en la dimensión de desarrollo evaluada.
- b. **Categoría 2: Riesgo de Retraso.** Corresponde a puntajes T entre 30 y 39 puntos, lo que corresponde a casos con rendimiento entre una y dos desviaciones estándar por debajo de los esperados para el grupo de referencia, configurando un grupo en el que se sospecha la posibilidad de un retraso, el cual debe ser verificado en nuevas observaciones.
- c. **Categoría 3: Normal.** Corresponde a los puntajes entre 40 y 59 puntos, y representan el comportamiento medio típicamente esperado en las escalas medidas. En esta categoría se puede hacer también una subdivisión en tres niveles: Normal con riesgo (puntajes entre 40 y 45), Normal (entre 46 y 55), y normal avanzado (entre 56 y 59)

- d. **Categoría 4: Consolidado.** Corresponde a puntajes de 60 puntos o más, en los que se registra un nivel de logro avanzado de acuerdo a lo esperado para la edad en las dimensiones evaluadas.

Esta categorización permite, por tanto, contar con un marco interpretativo para los puntajes obtenidos por los niños y niñas evaluados y permite además evaluar desde una perspectiva más cualitativa la relación entre los puntajes estandarizados TADI y los registros obtenidos en las otras escalas evaluadas. La distribución de los casos válidamente evaluados al usar este esquema clasificatorio puede verse en la tabla 6.

Tabla 6.

Frecuencia y porcentaje de incidencia de cada una de las categorías de Interpretación del puntaje TADI en la clasificación de rezago

	Dimensión TADI								TADI	
	Cognición		Lenguaje		Motricidad		Socioemocional		Total	
	n	%	n	%	n	%	n	%	n	%
Retraso	352	3.12	426	3.78	342	3.03	378	3.35	140	1.24
Riesgo	1493	13.23	1565	13.87	1266	11.22	1493	13.23	1041	9.23
Normal con Riesgo	1930	17.11	1930	17.11	1695	15.02	1997	17.70	2201	19.51
Normal	4356	38.61	4186	37.10	4640	41.12	4181	37.06	5551	49.20
Normal Avanzado	1554	13.77	1438	12.74	896	7.94	1491	13.21	1361	12.06
Consolidado	1598	14.16	1738	15.40	2444	21.66	1743	15.45	989	8.77

Al calificar los puntajes obtenidos en el TADI en base a esta categorización vemos una mayor incidencia de déficit detectado en cada una de las escalas del test (alrededor de un 16% de los casos), en comparación con el puntaje total, en el que el 10.47% de los casos caen en la zona de riesgo o retraso. Los índices de rezago y riesgo reportado son menores a los publicados previamente por revisiones preliminares del test (Bedregal, 2014), pero debe tomarse en cuenta que en esos casos se incluyó como parte del grupo de desarrollo alterado a aquellos casos que entran en la categoría de Normal con rezago, que incluye aquellos casos en los que si bien el puntaje total obtenido por el niño o niña es normal,

hay alguna dimensión en déficit. Al incluir el grupo de rezago en la definición dada en este informe, se llega a un 29,68% de casos evaluados que podrían ser considerados como parte de un grupo con un desarrollo alterado, en base al puntaje total TADI, lo que calza con la estimación de retraso entregada por diversas pruebas usadas como estrategias de tamizaje para este fin.

La elección de si se considera como desarrollo alterado al grupo que une Riesgo y Retraso, o se agrega en esa categoría el grupo Normal con rezago responde a la forma en la que será utilizada la información. En el caso de tamizajes generales pareciera ser adecuada como primera instancia de detección de posibles retrasos la inclusión del grupo Normal con rezago. Por otro lado, si se quiere hacer una detección más fina de casos, y como mostraremos más adelante, nos parece que restringir la definición de desarrollo alterado a los casos que caen en las categorías de Riesgo o Retraso es más adecuada.

En relación a la reducción de la tasa de riesgo o retraso detectado a nivel global por el TADI, en comparación con las distintas escalas. Un posible interpretación viene dada por el hecho de que al surgir ese puntaje de un promedio agregado de los puntajes de las distintas dimensiones, aquellas en las que no hay retraso tienden a compensar las áreas en riesgo y que requieren una observación más detallada del niño o niña. En consideración a lo anterior, se propone como alternativa de detección de riesgo contar el número de dimensiones en las que el niño o niña presenta riesgo o retraso. Al aplicar este criterio se registra un total de 6935 casos sin áreas en riesgo detectadas, correspondiente al 61.46% de los casos evaluados, 2452 casos con un área en riesgo o retraso, correspondiente al 21.73% de los casos, 1066 (9.45%) casos con dos áreas de riesgo o retraso, 589 (5.22%) con tres áreas en situación de riesgo o retraso y 241 (2.14%) con todas las áreas en condición de riesgo o retraso. De esta forma, ya al registrarse 1 o más áreas de retraso debería indicarse una evaluación diagnóstica especializada que permitiera verificar la condición de posible rezago en desarrollo. La distribución de la cantidad de áreas en situación de riesgo o retraso detectadas se muestra en la tabla 7

Tabla 7.

Frecuencia y Porcentaje (en paréntesis) del número de áreas en situación de riesgo o retraso en cada tramo de edad

Tramo de edad	Áreas				
	0	1	2	3	4
2. 6 a 9 meses	60 (54.05%)	35 (31.53%)	12 (10.81%)	4 (3.60%)	0 (0.00%)
3. 9 a 12 meses	165 (60.66%)	74 (27.21%)	27 (9.93%)	5 (1.84%)	1 (0.37%)
4. 1 a 1 ½ años	365 (64.04%)	161 (28.25%)	34 (5.96%)	9 (1.58%)	1 (0.18%)
5. 1 ½ a 2 años	353 (61.61%)	133 (23.21%)	58 (10.12%)	20 (3.49%)	9 (1.57%)
6. 2 a 2 ½ años	347 (62.98%)	128 (23.23%)	45 (8.17%)	24 (4.36%)	7 (1.27%)
7. 2 ½ a 3 años	653 (67.88%)	196 (20.37%)	71 (7.38%)	35 (3.64%)	7 (0.73%)
8. 3 a 3 ½ años	1149 (71.99%)	245 (15.35%)	120 (7.52%)	60 (3.76%)	22 (1.38%)
9. 3 ½ a 4 años	905 (58.09%)	332 (21.31%)	188 (12.07%)	105 (6.74%)	28 (1.80%)
10. 4 a 4 ½ años	866 (60.86%)	304 (21.36%)	122 (8.57%)	95 (6.68%)	36 (2.53%)
11. 4 ½ a 5 años	863 (57.23%)	315 (20.89%)	168 (11.14%)	106 (7.03%)	56 (3.71%)
12. 5 a 6 años	1209 (55.99%)	529 (24.50%)	221 (10.24%)	126 (5.84%)	74 (3.43%)

En base a estos resultados se puede configurar una nueva interpretación de los resultados TADI, generada en base a la interpretación de la cantidad de áreas en situación de riesgo o retraso detectadas y que representan el nivel de riesgo de retraso de desarrollo presentado por el niño o niña, y que podemos agrupar en tres categorías distintas:

- a) **Nivel 0, Sin riesgo:** correspondientes a los casos en los que no hay áreas en riesgo o retraso detectado.
- b) **Nivel 1, Observación:** casos en los que hay un área en situación de riesgo o retraso, en los que se recomendaría una nueva observación tras un tiempo de estimulación del área en riesgo detectada

- c) **Nivel 2, Riesgo:** casos en los que hay 2 o más áreas en situación de riesgo o retraso, y en los que se recomendaría una reevaluación de los logros de desarrollo por medio de expertos.

La distribución de estas categorías en los distintos tramos de edad se muestra en la figura 6 en la página siguiente.

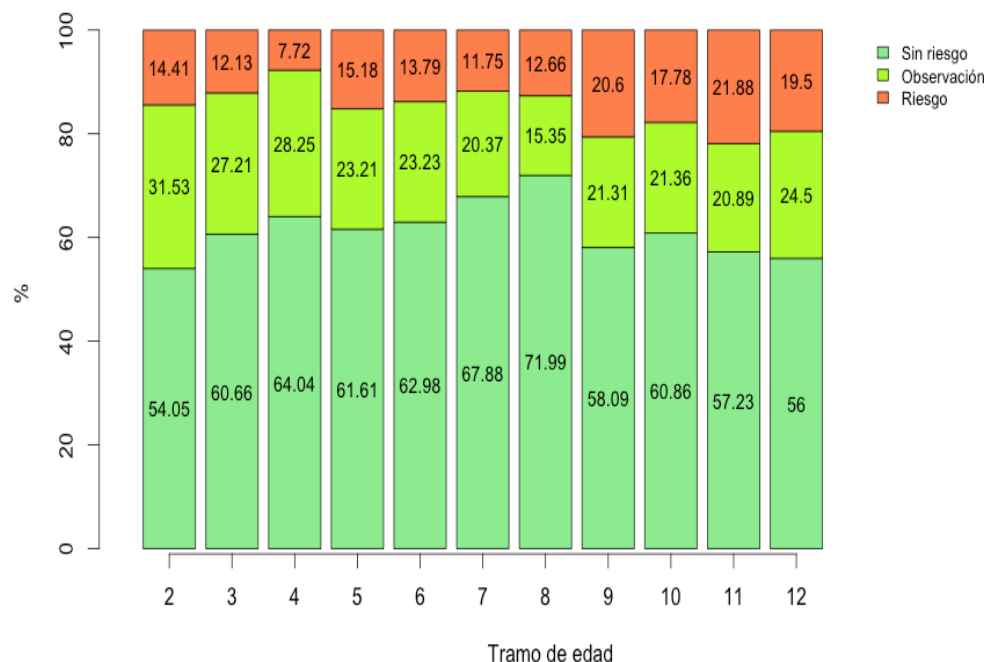


fig. 6. Distribución de categorías de riesgo según tramo de edad

La categorización de los resultados obtenidos en el puntaje Total del TADI y la categorización por Nivel de riesgo recién expuesta registran una asociación significativa entre ambas clasificaciones ($\chi^2=6595.996$, $gl=6$, $p<.001$). Específicamente, podemos observar que el tramo de puntaje normal del TADI total abarca casos de todos los niveles de riesgo (65% sin riesgo, 26% en observación, 8% en riesgo), mientras que la categoría de retraso incluye solo casos con 2 o más áreas en riesgo, y la de Riesgo incluye solo marginalmente algunos casos en zona de observación (3%), pero mayoritariamente casos con 2 o más áreas en riesgo (97%). Por su parte, aquellos casos considerados en un nivel

de desarrollo consolidado en la interpretación del puntaje TADI total se encuentran casi en su totalidad en sin áreas en riesgo (99,4%), y muy marginalmente en la zona de observación (0.6%) (ver Tabla 8)

Tabla 8.

Asociación entre categorización por interpretación del puntaje total del TADI con la categorización por Nivel de riesgo

	0. Sin riesgo	1. Observación	2. Riesgo	Total
Retraso	0 (0.00%)	0 (0.00%)	140 (100%)	140 (1.24%)
Riesgo	0 (0.00%)	32 (3.07%)	1009 (96.93%)	1041 (9.23%)
Normal	5952 (65.31%)	2414 (26.49%)	747 (8.20%)	9113 (80.77%)
Consolidado	983 (99.39%)	6 (0.61%)	0 (0.00%)	989 (8.77%)
Total	6935	2452	1896	11283

De esta forma, podemos ver que, desde el punto de vista de los análisis podemos considerar los puntajes totales en TADI que caen en la categoría de “Consolidado” como claro indicador de un nivel de desarrollo sin riesgo, mientras que los puntajes de las categorías “Riesgo” y “Retraso” como índices de posible retraso en el desarrollo del niño o niña evaluado.

En definitiva, se cuenta con tres alternativas para analizar el comportamiento del TADI:

1. Usando el puntaje estandarizado global obtenido en el test, o en cada una de sus escalas.
2. Usando el criterio de interpretación en base al puntaje estandarizado obtenido en el test total o en cada una de sus escalas, separado en los niveles: Retraso, Riesgo, Normal y Consolidado, la que además puede ser reagrupada en las categorías Retraso y Riesgo, como representantes de posibles retrasos en los logros de las metas de desarrollo de cada edad, versus el grupo Consolidado, como

representante de los casos en los que se observa un logro completo de las metas de desarrollo esperadas para el tramo de edad en el que está la niña o niño

3. Usando el criterio del nivel de riesgo detectado a partir del número de áreas que caen en la zona de riesgo o retraso, en tres categorías: Sin riesgo, Observación y Riesgo.

Estas alternativas de interpretación pueden ser usadas también como elementos de contraste de los resultados del test TADI al compararlo con otras escalas que miden el desarrollo infantil, o áreas de desarrollo específico, lo que permitirá analizar la validez convergente de los resultados del TADI.

Validez de Constructo

Dado que el estudio ELPI aplicó una serie de pruebas asociadas a medir el desarrollo infantil tanto a nivel de áreas específicas como en forma global, podemos usar dicha información para determinar la convergencia de los resultados TADI con la información entregada por esas otras pruebas.

