

Propiedades del Test de Aprendizaje y Desarrollo Infantil “TADI”, en el contexto de las políticas de validación de escalas para niños y niñas menores de 6 años en Chile y Latinoamérica

Properties of the Test of Early Learning and Development, second edition “TADI”, in the context of scale validation policies for children under 6 years of age in Chile and Latin America

Luisa Schonhaut^a, Marta Edwards^{b,e}, Marcela Pardo^{c,f}, Antonia Valdés^{d,e}

^aClínica Alemana, Facultad de Medicina Universidad del Desarrollo. Santiago, Chile.

^bCentro de Estudios de Desarrollo y Estimulación Psicosocial (CEDEP). Chile.

^cInstituto de Estudios Avanzados en Educación y Centro de Investigación Avanzada en Educación, Universidad de Chile. Santiago, Chile.

^dUniversidad del Desarrollo, Facultad de Educación. Santiago, Chile.

^ePsicóloga.

^fAntropóloga.

Recibido: 7 de febrero de 2024; Aceptado: 25 de abril de 2024

¿Qué se sabe del tema que trata este estudio?

La detección temprana junto con la intervención oportuna de los niños/as con rezagos en el desarrollo, tiene gran impacto para el futuro del infante, su familia y sociedad. En respuesta a la necesidad de contar con un nuevo instrumento construido en Chile para la evaluación del desarrollo infantil temprano a gran escala, se desarrolló el Test de Aprendizaje y Desarrollo Infantil (TADI), el que fue actualizado y robustecido como TADI-2.

¿Qué aporta este estudio a lo ya conocido?

Se describen los principales requerimientos metodológicos que deben ser considerados en la construcción de un instrumento de medición en salud. Se presenta la escala TADI creada en Chile, profundizando en su reciente actualización (TADI-2) y se analiza su construcción, validez y confiabilidad a través del cumplimiento de criterios de validación.

Resumen

En Latinoamérica se han construido distintas escalas para la evaluación del Desarrollo Infantil Temprano (DIT), la mayoría de aplicación directa. En Chile, en los años 70-80 se crearon la Escala de Evaluación del Desarrollo Psicomotor (EEDP) y el Test de Desarrollo Psicomotor (TEPSI), utilizadas hasta la actualidad en distintos países hispanoparlantes, sin que hayan sido actualizadas ni revalidadas. En respuesta a la necesidad de contar con un nuevo instrumento construido en Chile para la evaluación del DIT a gran escala, entre los años 2009 y 2012 se construyó y estandarizó el Test de

Palabras clave:

Tamizaje del Desarrollo; Psicometría; Evaluación del Desarrollo; Desarrollo