Es necesario considerar, en todo caso, que a excepción de la prueba de Battelle, las otras pruebas miden solo áreas específicas del desarrollo y no en todos los tramos de edad. Asimismo, debe tomarse en cuenta que los instrumentos aplicados, si bien corresponden a pruebas con buenos resultados en cuanto a sus capacidades diagnósticas en la literatura científica internacional, no han sido formalmente validados en nuestro país (a excepción del test CBCL de Achenbach, aunque en el informe ELPI no se especifica haber usadp esta versión), y que por lo mismo, los resultados entregados por ELPI se basan en la estandarización hecha en Estados Unidos, la que no necesariamente calza con el comportamiento de las misma pruebas registradas en nuestro país (Lecannelier et al., 2014; Medwave Estudios Limitadas & CIGES, 2013). Por este motivo, al momento de analizar el comportamiento de las pruebas se utilizarán distintas estrategias:

- a) Usando las categorías de interpretación disponibles en base a los puntajes estandarizados entregados por ELPI.

- b) Generando las categorías de resultados que surgen de la estandarización de los puntajes brutos reportados, en caso de estar disponibles, lo que puede ser considerado una aproximación al comportamiento poblacional esperado de esa prueba, dado el tamaño muestral utilizado por ELPI. Las categorías generadas en estos casos son **bajo promedio** (con puntajes bajo una desviación estándar del promedio registrado), **promedio** (entre una desviación estándar bajo promedio y una desviación estándar sobre promedio), y **sobre promedio** (más de una desviación estándar sobre el promedio).

El comportamiento registrado en los los puntajes obtenidos desde la encuesta ELPI para las distintas pruebas según e tramo de edad estudiado, se presentan en la tabla 9.

Tabla 9

Promedio y Desviación Estándar para escalas Battelle, TVIP, BDST, PTT, HTKS, SDT, CBCL External y CBCL Internal.

Tramo	Battelle M (DE)	TVIP M (DE)	BDST M (DE)	PTT M (DE)	HTKS M (DE)	SDT M (DE)	External M (DE)	Internal M (DE)
2	47.5 (6.4)							
3	48.7 (7.8)							
4	50.4 (6.4)						55.5 (12.1)	52.0 (11.6)
5	52.5 (5.5)			49.2 (8.1)		48.5 (9.3)	54.8 (11.9)	52.8 (11.0)
6	48.8 (4.0)	100 (9.9)		49.9 (9.6)		49.4 (9.7)	54.9 (12.1)	53.9 (11.5)
7	51.5 (4.6)	99.8 (11.0)	45.59 (3.99)	50.3 (9.9)	49.4 (9.9)	49.7 (9.5)	56.1 (12.0)	56.6 (10.9)
8	52.1 (4.3)	99.7 (13.7)	66.91 (2.91)		48.8 (9.3)		55.5 (11.6)	56.6 (11.0)
9	53.9 (4.5)	99.3 (16.6)	44.51 (3.52)		48.6 (9.8)		55.1 (12.1)	56.1 (11.4)
10	54.5 (4.6)	100.9 (18.6)	44.15 (5.34)		48.7 (10.6)		54.9 (11.7)	56.8 (11.4)
11	56.5 (4.3)	101.2 (19.8)	44.50 (7.42)		49.3 (10.9)		53.7 (11.8)	56.9 (11.1)
12	56.7 (3.8)	103.3 (18.7)	44.73 (9.16)		48.7 (11.2)		53.6 (11.5)	56.7 (11.5)

Nota: BATELLE = Inventario de Desarrollo Battelle – Screening Test (BDI-ST2).

TVIP= test de Vocabulario en Imágenes de Peabody.

BDST= Backward Digit Span Task.

PTT = Pencil Tapping Test.

HTKS= Head Toes Knees Shoulders.

SDT= Snack Delay Test (SDT).

EXTERNAL =Conducta Externalizante en test CBCL.

INTERNAL= Conducta internalizante en test CBCL.

Al correlacionar los puntajes registrados en TADI, tanto a nivel global, como en cada una de sus escalas, con los puntajes registrados en las escalas externas identificadas en la tabla, vemos que se registran asociaciones estadísticamente significativas de los puntajes TADI con respecto a los resultados de las distintas pruebas externas, particularmente con Battelle ($r=.49$) y TVIP ($r=.46$). Los resultados detallados se muestran en la tabla 10.

Tabla 10.

Correlaciones Observadas entre los puntajes TADI y las pruebas de evaluación de Desarrollo aplicadas por ELPI

	Battelle (n=11056)	TVIP (n=9295)	BDST (n=8431)	PTT (n=1326)	HTKS (n=8513)	SDT (n=1325)	External (n=10349)	Internal (n=10343)
TADI	.49	.46	.24	.18	.36	.18	-.17	-.18
Cognición	.37	.42	.22	.16	.32	.10	-.14	-.17
Lenguaje	.38	.43	.25	.14	.32	.18	-.11	-.12
Motricidad	.35	.21	.12	.12	.21	.11	-.11	-.11
Socioemocional	.40	.36	.16	.14	.29	.16	-.14	-.15

Nota: BATELLE = Inventario de Desarrollo Battelle – Screening Test (BDI-ST2). TVIP= test de Vocabulario en Imágenes de Peabody. BDST= Backward Digit Span Task . PTT = Pencil Tapping Test. HTKS= Head Toes Knees Shoulders. SDT= Snack Delay Test (SDT). EXTERNAL =Conducta Externalizante en test CBCL. INTERNAL= Conducta internalizante en test CBCL.

Se enmarcan las áreas en las que las escalas TADI y las escalas externas apuntan a evaluar aspectos similares de desarrollo. Todas las asociaciones son estadísticamente significativas ($p<.01$).

Todas las escalas del TADI muestran asociaciones directas y estadísticamente significativas con los puntajes de la prueba de Battelle. Se registran también asociaciones de magnitud media, positivas y estadísticamente significativas con la prueba TVIP y HTKS, mientras que las escalas de conducta externalizante e internalizante del CBCL son la que muestran asociaciones de menor magnitud; esto es esperable en la medida en que el test CBCL está orientado a la detección de una variedad de trastornos conductuales que no

necesariamente se unen a logros de desarrollo, sino más bien a la forma de interacción y expresión emocional del niño o niña en la infancia temprana.

Las asociaciones registradas se adecúan en su dirección a lo esperado en la evaluación de las escalas TADI con respecto de las pruebas que se enfocan en forma más o menos directa sobre el comportamiento que mide cada una de ellas. No obstante lo anterior, el TADI a nivel global suele mostrar las mayores correlaciones con respecto a las pruebas externas aplicadas.

Análisis de convergencia detallado por tipo de prueba externa

Mostramos a continuación el análisis del comportamiento registrado respecto de cada prueba externa por separado, de manera de evaluar la coherencia de la información entregada por TADI y cada una de las escalas externas. En el caso de las pruebas que miden áreas específicas, se evaluó su comportamiento en relación a la dimensión del TADI que teóricamente debería estar más asociada con sus puntajes, además del puntaje TADI TOTAL.

La estrategia de análisis incluyó dos evaluaciones

1. Puntajes TADI registrados por personas que caen en distintas clasificaciones de la prueba externa, cuando éstas están disponibles, o con la clasificación derivada de la estandarización de los puntajes de la prueba externa, cuando éstos están disponibles.
2. Asociación entre las clasificaciones que surgen de la prueba externa y las categorías de interpretación de puntaje extraídas de los puntajes TADI, en base a dos criterios:
 - a. En base a la clasificación de nivel de logro de las tareas del desarrollo para cada tramo de edad, medido por el puntaje estandarizado total TADI, en 4 categorías: Retraso (puntaje T bajo 30) , Riesgo (puntaje T entre 30 y 39) , Normal (puntaje T entre 40 y 60) y Consolidado (puntaje T sobre 60).

- b. En base a la clasificación de Nivel de Riesgo, que depende de la cantidad de áreas clasificadas en riesgo o rezago en cada niño o niña, expresado en tres categorías: Nivel 0: Sin riesgo detectado en alguna de las escalas TADI, Nivel 1: Observación, cuando se observa 1 área en condición de riesgo o rezago, Nivel 2: Riesgo, cuando se registran 2 o más escalas en condición de riesgo o rezago.

Análisis de convergencia con los resultados del test Battelle

Comportamiento de puntajes

Se observa la presencia de diferencias estadísticamente significativas en los puntajes TADI obtenidos por los casos clasificados como sobre promedio, en promedio o bajo promedio de acuerdo a la prueba de Battelle, registrando un aumento progresivo de puntajes en el TADI a medida que se progresa en las categorías de puntaje en la prueba de Battelle.

Tabla 11

Nivel de ajuste entre TADI Total y Categorías Battelle.

	Categorías BATTELLE						F	Diferencias
	Bajo Promedio		Promedio		Sobre Promedio			
	N=1731	%=15.7	N=7915	%=71.6	N=1410	%=12.8		
	M (DE)		M (DE)		M (DE)			
TADI Total	43.82 (8.03)		49.68 (6.90)		56.51 (5.60)		1299**	Todas las categorías

Nota: **: p menor a 0.01; *: p menor a 0.05. M= Promedio, DE = Desviación Estándar

A nivel numérico, es interesante constatar que en aquellos casos que obtienen puntajes bajo promedio en Battelle, la media en puntaje global TADI cae en la categoría de Normal con riesgo (M=43.82, DE= 8.03), lo cual podría estar indicando una baja sensibilidad del TADI al momento de identificar casos con retrasos del desarrollo. No obstante lo anterior, los puntajes T de la misma prueba de Battelle, en la estandarización usada en ELPI2, deja

al 75% de los casos sobre un puntaje T de 50 puntos, y solo un 1.6% bajo el puntaje T 40, con una clara asimetría negativa en los puntajes registrados (fig.7) lo que indicaría una tendencia a la sub-detección de rezago (en comparación con lo esperado) si nos basamos exclusivamente en la interpretación clásica de los puntajes Battelle.

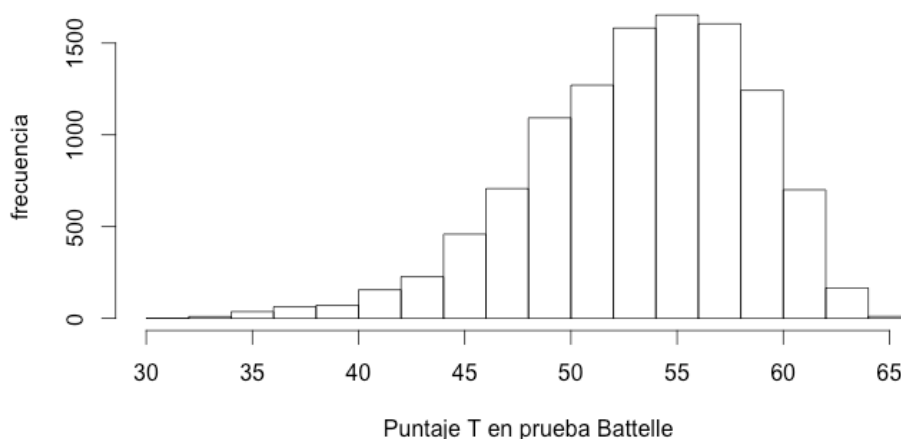


fig. 7 Distribución observada en puntajes estandarizados T de la prueba de Battelle

Considerando lo anterior, la clasificación de los casos usando en puntaje T reportado en Battelle como indicador no estandarizado, y la generación de la clasificación en tres niveles reportada en la tabla 11 permite una mirada que respeta de mejor forma el comportamiento de los puntajes registrados en la muestra nacional y hace necesario, además, un análisis más específico de la forma en que se comportan las clasificaciones asociadas a los puntajes obtenidos en ambas pruebas, el que se presenta a continuación.

Acuerdo entre categorías de clasificación Battelle y TADI

*Comparación en base a distribución de puntajes T de la prueba Battelle
(bajo promedio, promedio, sobre promedio)*

La tabla 12 muestra el comportamiento registrado al comparar las categorías de clasificación Battelle con las clasificaciones de Nivel de Desarrollo y Nivel de Riesgo asociadas a los puntajes TADI. Como puede verse, se detectan asociaciones estadísticamente significativas entre la clasificación de casos asociada al test de Battelle y las categorías de interpretación asociadas a los puntajes TADI. En particular, podemos ver que solo el 2.8% de los casos clasificados como de desarrollo consolidado caen bajo el promedio en la prueba de Battelle, en forma similar a lo que ocurre en la clasificación por nivel de riesgo, donde solo el 0.8% de los casos clasificados como riesgo en TADI caen sobre el promedio en Battelle. Por su parte, en los casos clasificados como retraso por TADI, el 80% se encuentra bajo el promedio en Battelle y 0% sobre el promedio, y en los casos clasificados en Nivel 2 de riesgo según TADI, el 36.5% se encuentra bajo el promedio en Battelle.

Tabla 12

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a puntajes observados de Battelle

Categorías TADI	Categorías Battelle			
	Bajo Promedio	Promedio	Sobre Promedio	
Retraso	105 (80.2%)	26 (19.8%)	0 (0.0%)	$\chi^2_{(g=6)}=1726.68^{***}$
Riesgo	413 (40.4%)	606 (59.3%)	3 (0.3%)	
Normal	1185(13.3%)	6728 (75.3%)	1016 (11.4%)	
Consolidado	28 (2.9%)	555 (57.0%)	391 (40.1%)	
Nivel Riesgo TADI				
0: Sin riesgo	603 (8.9%)	4957 (72.8%)	1253 (18.4%)	$\chi^2_{(g=4)}= 1221.9^{***}$
1: Observación	453 (18.9%)	1800 (75.2%)	142 (5.9%)	
2: Riesgo	675 (36.5%)	1158 (62.7%)	15 (0.8%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

En base a la clasificación por niveles de desarrollo del TADI (retraso, riesgo, normal o consolidado), si se analiza la capacidad de detección de rezago del TADI (categorías Retraso o Riesgo) en comparación con Battelle (asumiendo como rezago la categoría Bajo

promedio), se obtienen valores altos en especificidad pero bajos en relación a la capacidad de detección de riesgo de retraso del desarrollo por parte del TADI (Sensibilidad = .30, Especificidad =.93, VPP =.45, VPN=.88¹). Estos indicadores mejoran cuando se considera como muestra sin rezago según el TADI solo a aquellos casos que caen en la categoría de “Consolidado” en su Nivel de Desarrollo (sensibilidad = .95, especificidad =.60, VPP =.45, VPN=.97). Estos resultados, sin embargo, pueden estar alterados por el comportamiento registrado en los puntajes T de Battelle, que hacen que la categoría promedio puede estar integrando un número indeterminado de casos con rezago, que por lo tanto no están siendo detectados. En este sentido, considerar el grupo promedio como grupo sin riesgo de rezago sería un error. De hecho, si excluimos las categorías medias de ambos test (las que pueden incluir casos de riesgo de rezago no detectado) y analizamos aquellos grupos en que se puede asumir más claramente un posible rezago (categorías Retraso o riesgo en el TADI, y categoría bajo promedio en Battelle), o un logro claro de metas del desarrollo (categoría consolidado en el TADI), se registra una clara concordancia entre las pruebas (96% de los casos) y con un aumento notable en la capacidad de detección de rezago del TADI al usar a Battelle como referencia (Sensibilidad=.95, Especificidad=.99, VPP=.99, VPN=.93).

Por otro lado, al analizar la capacidad de detección de rezago de la categorización por nivel de Riesgo en TADI comparado con las categorías Battelle Generales en la muestra global, encontramos índices aceptables (Sensibilidad=.65, Especificidad =.67, VPP=.27, VPN=.91), los que al excluir las categoría promedio del Battelle, mejoran en especificidad pero no en sensibilidad (Sensibilidad=.65, especificidad=.88, VPP=.88, VPN=.67).

¹

Sensibilidad = Capacidad del TADI de detectar casos con retraso

Especificidad= capacidad del TADI para detectar casos sin retraso

VPP = Valor Predictivo Positivo, o probabilidad de que los casos que TADI diga que tienen rezagos efectivamente los tienen,

VPN, Valor Predictivo Negativo, o probabilidad de que los casos que TADI diga que no tienen retraso efectivamente no los tengan

Todos estos indicadores se calculan en referencia a la prueba contra la que se compara el TADI

Comparación de acuerdo a puntos fuertes y débiles en Battelle

El test de Battelle permite también la codificación de los casos evaluados de acuerdo a la predominancia de puntos fuertes o débiles del desarrollo, que corresponde a puntajes globales por sobre o bajo 1,5 desviaciones estándar de la media obtenida, de acuerdo a lo definido por el equipo encargado de la encuesta ELPI en sus dos primeras aplicaciones (Centro Micro Datos Departamento de Economía, 2012). Al comparar la interpretación asociada a los puntajes TADI con la interpretación asociada al test de Battelle, en base a la predominancia de puntos fuertes o débiles en el desarrollo (tabla 13), vemos que, en los casos clasificados como retraso por el TADI, el 90.8% queda clasificado con predominancia de puntos débiles por Battelle. Mientras que, en el caso del desarrollo consolidado, el 40.1% queda clasificado con predominancia de puntos fuertes en Battelle. En este contexto, considerando como indicador de rezago los casos en los que predominan los puntos débiles de Battelle (que corresponden al 25.54% de los casos evaluados), y al resto como casos sin riesgo, el TADI logra niveles bajos de sensibilidad y buenos en especificidad de capacidad diagnóstica (Sensibilidad=.25, Especificidad=.94, VPP=.62, VPN=.79). Estas propiedades se mantienen al excluir del análisis a casos clasificados en el rango normal en Battelle (Sensibilidad=.25, Especificidad=.99, VPP=.99, VPN=.39). Al igual que en el análisis previo, en que se usaba el promedio como criterio para clasificar los casos normales vs los retrasos en Battelle, en este caso, si consideramos solos los grupos que claramente indican retraso y logro de desarrollo (Riesgo-Retraso vs Consolidado en TADI y Débiles vs fuertes en Battelle, los indicadores aumentan considerablemente (Sensibilidad=.93, especificidad=.99, VPP=.99, VPN=.88)

Tabla 13

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo puntos fuertes y débiles de Battelle

Categorías TADI	Interpretación Battelle			
	Puntos Fuertes	Normal	Puntos Débiles	
Retraso	0 (0.0%)	12 (9.2%)	119 (90.8%)	$\chi^2_{(gl=6)}=1698.2^{***}$
Riesgo	3 (0.3%)	425 (41.6%)	594 (58.1%)	
Normal	1016 (11.4%)	5859 (65.6%)	2054 (23.0%)	
Consolidado	391 (40.1%)	526 (54.0%)	57 (5.9%)	
Nivel Riesgo TADI				
0: Sin riesgo	1253 (18.4%)	4443 (65.2%)	1117 (16.4%)	$\chi^2_{(gl=4)}= 12.9^{***}$
1: Observación	142 (5.9%)	1503 (62.8%)	750 (31.3%)	
2: Riesgo	15 (0.8%)	876 (47.4%)	957 (51.8%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

Por otro lado, al revisar la clasificación en base a nivel de riesgo según TADI, vemos que solo el 16% de los casos clasificados como sin riesgo por TADI caen en la categoría de predominancia de puntos débiles según Battelle, mientras que en el nivel de riesgo 2, indicador de retraso en el desarrollo, esta cifra suba al 51.8%. En base a este criterio, se observa una capacidad diagnóstica media del TADI en comparación con Battelle al hacer el análisis con el total de la muestra (Sensibilidad=.60, especificidad=.69, VPP=.40, VPN=.84), los que se mantienen al excluir de la muestra los casos clasificados como normales en Battelle (Sensibilidad=.60, especificidad=.88, VPP=.92, VPN=.53).

En definitiva, al usar Battelle como punto de referencia, podemos observar una capacidad aceptable de identificación de casos con rezago, particularmente usando el criterio de nivel de riesgo, pero menores al usar la clasificación en categorías de logro de desarrollo. Sin embargo, al usar como referencia los grupos que claramente indicarían rezago o logro en el desarrollo en Battelle, observamos un aumento considerable en el valor predictivo positivo del TADI, lo que indicaría que efectivamente los casos identificados como en riesgo serían casos que efectivamente muestran ese rezago según Battelle. Al considerar

solo los casos que son clasificados claramente como en riesgo versus los casos con logro consolidado por parte del TADI, por su parte, todos los valores llegan a niveles óptimos. Esto indicaría que mayor investigación debería hacerse sobre los casos promedio, lo cual se aplica también en relación a los puntajes Battelle, dada la forma en la que se distribuyeron los puntajes estandarizados de esa prueba en la muestra nacional.

Análisis de convergencia con los resultados de pruebas asociadas a Tareas Cognitivas

Comparación en relación a puntajes en el Test de Vocabulario en Imágenes de Peabody (TVIP)

La prueba TVIP se orienta a medir el vocabulario receptivo, sin requerir la capacidad lectora del examinado y permite hacer una estimación de la inteligencia verbal del niño o niña. La prueba fue validada en Estados Unidos y en población hispanoparlante procedente de México y Puerto Rico. Si bien existe una validación en Chile (Echeverría, Herrera, & Segure, 2002), no se especifica que esta versión haya sido utilizada en ELPI, lo que hace pensar que se utilizó la versión norteamericana adaptada para Hispanoamérica (Dunn, Padilla, Lugo, & Dunn, 1986), la que es usada regularmente como instrumento de evaluación de la comprensión de lenguaje (Strasser, Larraín, López de Lérída, & Lissi, 2010). La distribución de puntajes TVIP estandarizados registrados se aproxima a la normal, tal como se muestra en la figura 8.

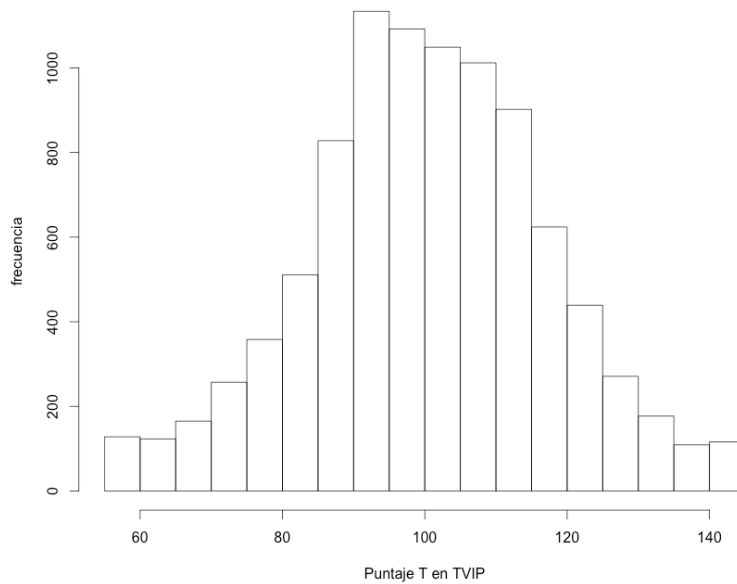


fig. 8. Distribución observada de puntajes estandarizados TVIP

En teoría, los casos que muestran niveles bajo lo esperado en TVIP deberían tender a coincidir con las categorías de riesgo en los puntajes tanto en la escala Total como en las dimensiones de Lenguaje y Cognición. La tabla 14 muestra el comportamiento de los puntajes TADI usando como clasificación la interpretación del puntaje estandarizado TVIP, en tres categorías: “bajo promedio” (que incluye las categorías “extremadamente bajo” y “moderadamente bajo” de la interpretación normativa TVIP), “promedio” (que corresponde a las categorías “promedio bajo”, “promedio” y “promedio alto” TVIP) y “Sobre Promedio” (correspondiente a las categorías “moderadamente alto” y “extremadamente alto” de TVIP).

Tabla 14

Nivel de ajuste entre TADI Total, TADI Cognición y TADI Lenguaje y Categorías TVIP.

	<u>Categorías TVIP</u>						F	Diferencias
	<i>Bajo Promedio</i>		<i>Promedio</i>		<i>Sobre Promedio</i>			
	N=1489	%=16.0	N=5879	%=63.25	N=1927	%=20.73		
	M (DE)		M (DE)		M (DE)			
TADI Total	43.12 (8.45)		49.37 (7.00)		53.87 (6.28)		957.7**	Todas las categorías
TADI Cognición	42.40 (10.26)		49.04 (9.11)		54.84 (8.64)		776.8**	Todas las categorías
TADI Lenguaje	41.07 (10.57)		48.74 (9.42)		54.45 (9.24)		830.3**	Todas las categorías

Nota: **: $p < 0.01$; *: $p < 0.05$

Los puntajes TADI registran diferencias estadísticamente significativas al comparar las categorías TVIP. De acuerdo a lo esperado, los puntajes TADI son inferiores en la categoría bajo promedio en TVIP, y aumentan a medida que se sube en rendimiento en vocabulario según las medidas de ese test. Los puntajes medios de TADI observados en las categorías inferiores de TVIP, sin embargo, caen en la categoría Normal con Riesgo, aunque con una desviación estándar que hace suponer que igualmente una proporción de los casos detectados como bajo promedio por TVIP caen también en las categorías de riesgo TADI. Una mirada en detalle del comportamiento de las clasificaciones TADI en comparación con TVIP se muestra en la tabla 15.

Tabla 15

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a puntajes TVIP

Categorías TADI	<u>Interpretación TVIP</u>			$\chi^2_{(g =6)}=1375.04***$
	<i>Bajo Promedio</i>	<i>Promedio</i>	<i>Promedio Alto</i>	
Retraso	111 (84.1%)	21 (15.9%)	0 (0.0%)	
Riesgo	413 (43.8%)	498 (52.9%)	31 (3.3%)	
Normal	933 (12.5%)	4943 (66.3%)	1575 (21.1%)	
Consolidado	32 (4.2%)	417 (54.2%)	321 (41.7%)	

Nivel Riesgo TADI				
0: Sin riesgo	487 (8.6%)	3653 (64.2%)	1548 (27.2%)	$\chi^2_{(g=4)} = 1209.09^{***}$
1: Observación	324 (16.7%)	1321 (67.9%)	300 (15.4%)	
2: Riesgo	678 (20.8%)	905 (54.4%)	79 (17.9%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

Vemos que los casos clasificados como retraso por TADI se concentran en la categoría bajo promedio del TVIP (84.1%), mientras que los casos clasificados en riesgo se clasifican en un 96.7% en las categorías bajo promedio o promedio TVIP, mientras que la categoría Normal o Consolidado registran un 12.5% y un 4.2% de casos en la categoría bajo promedio del TVIP. Dado que TVIP se enfoca únicamente en vocabulario, las alteraciones detectadas por el test no se corresponden necesariamente con un retraso a nivel global, de modo que el comportamiento observado no está fuera de lo esperado a la luz de los análisis hechos. En cuanto a los índices de capacidad diagnóstica, el mejor comportamiento se logra al usar como criterio de comparación la separación entre bajo y alto promedio en TVIP y los grupos Riesgo y Retraso vs Consolidado en TADI (Sensibilidad=.94, especificidad=.91, VPP=.94, VPN=.91), perdiendo sensibilidad al considerar la muestra total. La clasificación por Nivel de Riesgo mantiene niveles de sensibilidad de .67, tal como se registró en relación al test de Battelle.

Comparación en relación a puntajes de la Tarea de Retención de Dígitos Inversa (Backward Digit Span Task: BDST)

La prueba BDST entrega indicios en relación a la función ejecutiva lograda por el niño, al combinar capacidades auditivas, memoria, e inhibición en una tarea simple, la que además, ha mostrado su sensibilidad al desarrollo del niño, con incrementos asociados a la edad (Hongwanishkul et al., 2010). La obtención de puntajes bajo el promedio en esta prueba puede ser considerado como un indicador de posible alteración en el desarrollo. La distribución observada de puntajes en esta prueba se muestra en la figura 9.

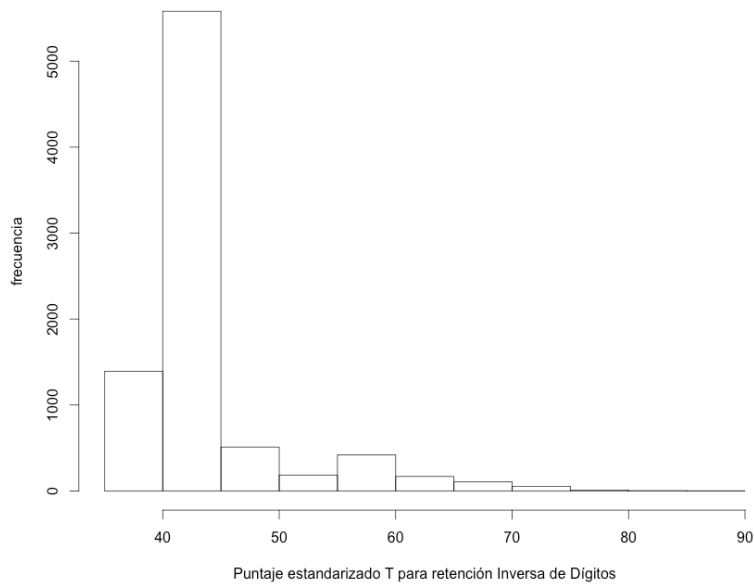


fig. 9. Distribución de puntajes estandarizados obtenidos en prueba de Retención Inversa de Dígitos (BDST)

El reporte de los puntajes estandarizados obtenidos por los casos evaluado por ELPI muestra una clara asimetría positiva, con resultados concentrados hasta el puntaje 45 (82.7% de los casos evaluados). Dado que este puntaje se usa normativamente para determinar los casos bajo promedio en esta prueba, esto genera que la gran mayoría de los casos queden clasificados bajo promedio. Aún así, se encontraron diferencias estadísticamente significativas en los puntajes TADI obtenidos por niños y niñas categorizados en los niveles bajo promedio, promedio y sobre promedio en la prueba BDST, con un aumento de puntaje tanto en la dimensión de cognición como en la prueba total.

Tabla 16

Nivel de ajuste entre TADI Total, TADI Cognición y TADI Lenguaje y Categorías BDST.

	<u>Categorías BDST</u>						F	Diferencias
	<i>Bajo Promedio</i>		<i>Promedio</i>		<i>Sobre Promedio</i>			
	N=6973	%=82.7	N=1347	%=16	N=111	%=1.3		
	M (DE)		M (DE)		M (DE)			
TADI Total	48.40 (7.88)		52.29 (6.42)		55 (6.56)		178.9**	Todas las categorías
TADI Cognición	48.19 (9.90)		52.81 (8.72)		56.28 (8.85)		158.7**	Todas las categorías

Nota: **: p menor a 0.01; *: p menor a 0.05

Al comparar con las categorías de nivel de desarrollo y nivel de riesgo en base a los puntajes TADI, se observa que la presencia de casos bajo promedio en BDST disminuye en forma estable a medida que se registra un mayor logro en desarrollo, y aumenta a medida que aumenta el riesgo (tabla 17).

Tabla 17

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo puntajes BDST

Categorías TADI	<u>Interpretación BDST</u>			
	<i>Bajo promedio</i>	<i>Promedio</i>	<i>Sobre Promedio</i>	
Retraso	126 (97.7%)	3 (2.3%)	0 (0.0%)	$\chi^2_{(gl=6)}=210.07***$
Riesgo	846 (95.9%)	34 (3.9%)	2 (0.2%)	
Normal	5534 (81.8%)	1153 (17.0%)	82 (1.2%)	
Consolidado	467 (71.7%)	157 (24.1%)	27 (4.1%)	
Nivel Riesgo TADI				
0: Sin riesgo	4030 (78.8%)	996 (19.5%)	89 (1.7%)	$\chi^2_{(gl=4)}= 210.49***$
1: Observación	1477 (83.6%)	271 (15.3%)	18 (1.0%)	
2: Riesgo	1466 (94.6%)	80 (5.2%)	4 (0.3%)	

Nota: *** = p <.001. Los porcentajes indican la incidencia de los casos en cada fila

A nivel diagnóstico, el uso de la categorización BDST como criterio de comparación no se considera adecuado, dado que en base a la interpretación propuesta para estos puntajes el 80% de los casos mostraría algún nivel de retraso. Por lo mismo, se decidió re categorizar los puntajes T de la prueba BDST, considerando como bajo promedio aquellos casos con puntajes hasta 40, los que equivalen al 16.5% de la muestra evaluada. Usando ese criterio y usando la muestra completa de casos evaluados por TADI se logran bajos niveles de sensibilidad (.22) y altos niveles de especificidad (.90), mientras que al usar la muestra restringida a casos detectados como riesgo o retraso versus desarrollo consolidado aumenta la sensibilidad (.90) y baja la especificidad (.47). El uso del criterio de nivel de riesgo, por su parte, obtiene valores moderados de sensibilidad (.57) y especificidad (.64).

Análisis de convergencia con los resultados de pruebas asociadas a función ejecutiva.

Comparación en relación a puntajes en Pencil Tapping Task (PTT)

El Pencil Tapping test (PTT) es otra escala asociada a la evaluación de función ejecutiva, y se utiliza en escalas de evaluación de autorregulación en pre-escolares (Smith-Donald, Raver, Hayes, & Richardson, 2007) y ha mostrado asociaciones significativas con el rendimiento académico posterior del niño (Willoughby, Kupersmidt, & Voegler-Lee, 2012). La prueba se orienta a medir el control inhibitorio, ya que el niño o niña debe realizar una acción distinta a la que se le indica verbalmente. La distribución de puntajes en observada en la muestra presenta, al igual que en la prueba BDST, un claro sesgo positivo con una condensación de un 69% de los casos registrados en los puntajes 45 (el mínimo observado) y 46.

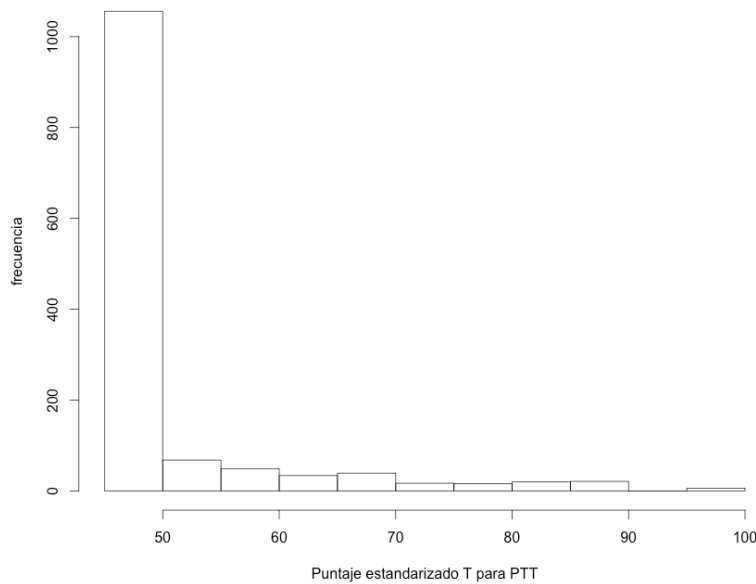


fig. 10. Distribución de puntajes estandarizados observados en Pencil Tapping Task

La clasificación en base a puntajes T considera bajo promedio a cualquier puntaje estandarizado bajo 50, lo que deja casi al 80% de la muestra en esa condición. Nuevamente, considerar ese indicador como criterio de retraso en desarrollo no parece adecuado, dado el comportamiento observado. No obstante lo anterior, los puntajes TADI muestran diferencias estadísticamente significativas al momento de comparar los grupos de rendimiento bajo promedio, promedio y sobre promedio en PTT (tabla 18), pero con puntajes que caen siempre en la categoría de normalidad para los criterios TADI.

Tabla 18

Nivel de ajuste entre TADI Total y TADI Motricidad y Categorías PTT .

	<u>Categorías PTT</u>						F	Diferencias
	<i>Bajo Promedio</i>		<i>Promedio</i>		<i>Sobre Promedio</i>			
	N=1056	%=79.6	N=243	%=18.3	N=27	%=2		
	M (DE)	M (DE)	M (DE)					
TADI Total	49.99 (7.50)	53.79 (7.80)	53.83 (6.56)	27.36**			Bajo Promedio-Promedio**	
TADI Motricidad	49.81 (10.50)	53.50 (10.51)	52.41 (12.54)	12.51**			Bajo Promedio-Sobre Promedio* Promedio-Bajo Promedio*	

Nota: **: p menor a 0.01; *: p menor a 0.05

Al analizar el ajuste entre las categorías de interpretación TADI y los grupos de rendimiento en PTT vemos nuevamente que la presencia del grupo considerado bajo promedio en PTT (el 80% de la muestra total), va disminuyendo a medida que se aumenta el nivel de desarrollo detectado por TADI, y aumenta a medida que aumenta el nivel de riesgo (tabla 19)

Tabla 19

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a categorías PTT

Categorías TADI	Interpretación PTT			
	Bajo promedio	Promedio	Sobre Promedio	
Retraso	8 (88.9%)	1 (11.1%)	0 (0.0%)	$\chi^2_{(gl=6)}=47.85^{***}$
Riesgo	90 (92.8%)	6 (6.2%)	1 (1.0%)	
Normal	857 (81.2%)	177 (16.8%)	21 (2.0%)	
Consolidado	101 (61.2%)	59 (35.8%)	5 (3.0%)	
Nivel Riesgo TADI				
0: Sin riesgo	665 (76.6%)	184 (21.2%)	19 (2.2%)	$\chi^2_{(gl=4)}= 14.49^{***}$
1: Observación	235 (83.6%)	40 (14.2%)	6 (2.1%)	
2: Riesgo	156 (88.1%)	19 (10.7%)	2 (1.1%)	

Nota: *** = $p < .001$, ** = $p < .01$.

Los porcentajes indican la incidencia de los casos en cada fila.

La significancia de la prueba de chi cuadrado se ajustó para evitar los sesgos asociados a las celdas con pocos casos

Como es de esperarse en base a lo expuesto previamente en relación al sesgo de distribución de la prueba PTT, los índices de sensibilidad de TADI respecto de esta prueba son bajos (.09 usando la muestra completa a .49 al excluir el grupo medio), pero el comportamiento mismo de la prueba PTT hace dudar si esa prueba está captando efectivamente desarrollos alterados o no.

Comparación en relación a puntajes en prueba Head Toe Knee Shoulder (HTKS)

La prueba HTKS también se orienta a la evaluación del control inhibitorio del niño o niña (al inhibir la respuesta imitadora primaria), memoria de trabajo (al recordar correctamente las instrucciones) y atención (debe mantener el foco de atención en las instrucciones dadas por el examinador). La prueba fue desarrollada y validada con sus instrucciones en inglés y español y en muestras norteamericanas ha mostrado un valor predictivo sobre el rendimiento en niños y niñas a nivel de jardín infantil (Ponitz, McClelland, Matthews, & Morrison, 2009). Los resultados estandarizados en la muestra

chilena muestran un leve sesgo positivo, pero con una mejor distribución que lo observado en PTT (en cuanto a variabilidad de puntajes)

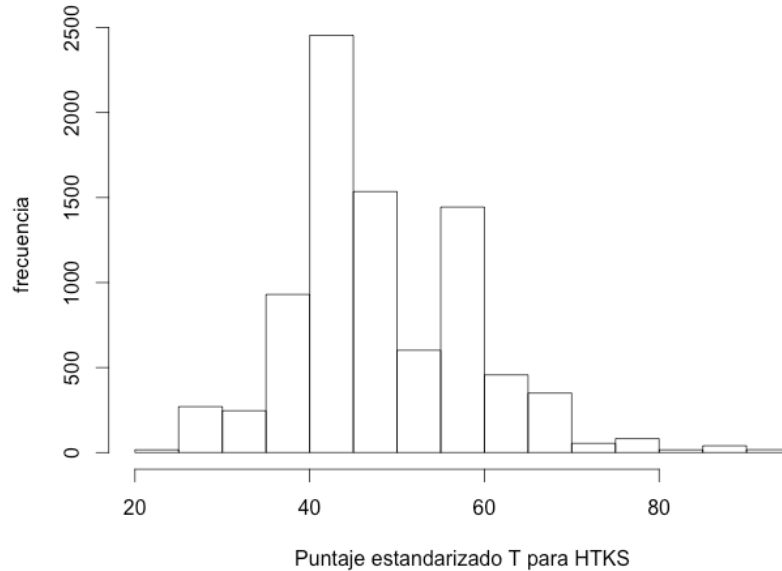


fig. 11. Distribución de puntajes estandarizados observados en Head Toe Knee Shoulder

En la interpretación usada para dichos puntajes por ELPI, se consideró bajo promedio puntajes T hasta 46 (58% de la muestra), en promedio puntajes T entre 47 y 54 (12% de la muestra) y sobre promedio puntajes 55 o superior (30% de la muestra). Al comparar los puntajes TADI obtenidos por niños y niñas que pertenecen a esas categorías en HTKS, se observan diferencias estadísticamente significativas, tanto en la dimensión de Motricidad ($F_{(2,8510)}=109.8, p<.001$) como en el puntaje TOTAL ($F_{(2,8510)}=420.5, p<.001$). tal como se espera, los puntajes TADI mejoran a medida que se pasa de la categoría bajo promedio a promedio y sobre promedio, consecutivamente (tabla 20).

Tabla 20

Nivel de ajuste entre TADI Total y TADI Motricidad y Categorías HTKS.

	<u>Categorías HTKS</u>						F (gl=2,8510)	Diferencias
	<i>Bajo Promedio</i>		<i>Promedio</i>		<i>Sobre Promedio</i>			
	N=4920	%=57.8	N=1020	%=12	N=2573	%=30.2		
	M (DE)		M (DE)		M (DE)			
TADI Total	47.25 (7.97)		49.21 (6.57)		52.48 (6.72)		420.5**	Todas las categorías**
TADI Motricidad	49.02 (10.14)		50.78 (9.49)		52.59 (9.81)		109.8**	Todas las categorías**

Nota: **: p menor a 0.01; *: p menor a 0.05

A nivel de las categorías utilizadas para la interpretación de los puntajes TADI (tabla 21), vemos que el porcentaje de casos bajo promedio en HTKS baja en forma consistente a la categoría de nivel de desarrollo TADI (desde un 97% entre niños y niñas en categoría Retraso en TADI a un 39% entre los casos de desarrollo Consolidado según TADI). Al considerar la categoría “bajo promedio” del HTKS como criterio de rezago en el desarrollo, se registran bajos niveles de sensibilidad y altos niveles de especificidad de los puntajes TADI, usando las categorías TADI Retraso + Riesgo como indicadores de rezago, en comparación con las categorías TADI Normal + Consolidado como indicativas de ausencia de déficit (Sensibilidad=.17, especificidad=.95, VPP=.83, VPN=.46). Los indicadores mejorar si se restringe a la categoría TADI de nivel de desarrollo “Consolidado” el grupo en el que se asume que no hay rezagos de desarrollo (Sensibilidad=.76, especificidad=.70, VPP=.83, VPN=.62).

Tabla 21

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a categorías HTKS

Categorías TADI	Interpretación HTKS			
	Bajo promedio	Promedio	Sobre Promedio	
Retraso	125 (96.9%)	2 (1.6%)	2 (1.6%)	$\chi^2_{(gl=6)}=440.56^{***}$
Riesgo	718 (80.6%)	70 (7.9%)	103 (11.6%)	
Normal	3823 (55.9%)	887 (13.0%)	2123 (31.1%)	
Consolidado	254 (38.5%)	61 (9.2%)	345 (52.3%)	
Nivel Riesgo TADI				
0: Sin riesgo	2603 (50.4%)	650 (12.6%)	1914 (37.0%)	$\chi^2_{(gl=4)}=436.22^{***}$
1: Observación	1088 (61.1%)	236 (13.2%)	458 (25.7%)	
2: Riesgo	1229 (78.6%)	134 (8.6%)	201 (12.9%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

Al analizar el comportamiento de la clasificación del nivel de Riesgo según TADI, vemos que la incidencia de casos bajo promedio en HTKS sube desde un 50% en el caso del grupo sin riesgo, a un 79% en el caso del grupo de Riesgo. Al igual que en otros casos, la clasificación de riesgo logra niveles moderados de capacidad diagnóstica, al considerar el nivel 0 como indicador de ausencia de desarrollo alterado y los niveles 1 y 2 como presencia de alteración del desarrollo (Sensibilidad=.47, especificidad=.71, VPP=.69, VPN=.50).

Análisis de convergencia con los resultados de pruebas asociadas a comportamiento Socioemocional

Comparación en relación a puntajes en Snack Delay Task (SDT)

La prueba SDT fue diseñada como parte de una batería orientada a la evaluación de la capacidad de control esforzado, el cual se refiere al aspecto auto regulatorio del temperamento y se observa específicamente como la habilidad de suprimir una respuesta dominante para ejecutar una subdominante (Posner & Rothbart, 2000). Se propone este rasgo como una característica de personalidad que se estabiliza cerca de los 45 meses de vida y se asocia directamente con el desarrollo emocional y social en niñez y adolescencia (Kochanska & Knaack, 2003). La presencia de puntajes alterados en esta prueba puede ser considerado entonces como un posible índice de alteraciones de desarrollo. La distribución de los puntajes estandarizados para esta prueba en la muestra evaluada se muestra a continuación

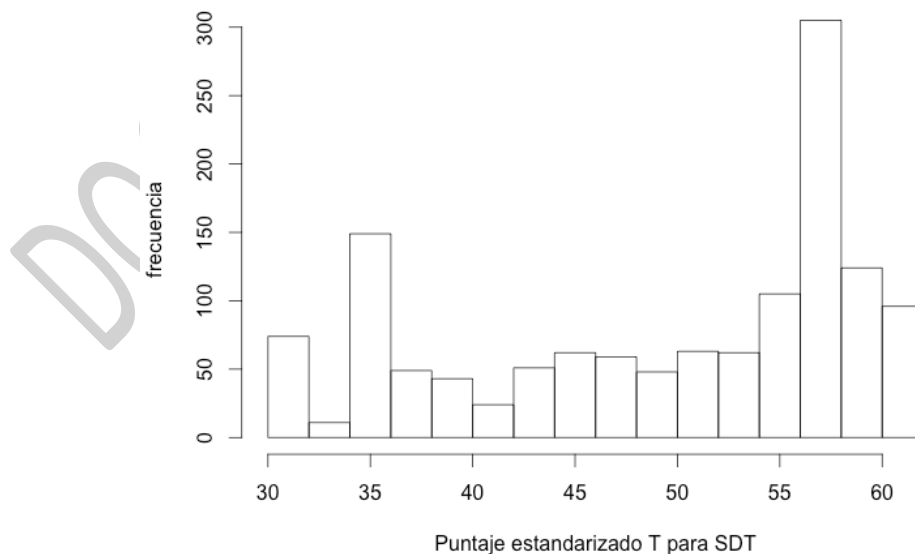


fig. 12. Distribución de puntajes estandarizados en prueba Snack Delay Task (n=1325)

En el estudio ELPI, se categorizó estos puntajes en tres niveles: “bajo promedio” para puntajes bajo 39, “promedio” para puntajes entre 39 y 56, y “sobre promedio” para puntajes desde 57 en adelante. No hay una explicación de por qué se usó esa clasificación de los puntajes t y no la tradicional en los tramos bajo 40, de 40 a 60 y 60 en adelante, respectivamente, pero se decidió mantener la clasificación entregada en la base de datos para ser fieles a la forma de procesamiento de los datos original del estudio ELPI. Usando esa clasificación como factor de agrupación, se detectan diferencias estadísticamente significativas entre todas ellas al momento de comparar los puntajes TADI Total ($F_{(2,1322)}=25.12, p<.001$), y entre la categoría sobre promedio y las otras dos al comparar los puntajes obtenidos en la escala socioemocional del TADI ($F_{(2,1322)}=19.89, p<.001$). Este comportamiento está en línea con lo esperado en cuanto a convergencia de resultados.

Tabla 22

Nivel de ajuste entre TADI Total y TADI Socioemocional y Categorías SDT.

	Categorías SDT						F	Diferencias
	Bajo Promedio		Promedio		Sobre Promedio			
	N=283	%=21.4	N=517	%=39	N=525	%=39.6		
	M (DE)		M (DE)		M (DE)			
TADI Total	48.29 (7.75)		50.62 (7.73)		52.23 (7.26)		25.12**	Todas las categorías
TADI Socioemocional	49.08 (9.63)		50.62 (9.32)		53.15 (9.03)		19.89**	Bajo Promedio-Sobre Promedio**
								Promedio-Sobre Promedio**

Nota: **: p menor a 0.01; *: p menor a 0.05

En cuanto a las categorías de clasificación de logro de desarrollo del TADI, éstas también muestran una asociación estadísticamente significativa con las categorías de

interpretación de los puntajes SDT, con una disminución de los casos con puntajes SDT bajo promedio y un aumento de los casos con puntajes SDT sobre promedio a medida que se avanza desde la categoría de retraso a la de desarrollo consolidado en TADI (tabla 23). Coherentemente, vemos que el porcentaje de casos clasificados como bajo promedio en SDT aumentan cuando el nivel de riesgo según TADI aumenta.

Tabla 23

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a categorías SDT

Categorías TADI	Interpretación SDT			
	Bajo promedio	Promedio	Sobre Promedio	
Retraso	5 (55.6%)	3 (33.3%)	1 (11.1%)	$\chi^2_{(g=6)}=41.72^{***}$
Riesgo	38 (39.2%)	39 (40.2%)	20 (20.6%)	
Normal	220 (20.9%)	413 (39.2%)	421 (39.9%)	
Consolidado	20 (12.1%)	62 (37.6%)	83 (50.3%)	
Nivel Riesgo TADI				
0: Sin riesgo	171 (19.7%)	325 (37.4%)	372 (42.9%)	$\chi^2_{(g=4)}= 29.02^{***}$
1: Observación	53 (18.9%)	116 (41.4%)	111 (39.6%)	
2: Riesgo	59 (33.3%)	76 (42.9%)	42 (23.7%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

La capacidad diagnóstica del TADI, usando como referencia la clasificación de puntajes estandarizados SDT e identificando los casos bajo promedio como indicadores de posibles alteraciones de desarrollo, logran niveles bajos de sensibilidad y altos en especificidad cuando se trabaja con la muestra completa (Sensibilidad=.15 especificidad=.94, VPP=.41, VPN=.80). La sensibilidad aumenta al excluir los casos TADI en la categoría normal y comparando solo los casos de riesgo y retraso versus aquellos con puntajes categorizados como de desarrollo consolidado (Sensibilidad=.68, especificidad=.70, VPP=.41, VPN=.88). La clasificación en base a niveles de riesgo TADI, por su parte, muestra niveles medios de sensibilidad y especificidad (Sensibilidad=.40, especificidad=.67, VPP=.25, VPN=.80).

Comparación en relación a puntajes en conductas externalizante e internalizante en el Children's Behavior Checklist de Achenbach(CBCL)

Los dominios internalizante y externalizante comprenden el conjunto de problemas conductuales usualmente registrados en niños, desde la primera infancia en adelante. En la dimensión externalizante se consideran problemas asociados al control de la conducta, incluyendo hiperactividad, impulsividad y problemas de conducta orientados al mundo externo del niño o niña. En la dimensión internalizante, por su parte, se incluyen aquellas experiencias de malestar de reacciones orientadas internamente, tales como pena, preocupación y ansiedad. El balance entre ambas dimensiones define un estilo temperamental en el niño o niña. Los problemas externalizantes suelen asociarse a rechazo de pares y bajo rendimiento académico, mientras que los internalizantes tienden a facilitar el desarrollo de problemas de salud mental posterior, principalmente en el área depresiva, y pueden asociarse a rasgos específicos de temperamento, como Rebeldía, Emocionalidad Negativa y Autorregulación (Gartstein, Putnam, & Rothbart, 2012). En definitiva, puntajes alterados en esta escala deberían asociarse en términos generales a posibles desarrollos alterados en el área emocional y social de los niños y niñas. La distribución observada de las escalas de conducta externalizante e internalizante registradas en el estudio ELPI pueden verse en la figura 13

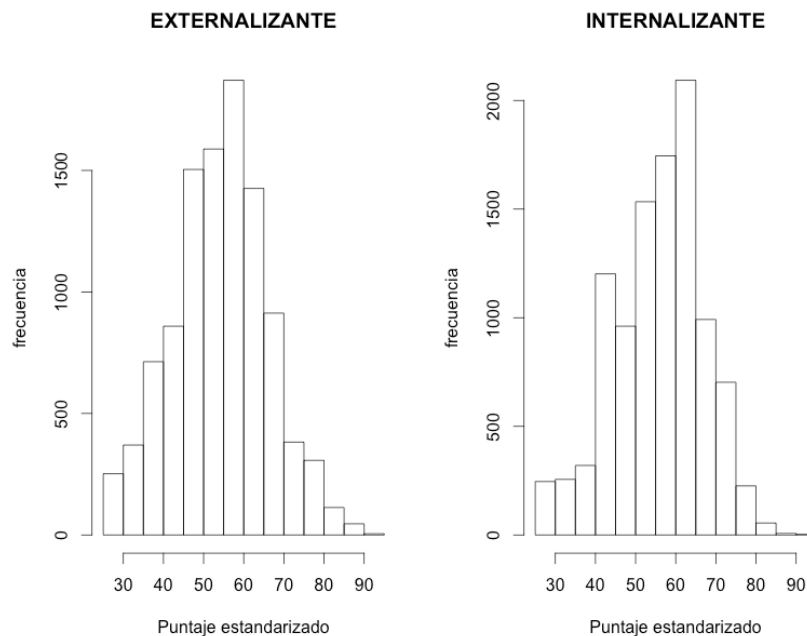


fig. 13. Distribución de puntajes estandarizados en las dimensiones Internalizante y externalizante del CBCL

Para efectos de interpretación, en el estudio ELPI se consideró puntajes normales a todos aquellos menores de 60 puntos, En riesgo, aquellos que están entre 60 y 63 y en rango clínico aquellos puntajes 64 o superior. Al usar esa clasificación como factor de agrupación se detectan diferencias estadísticamente significativas en los puntajes totales del TADI ($F_{(2,10346)}=117.8, p<.001$), con puntajes mayores en TADI en el grupo Normal según puntajes externalizantes, y menores en los grupos de Riesgo y Rango clínico. Sin embargo las diferencias observadas a nivel numérico no reflejan cambios en la interpretación de los puntajes TADI, ya que en principio las medias de todos los grupos están dentro de lo esperado en el grupo de desarrollo normal (tabla 24). Esto también ocurre en la escala Socioemocional.

Tabla 24

Nivel de ajuste entre TADI Total y TADI Socioemocional y Categorías CBCL EXTERNALIZANTE

	<u>Categorías EXTERNALIZANTE</u>			F	Diferencias
	<i>Normal</i>	<i>Riesgo</i>	<i>Rango Clínico</i>		
	N=6794 %=65.7	N=1283 %=12.4	N=2272 %=21.9		
	M (DE)	M (DE)	M (DE)		
TADI Total	50.26 (7.66)	48.5 (7.91)	47.50 (7.71)	117.8**	Todas las categorías
TADI Socioemocional	50.04 (9.77)	48.38 (10.01)	47.02 (10.06)	84.4**	Todas las categorías

Nota: **: p menor a 0.01; *: p menor a 0.05. La escala EXTERNALIZANTE se encuentra invertida.

Al analizar el comportamiento del puntaje externalizante en CBCL en relación a las categorías de logro de desarrollo en TADI, vemos que el grupo con puntajes externalizante normal aumenta de un 51.5% en el grupo de riesgo a un 77.6% en el grupo consolidado. A nivel de propiedades diagnósticas, usando el grupo normal de CBCL como representante de ausencia de desarrollo alterado y a los otros dos grupos como posibles indicadores de alteraciones de desarrollo, se registran altos niveles de sensibilidad y baja especificidad (Sensibilidad=.15, especificidad=.91, VPP=.47, VPN=.67) al compararlo con la muestra total. La sensibilidad aumenta al excluir del análisis al grupo Normal, pero disminuye la especificidad (Sensibilidad=.73, especificidad=.53, VPP=.47, VPN=.78). Al comparar las categorías CBCL con las categorías de nivel de riesgo del TADI, vemos que el grupo con un puntaje normal en la escala externalizante ocupa un 70% del grupo con nivel de riesgo 0 y baja hasta un 55% en el grupo de nivel de riesgo 2, logrando niveles moderados de capacidad diagnóstica (Sensibilidad=.46, especificidad=.65, VPP=.41, VPN=.70).

Tabla 25

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a categorías CBCL EXTERNALIZANTE

Categorías TADI	Interpretación CBCL EXTERNALIZANTE			
	Normal	Riesgo	Rango Clínico	
Retraso	71 (51.5%)	24 (17.4%)	43 (31.2%)	$\chi^2_{(gl=6)}=141.99^{***}$
Riesgo	540 (53.5%)	159 (15.8%)	310 (30.7%)	
Normal	5495 (66.1%)	1008 (12.1%)	1812 (21.8%)	
Consolidado	688 (77.6%)	92 (10.4%)	107 (12.0%)	
Nivel Riesgo TADI				
0: Sin riesgo	4445 (69.9%)	737 (11.6%)	1176 (18.5%)	$\chi^2_{(gl=4)}= 165.35^{***}$
1: Observación	1360 (61.9%)	272 (12.4%)	562 (25.6%)	
2: Riesgo	989 (55.0%)	274 (15.2%)	534 (29.7%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

Al analizar los puntajes TADI Total y en la escala socioemocional obtenidos por niños y niñas pertenecientes a las categorías Normal, Riesgo y Rango Clínico de acuerdo al puntaje internalizante se registran diferencias estadísticamente significativas ($F_{(2,10340)}=182.2$, $p < .001$ para TADI Total, $F_{(2,10340)}=144.1$, $p < .001$ para TADI Socioemocional), con una disminución de puntaje desde el rango normal al rango clínico (tabla 26).

Tabla 26

Nivel de ajuste entre TADI Total y TADI Socioemocional y Categorías CBCL INTERNALIZANTE

	Categorías INTERNALIZANTE			F	Diferencias
	Normal	Riesgo	Rango Clínico		
	N=5856 %=56.6	N=1610 %=15.6	N=2877 %=27.8		
	M (DE)	M (DE)	M (DE)		
TADI Total	50.62 (7.46)	48.98 (8.00)	47.32 (7.86)	182.2**	Todas las categorías
TADI Socioemocional	50.52 (9.60)	48.60 (10.07)	46.77 (10.07)	144.1**	Todas las categorías

Nota: **: p menor a 0.01; *: p menor a 0.05. La escala INTERNALIZANTE se encuentra invertida.

Al analizar la relación entre las categorías de interpretación de los puntajes en la dimensión internalizante del CBCL con las categorías de interpretación de logro de desarrollo TADI, vemos un aumento de la presencia de niños y niñas con puntaje internalizante normal desde un 28.9% en el grupo de retraso TADI a un 69.2% en el grupo de desarrollo consolidado según TADI. Las capacidad diagnóstica del TADI usando como referencia las categorías del puntaje Internalizante son bajos en sensibilidad y altos en especificidad (Sensibilidad=.15, especificidad=.92, VPP=.59, VPN=.59). La sensibilidad, como es esperable, aumenta al excluir del análisis a la categoría de logro de desarrollo normal (Sensibilidad =.71). En el caso de las categorías de Nivel de Riesgo, efectivamente se observa un patrón de aumento porcentual de casos de niños y niñas en Rango clínico, desde un 23.3% en el nivel 0 a 39.8% en el nivel 2 (tabla 27). La capacidad diagnóstica de este criterio, al igual que en análisis previos obtiene niveles moderados de sensibilidad y especificidad (Sensibilidad=.46, especificidad=.67, VPP=.51, VPN=.62)

Tabla 27

Ajuste de categorías TADI a la clasificación de desarrollo de acuerdo a categorías CBCL INTERNALIZANTE

Categorías TADI	Interpretación CBCL INTERNALIZANTE			
	Normal	Riesgo	Rango Clínico	
Retraso	40 (28.9%)	26 (18.8%)	72 (52.2%)	$\chi^2_{(g=6)}=209.27^{***}$
Riesgo	433 (42.9%)	173 (17.1%)	403 (39.9%)	
Normal	4769 (57.4%)	1280 (15.4%)	2260 (27.2%)	
Consolidado	614 (69.2%)	131 (14.8%)	142 (16.0%)	
Nivel Riesgo TADI				
0: Sin riesgo	3915 (61.6%)	961 (15.1%)	1478 (23.3%)	$\chi^2_{(g=4)}= 236.27^{***}$
1: Observación	1152 (52.5%)	356 (16.2%)	685 (31.2%)	
2: Riesgo	789 (43.9%)	293 (16.3%)	714 (39.8%)	

Nota: *** = $p < .001$. Los porcentajes indican la incidencia de los casos en cada fila

Análisis Preliminar de Sensibilidad y Especificidad del TADI

La tabla 28, en la página siguiente, muestra un resumen de los índices de capacidad diagnóstica calculados para TADI en comparación a distintos criterios externos, definidos de en base a las pruebas aplicadas en ELPI. El uso de diversos criterios externos permite, en parte, hacerse cargo del hecho de que no existe a nivel nacional una prueba que pueda ser considerada un *gold standard* para la evaluación del desarrollo, lo que lleva al uso de información parcial que puede confluir desde las distintas pruebas utilizadas como una forma de determinar el nivel global de los índices diagnósticos del TADI. Por otro lado implica asumir también que cada prueba entrega solo una mirada parcial al desarrollo (a excepción de Battelle), pero además requieren una revisión de la estandarización hecha, dadas las anormalidades encontradas en algunas de ellas (BSDT y PTT), en la medida en que se cuenta, gracias a ELPI, con una muestra de gran representatividad poblacional.

En términos globales, la prueba TADI muestra consistentemente altos niveles de especificidad (Media= .93), lo que indica una gran capacidad del test para detectar logros de desarrollo en las distintas áreas de evaluación testeadas. La sensibilidad, en cambio es baja al usar el criterio de los niveles de logro del desarrollo (Media =.20). Usando el criterio de Nivel de riesgo en TADI, por otro lado, se obtienen niveles moderados de sensibilidad (Media =.51), lo que estaría indicando que, en caso de querer usar TADI como prueba de tamizaje de alteraciones de desarrollo, el criterio de Nivel de riesgo (basado en la cantidad de áreas bajo promedio detectadas) sería más confiable que el uso de categorías extraídas directamente desde el puntaje total. Los valores predictivos medios positivo y negativo son también solo moderados.

El comportamiento observado debe interpretarse con cautela, ya que al estar trabajando con puntajes estandarizados en base a normas internacionales se genera una zona intermedia en la que confluyen casos que tienen alteraciones del desarrollo con aquellos que no lo tienen. De hecho, al centrar el análisis solo en las categorías de las pruebas

externas que reflejan la mayor probabilidad de desarrollo alterado versus consolidado, y usando ese mismo criterio en las categorías TADI se llega a niveles óptimos de sensibilidad y especificidad. Los resultados apuntan, en esta línea, a que se debe continuar el estudio para definir los casos que efectivamente registran un desarrollo alterado. Además, muestran como el TADI logra sus mejores resultados en la detección de logros del desarrollo. En el contexto antes señalado de que no existe un criterio definitivo de qué y cuándo nos encontramos con un rezago de desarrollo, el contar con una prueba que detecta confiablemente el logro es también un buen punto de partida en la evaluación del impacto de las distintas políticas que se implementen para promover el desarrollo saludable de niños y niñas en edad de infancia temprana en nuestro país.

Tabla 28.

Análisis de Capacidad diagnóstica de los puntajes TADI Total usando como criterio externo los resultados de las pruebas asociadas a desarrollo usadas en ELPI

	GRUPOS TADI														
	Categorías					Categorías					Nivel de Riesgo				
	(Riesgo+Retraso) vs (Normal+Consolidado)					(Riesgo + Retraso) vs. Consolidado					Sin Riesgo vs Con riesgo				
	n	Sens	Esp	VPP	VPN	n	Sens	Esp	VPP	VPN	n	Sens	Esp	VPP	VPN
1. Battelle: criterio puntajes amplio <i>bajo promedio vs promedio + sobre promedio</i>	11056	.30	.93	.45	.88	2127	.95	.60	.45	.97	11056	.65	.67	.27	.91
2. Battelle: criterio puntajes estricto <i>bajo promedio vs sobre promedio</i>	3141	.30	.99	.99	.54	940	.95	.99	.99	.93	3141	.65	.89	.88	.68
3. Battelle: Clasificación amplia <i>Puntos débiles vs (Normal + Puntos fuertes)</i>	11056	.25	.95	.61	.77	2127	.93	.68	.62	.94	11056	.60	.69	.40	.84
4. Battelle: Clasificación estricta <i>Puntos débiles vs Puntos fuertes</i>	4234	.25	.99	.99	.40	1164	.93	.99	.00	.87	4234	.60	.89	.92	.53
5. TVIP: criterio amplio <i>bajo promedio vs promedio + sobre promedio</i>	9295	.35	.93	.49	.88	1844	.94	.57	.49	.96	9295	.67	.67	.28	.91
6. TVIP: criterio estricto <i>bajo promedio vs sobre promedio</i>	3416	.35	.98	.94	.66	908	.94	.91	.94	.91	3416	.67	.80	.73	.76
7. BSDT: criterio reajustado^{b, c} <i>bajo promedio vs promedio + sobre promedio</i>	8431	.22	.90	.30	.85	1662	.90	.47	.30	.95	8431	.57	.64	.24	.88
8. PTT^b <i>bajo promedio vs promedio + sobre promedio</i>	1326	.09	.97	.92	.22	271	.49	.89	.92	.39	1326	.37	.75	.85	.23

^b Los puntajes estandarizados BDST por el sesgo detectado en la distribución observada en la muestra. Se considera bajo promedio a los casos con puntajes bajo 40 puntos y no hasta 45 como en la interpretación original

^c Pruebas que muestran alto sesgo positiva en su distribución y que por lo tanto no se consideran recomendables al momento de definir los casos en riesgo, dado que pueden reflejar problemas en el uso de normas extranjeras

Estudio validez TADI

9. HTKS

bajo promedio vs promedio + sobre promedio

8513 .17 .95 .83 .46 1680 .77 .69 .83 .62 8513 .47 .71 .69 .50

10. SDT

bajo promedio vs promedio + sobre promedio

1325 .15 .94 .41 .80 271 .68 .70 .41 .88 1325 .40 .67 .25 .80

11. CBCL Externalizante

Normal vs Riesgo + rango Clínico

10349 .15 .91 .47 .67 2034 .73 .53 .47 .78 10349 .46 .65 .41 .70

12. CBCL Internalizante

Normal vs Riesgo + rango Clínico

10343 .15 .92 .59 .59 2034 .71 .56 .59 .69 10343 .46 .67 .51 .62

Promedio pruebas seleccionadas (3,5,9,10,11,12)

.20 .93 .57 .69 .79 .62 .57 .81 .51 .68 .42 .72

DOCUMENTO DE TRABAJO

Análisis de Convergencia con otras variables medidas en ELPI, en las que se ha encontrado evidencia científica en relación a su influencia sobre los logros del desarrollo.

En el contexto de ausencia de un *Gold* estándar único que defina el logro de desarrollo en los niños y niñas en forma integral, una fuente de evidencia adicional para la validez de un test orientado a medir dicho constructo es el análisis del impacto de variables de riesgo asociadas a factores contextuales y médicos. El estudio de estas variables, que generan inequidades intra e inter poblacionales, se expresan por medios de experiencias tempranas consideradas como adversas y que afectan el desarrollo temprano y posterior (Walker et al., 2011). Algunos de los indicadores identificados por la literatura fueron incluidos dentro del estudio ELPI, ya sea en forma directa, como parte de la caracterización muestral, o dentro de las variables incluidas en el inventario Home Observation for Measurement of the Environment: HOME (Bradley, Caldwell, Rock, Hamrick, & Harris, 1988; Totsika & Sylva, 2009).

Las variables a evaluar, disponibles en la base de datos del estudio ELPI, que han mostrado evidencia empírica de impacto sobre el desarrollo en la literatura, son las siguientes

- Factores Económicos, operacionalizado por medio del quintil de ingreso del grupo familiar al que pertenece el niño o niña evaluado. Se considera riesgosa la pertenencia a grupos de menores ingresos. Para efectos del análisis, consideraremos en riesgo aquellos niños que pertenecen a hogares con ingreso per cápita mensual menor a \$88.500, correspondiente a los dos primeros quintiles poblacionales de ingreso.
- Factores del desarrollo biológico del niño y que pueden resultar en retrasos del desarrollo psicomotor, operacionalizado en las variables
 - o Retraso en peso de acuerdo a la edad
 - o Retraso en la talla respecto de la edad
 - o Puntaje bajo 8 puntos en el segundo APGAR al momento de nacer
 - o Presencia de problemas de salud perinatal mostrados por los niños

- Factores asociados al cuidador principal, en el caso de la muestra, la madre, específicamente
 - o Máxima Escolaridad alcanzada
 - o Presencia de Problemas Psiquiátricos
 - o Bajo Rendimiento intelectual, en base a las escalas de dígitos y vocabulario del WAIS
- Factores asociados al ambiente del hogar que pueden afectar el desarrollo psicomotor, en base a los indicadores HOME incluidos en ELPI en todas las edades medidas
 - o Respuesta Emocional y Verbal de la Madre o Cuidador primario, asociado a las interacciones comunicativas y afectivas entre el cuidador y el niño
 - o Evitación de Restricción y Castigo, asociado a disciplina.
 - o Provisión de material de juego
 - o Involucración parental con el niño (interacción física)

El impacto de cada una de estas variables sobre los puntajes TADI se muestra en la tabla 29. En cada caso la variable estudiada fue codificada en el sentido de si el valor registrado para cada caso estaba en el rango considerado riesgoso o no. Así, se puede determinar la presencia relativa de cada factor de riesgo, en base al porcentaje de casos en condición de riesgo de cada factor y, además, se puede generar una medida estándar de impacto, en base al coeficiente de regresión asociado a cada factor, el que indica la cantidad media de disminución de puntaje TADI asociado a la presencia de cada riesgo.

Tabla 29.

Incidencia e impacto de riesgos asociados a factores sociodemográficos y a variables propias del ambiente del hogar que afectan el desarrollo infantil en primera infancia.

Factor de Riesgo	Presencia Riesgo		Impacto
	f	(%)	b
RIESGO SOCIOECONÓMICO			
<i>Ingreso per cápita (n=7609)</i>			
Ingreso per Cápita mensual < \$188500	3161	(41.54%)	-1.35 (p<.001)
RIESGO DESARROLLO BIOLÓGICO NIÑO(A)			
<i>Relación peso /edad (n= 10605)</i>			
Presencia de retraso en peso	62	(0.58%)	-5.03 (p<.001)
<i>Relación talla /edad (n= 10722)</i>			
Presencia de retraso en talla	383	(3.57%)	-1.84 (p<.001)
<i>APGAR a los 5 minutos de nacer (n= 1784)</i>			
APGAR < 8 puntos	21	(1.18%)	-4.60 (p=.003)
<i>Número de antecedentes médicos perinatales (n= 2614)</i>			
Presencia de al menos un antecedente médico perinatal	983	(37.61%)	-0.33 (p=.239)
RIESGOS ASOCIADOS A LA MADRE			
<i>Escolaridad de la Madre (n= 11063)</i>			
Baja escolaridad Materna	3633	(32.84%)	-2.98 (p<.001)
<i>Antecedentes Psiquiátricos en la Madre (n=2614)</i>			
Presencia de Antecedentes Psiquiátricos	312	(11.94%)	-0.30 (p=.481)
<i>Rendimiento Intelectual (dígitos) (n= 2722)</i>			
Rendimiento Moderada o Fuertemente Disminuido	703	(25.83%)	-1.84 (p<.001)

Rendimiento Intelectual (vocabulario) (n= 2722)

Rendimiento Moderada o Fuertemente Disminuido	918	(33.73%)	-0.86 (p=.003)
---	-----	----------	----------------

RIESGOS DEL AMBIENTE DE CRIANZA (HOME) (n=11190)

Baja Respuesta Emocional Cuidador	683	(6.10%)	-3.27 (p<.001)
Baja Evitación del Castigo	1670	(14.92%)	0.06 (p=.772)
Escasez de Materiales de Juego/Aprendizaje	5888	(52.62%)	-3.29 (p<.001)
Bajo Compromiso Parental	368	(3.29%)	-1.78 (p<.001)

Nota: n= indica el número total de casos en los que se registró la presencia o ausencia de cada uno de los factores de riesgo

f= indica el número de casos que reportan la presencia del riesgo entre los casos registrados

% = indica la presencia porcentual del riesgo

b = coeficiente de regresión. Indica la variación media esperada en el puntaje TADI Total asociada a la presencia del factor de riesgo identificado.

Tal como puede observarse, los puntajes TADI muestran variaciones estadísticamente significativas asociadas a la presencia de los distintos factores de riesgo identificados, con una disminución de puntaje total del test cada vez que alguno de estos factores de riesgo está presente.

Dado que todos los factores de riesgo fueron dicotomizados en las categorías presencia/ausencia del riesgo, la comparación entre los coeficientes de regresión permite determinar aquellos factores ambientales que tienen mayor impacto sobre los puntajes TADI. En esta línea de pensamiento, consideramos como riesgos de alto impacto a aquellos factores cuya presencia puede generar una disminución de al menos 3 puntos en el puntaje estándar total de TADI, que equivalen a un tercio de desviación estándar, lo que puede llegar a significar una disminución percentil de hasta 8 puntos para la persona evaluada. En otras palabras, en este caso la presencia del riesgo puede hacer que una persona pase del percentil 50 al percentil 42 en su evaluación de desarrollo, esperándose un mayor retroceso a medida que el impacto del factor de riesgo es mayor. Asimismo, consideramos como riesgos de impacto medio a aquellos factores cuya presencia puede generar disminuciones medias de entre 1 y 3 puntos TADI (que pueden implicar retrocesos de entre 4 y 8 puntos percentil), y como riesgos de impacto bajo aquellos cuya presencia puede generar disminuciones de menos de 1 punto en el puntaje TADI. Se consideran como factores sin impacto a aquellos que no registran un coeficiente de regresión estadísticamente significativo asociado a su presencia.

Revisando los resultados en la tabla 29, destacan como riesgos de alto impacto los factores de riesgo biológico, como retraso en el peso respecto de la talla ($b=-5.03$, $p<.001$) y presencia de un APGAR alterado a los 5 minutos del nacimiento ($b= -4.60$, $p<.001$). También se registraron efectos significativos del ambiente de crianza, en particular la escasez de materiales de juego o Aprendizaje ($b=-3.29$, $p<.001$) y la baja respuesta emocional del cuidador principal ($b=-3.27$, $p<.001$). Asociado a lo anterior, en características de la madre, el mayor efecto sobre los puntajes TADI se deriva de la presencia de baja escolaridad materna ($b=-2.98$, $p<.001$). Los efectos mencionados en el área biológica al nacer concuerdan con aquellos aspectos que han mostrado efectos significativos sobre el desarrollo posterior en niños y niñas, seguidos desde el nacimiento hasta los 7 años, particularmente en lo que respecta a alteraciones en el APGAR a los 5 minutos de nacer (Gampel & Nomura, 2014). Los efectos de baja escolaridad materna y poco material estimulante ambiental apuntan a la presencia de un ambiente de estimulación empobrecido para el logro de metas en el desarrollo temprano, o que unido a una baja respuesta emocional del cuidador generaría un entorno de baja estimulación global al desarrollo, que reduce las potencialidades de desarrollo del niño o niña criado(a) en dicho ambiente (Walker et al., 2011).

En cuanto a riesgos de impacto medio, se identifican en este tramo, en el área de factores de riesgo biológico, al Retraso en la talla en relación a la edad ($b=-1.84$, $p<.001$), el cual se encuentra asociado al retraso en peso observado ($\chi^2=91.04$, $gl=1$, $p<.001$, calculado con corrección de Yates), y que puede estar indicando la presencia factores de déficit nutricional o malformación constitucional que potencialmente puede afectar el desarrollo infantil. En el área de características de la madre aparece el efecto del Rendimiento intelectual disminuido en dígitos ($b=-1.84$, $p<.001$), el cual se asocia a memoria de trabajo y capacidad de atención y que se asocia a la capacidad intelectual general mostrado por el cuidador. En características del ambiente aparece el impacto del Bajo compromiso parental ($b=-1.78$, $p<.001$), el que implica poca interacción y contacto visual entre adultos y el niño(a) evaluado, condición que impactaría las posibilidades de desarrollo cognitivo, social y preescolar de los niños y niñas viviendo en dicha condición (DeLoatche, Bradley-Klug, Ogg, Kromrey, & Sundman-Wheat, 2014). En general, el impacto de la involucración parental se hace más visible en la época escolar, de modo que encontrar un impacto medio en la muestra estudiada (hasta los 6 años), es esperable en la medida en que aún no entran en un régimen escolar regular. Finalmente, se encontró un efecto medio del factor de riesgo asociado a bajo ingreso per cápita familiar ($b=-1.35$,

$p < .001$). Si bien la consideración del ingreso económico como factor de riesgo al desarrollo ha sido bien documentado en la literatura, el impacto detectado no pareciera en principio ser tan importante. Sin embargo, al hacer el análisis detallado por quintil de ingreso vemos un efecto de disminución progresiva de los puntajes TADI, los que se acentúan al tomar como punto de comparación los resultados de quienes pertenecen al quintil de ingreso superior (desde \$200000 a \$2750000 per cápita mensual), tal como se muestra en la tabla 30. Considerando que el puntaje medio TADI del quintil de mayor ingreso es de 51.6 puntos, vemos que la diferencia registrada asociada a pertenecer a quintiles de menor ingreso es significativa en todos los casos, y tiene su impacto mayor al compararlo con el quintil de menor ingreso. Considerando que el rango de ingresos del quintil superior es muy superior al de cualquiera de los otros quintiles, el hecho de que sea el grupo que genera mayores diferencias con todo el resto en cuanto a su impacto en el desarrollo reportado en base a puntajes TADI puede estar reflejando efectos asociados a ambientes de mayor protección en todos los factores de riesgo, en comparación al resto de la población, siendo un tema que requiere mayor profundización en estudios posteriores. En el caso de este reporte, el resultado encontrado es concordante con el efecto de la pertenencia a grupos de bajos ingresos sobre el desarrollo temprano y la evolución posterior en el sistema educativo (Schoon, Hope, Ross, & Duckworth, 2010; Waldfogel & Washbrook, 2010).

Tabla 30.

Efectos sobre el resultado TADI asociado a la pertenencia a los distintos quintiles de ingreso, tomando como base el Quintil de mayor Ingreso

Ingreso per cápita reportado	Efecto sobre Puntaje TADI (b)
Cuarto Quintil: \$ 120000 a \$19250	-1.50 ($p < .001$)
Tercer Quintil: \$88500 a \$ 119800	-1.92 ($p < .001$)
Segundo Quintil: \$59000 a \$88400	-1.92 ($p < .001$)
Primer Quintil: \$0 a \$58889	-3.03 ($p < .001$)

Nota: Se toma como referencia al Quintil de mayor ingreso, que obtiene una media en TADI de 51.6 puntos

Finalmente, el factor de inteligencia verbal de la madre mostro efectos estadísticamente significativos, pero bajos sobre los puntajes TADI ($b = -.86$, $p = .003$). La inteligencia en el área verbal fue

medida usando la prueba de Vocabulario de la escala de Wechsler, la que se centra principalmente en el conocimiento de conceptos, de modo que sus efectos probablemente pueden hacerse más visibles en la edad escolar, y pueden también ser parcialmente compensados por la interacción que los niños tengan en su relación con el sistema preescolar. En cualquier caso, el vocabulario materno ha mostrado ser un factor influyente en el desarrollo lingüístico de los niños y niñas, mediando incluso el impacto de las diferencias socioeconómicas (Hoff, 2003), de modo que el que aparezca como un factor de influencia asociado al logro de desarrollo es también coherente con la evidencia empírica.

En relación a factores que no registraron efectos significativos, la presencia de antecedentes médicos perinatales ($b=-0,33$, $p=.239$), presentes en el 37.61% de los 2614 casos que entregaron los antecedentes correspondientes y que puede corresponder a antecedentes de preclampsia (2.83% de los casos), colestasia (1.45%), infecciones urinarias (10.79%), hemorragias (2.03%), hipertensión (5.89%), placenta previa (2.98%), diabetes gestacional (8.53%), anemia (9.87%), toxoplasmosis (0.08%), rubeola (0.00%), sífilis (0.15%) u otra condición médica (3.79%). Se consideraba riesgosa la presencia de uno o más de estos antecedentes. Ninguno de los antecedentes por separado mostró tampoco un impacto significativo sobre los puntajes TADI. El resultado no es sorprendente, en todo caso, dado que los antecedentes de riesgo asociados a salud aún no han demostrado sus efectos sobre el desarrollo infantil temprano (Walker et al., 2011), incluso en condiciones de baja nutrición, anemia, escasez de nutrientes como hierro y ácido fólico, y baja concentración de omega 3 en la alimentación, todas condiciones que han en el contexto chileno son relativamente escasas. La presencia de antecedentes psiquiátricos en la madre ($b=-.30$, $p=.481$), presente en el 11.94% de los 2614 casos que entregaron información de esa área. Se consideró como antecedente de riesgo la presencia de uno o más de los siguientes antecedentes: Depresión (10.18%), Trastorno Bipolar (0.42%), trastorno de Ansiedad (0,92%), trastorno obsesivo compulsivo (0,19%), Fobias (0,27%), Trastorno de Pánico (0,69%), Trastorno de estrés postraumático (0,46%) u otros trastornos (0.27%). En este caso, el antecedente depresivo es el más preponderante, y por sí solo tampoco muestra efectos significativos sobre los puntaje TADI. En este caso, si bien el antecedente de depresión materna suele considerarse como un factor de riesgo sobre el desarrollo infantil en primaria infancia, el mecanismo de influencia actúa primariamente por los efectos de la depresión sobre sensibilidad reducida de la madre frente a las necesidades del niño(a), pero sin implicar un efecto directo o lineal. Dado que dicha influencia se registró directamente mediante el inventario HOME, y que dicho índice

efectivamente muestra los efectos sobre el desarrollo reportados en la literatura, podemos pensar que el mero antecedente de depresión no puede considerarse como razón suficiente para influir negativamente sobre el desarrollo de los niños y niñas estudiados.

DOCUMENTO DE TRABAJO

Síntesis de Resultados

El presente informe reportó diferentes análisis orientados a chequear la validez de la prueba TADI como herramienta de evaluación del desarrollo en infancia temprana.

A partir de los análisis realizados se concluye que el test reporta buenos indicadores de validez, incluyendo un comportamiento coherente tanto en su estructura interna como en relación con otras pruebas orientadas a evaluar los logros de desarrollo tanto a nivel parcial como global.

El test entrega además, distintas posibilidades de interpretación, que pueden ser usadas dependiendo del contexto en que sea aplicado. Estas posibilidades son:

- a) Usando el puntaje estándar directo obtenido por el niño o niña, que permite hacer un seguimiento de su evolución a través de las distintas etapas del crecimiento
- b) En base a la categorización del nivel de logro de las tareas del desarrollo, en categorías que se derivan directamente de los puntajes estandarizados y que permiten determinar en forma precisa los grupos de riesgo en cuanto a logro de tareas de desarrollo, mostrando los mejores niveles de especificidad al ser comparado con los resultados obtenidos de otras pruebas
- c) En base al nivel de riesgo, derivado del número de áreas consideradas en riesgo del desarrollo y que permite determinar casos que requieren mayor observación, en la medida en que es la estrategia de categorización que mostró el mejor nivel medio de sensibilidad en relación a los resultados entregados por otras pruebas incorporadas en el TADI.

En relación a la validez convergente, los resultados del TADI se encuentra alineados con los resultados obtenidos en todas las pruebas aplicadas en el estudio ELPI asociadas a evaluación de desarrollo y muestra niveles adecuados de ajuste al comparar las categorías diagnósticas usadas en cada una de ellas. En este punto, es importante destacar el hecho de que ninguna de las pruebas utilizadas puede ser considerada como un *gold* estándar en relación a la evaluación de desarrollo en la forma en la que lo evalúa TADI, ya que no cuentan con validaciones actualizadas en nuestro país, o que miden solo aspectos específicos que se relacionan en forma indirecta con los logros del desarrollo, mostrando incluso comportamientos estadísticos que hacen pensar en la necesaria revisión de la utilidad de los puntajes estandarizados utilizados y su interpretación.

Por otro lado, el comportamiento del TADI es coherente en relación con los efectos de factores sociodemográficos y de entorno de crianza que han sido evidenciados como riesgosos para el desarrollo psicomotor, particularmente en aspectos que pueden asociarse a nutrición general y a ambiente de estimulación y atención a las necesidades de él o la niño o niña evaluada. En la medida en que TADI ha sido una prueba desarrollada y validada enfocándose en estándares locales, incluyendo una amplia muestra nacional y considerando factores sociodemográficos y étnicos propios de nuestra cultura, creemos que se configura como una prueba que puede ser utilizada confiablemente en la evaluación de los logros de desarrollo en infancia temprana, y también como estrategia de detección de posibles casos en riesgo, complementando la información de medidas generales de tamizaje y apoyando la labor diagnóstica de dichas casos.

DOCUMENTO DE TRABAJO

Referencias

- Bedregal, P. (2014). Hacia la renovación en las políticas de infancia en Chile. *Rev. Chilena de Pediatría*, 85(1), 7–11. <http://doi.org/10.4067/S0370-41062014000100001>
- Bradley, R. H., Caldwell, B. M., Rock, S. L., Hamrick, H. M., & Harris, P. (1988). Home Observation for Measurement of the Environment: Development of a Home Inventory for use with families having children 6 to 10 years old. *Contemporary Educational Psychology*, 13(1), 58–71. [http://doi.org/10.1016/0361-476X\(88\)90006-9](http://doi.org/10.1016/0361-476X(88)90006-9)
- Centro Micro Datos Departamento de Economía. (2012). *Anexo Metodológico Instrumentos Utilizados para la Evaluación. Encuesta longitudinal de la Primera infancia*. Santiago. Retrieved from http://www.dev-out.cl/sites/default/files/Anexo_Metodologico_Instrumentos_Psicologicos.pdf
- DeLoatche, K. J., Bradley-Klug, K. L., Ogg, J., Kromrey, J. D., & Sundman-Wheat, A. N. (2014). Increasing Parent Involvement Among Head Start Families: A Randomized Control Group Study. *Early Childhood Education Journal*, 43(4), 271–279. <http://doi.org/10.1007/s10643-014-0660-7>
- Dunn, L., Padilla, E., Lugo, D., & Dunn, L. (1986). *Test de Vocabulario en Imágenes Peabody: TVIP: Adaptación Hispanoamericana (Peabody Picture Vocabulary Test: PPVT: Hispanic-American Adaptation)*. Circle Pines: American Guidance Service (AGS), Inc.
- Echeverría, M., Herrera, M. olivia, & Segure, J. T. (2002). *TEST DE VOCABULARIO EN IMÁGENES (TEVI-R)*. Concepción: Universidad de Concepción.
- Gampel, S., & Nomura, Y. (2014). Short and Long-Term Effects of Compromised Birth Weight, Head Circumference, and Apgar Scores on Neuropsychological Development. *Journal of Psychological Abnormalities in Children*, 03(3). <http://doi.org/10.4172/2329-9525.1000127>
- Gartstein, M. A., Putnam, S. P., & Rothbart, M. K. (2012). ETIOLOGY OF PRESCHOOL BEHAVIOR PROBLEMS: CONTRIBUTIONS OF TEMPERAMENT ATTRIBUTES IN EARLY CHILDHOOD. *Infant Mental Health Journal*, 33(2), 197–211. <http://doi.org/10.1002/imhj>.
- Hoff, E. (2003). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*, 74(5), 1368–1378. <http://doi.org/10.1111/1467-8624.00612>

- Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., Zelazo, P. D., Wendy, S., Lee, C., & Lee, S. C. (2010). Developmental Neuropsychology Assessment of Hot and Cool Executive Function in Young Children : Age-Related Changes and Individual Differences. *Developmental Neuropsychology*, 28(October 2011), 37–41. <http://doi.org/10.1207/s15326942dn2802>
- Kochanska, G., & Knaack, A. (2003). Effortful Control as a Personality Characteristic of Young Children : Antecedents , Correlates , and Consequences. *Journal of Personality*, 71(6), 1087–1112. <http://doi.org/10.1111/1467-6494.7106008>
- Lecannelier, F., Pérez Ewert, J. C., Groissman, S., Gallardo, D., Bardet, A. M., Bascuñan, A., & Rodríguez, J. (2014). Validación del Inventario de Conductas Infantiles para niños de entre 1½-5 años (CBCL 1½-5) en la Ciudad de Santiago. *Universitas Psychologica*, 13(2), 491–500. <http://doi.org/10.11144/Javeriana.UPSY13-2.vici>
- Medwave Estudios Limitadas, & CIGES. (2013). *Evaluación costo efectividad de modalidades de apoyo al desarrollo infantil del sistema de protección a la infancia.*
- Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, 45(3), 605–19. <http://doi.org/10.1037/a0015365>
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, 12, 427–441. Retrieved from http://journals.cambridge.org/abstract_S0954579400003096
- R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.r-project.org/>
- Rosseel, Y. (2012). lavaan : An R Package for Structural Equation. *Journal of Statistical Software*, 48(2), 1–36. Retrieved from <http://www.jstatsoft.org/v48/i02/>
- Schoon, I., Hope, S., Ross, A., & Duckworth, K. (2010). Family hardship and children’s development: the early years. *Longitudinal and Life Course Studies*, 1(3), 209–222. <http://doi.org/10.14301/llcs.v1i3.109>
- Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22(2), 173–187. <http://doi.org/10.1016/j.ecresq.2007.01.002>

- Strasser, K., Larraín, A., López de Lérída, S., & Lissi, M. R. (2010). La Comprensión Narrativa en Edad Preescolar: Un Instrumento para su Medición. *Psyche (Santiago)*, 19(1), 75–87.
<http://doi.org/10.4067/S0718-22282010000100006>
- Totsika, V., & Sylva, K. (2009). The Home Observation for Measurement of the Environment revisited. *Child and Adolescent Mental Health*, 9(1), 25–35. <http://doi.org/10.1046/j.1475-357X.2003.00073.x>
- Waldfoegel, J., & Washbrook, E. (2010). *Low income and early cognitive development in the UK : A report for the Sutton Trust* . Retrieved from http://www.suttontrust.com/reports/Sutton_Trust_Cognitive_Report.pdf
- Walker, S. P., Wachs, T. D., Grantham-McGregor, S., Black, M. M., Nelson, C. A., Huffman, S. L., ... Richter, L. (2011). Inequality in early childhood: risk and protective factors for early child development. *The Lancet*, 378(9799), 1325–1338. [http://doi.org/10.1016/S0140-6736\(11\)60555-2](http://doi.org/10.1016/S0140-6736(11)60555-2)
- Willoughby, M. T., Kupersmidt, J. B., & Voegler-Lee, M. E. (2012). Is preschool executive function causally related to academic achievement? *Child Neuropsychology*, 18(1), 79–91.
<http://doi.org/10.1080/09297049.2011.578572>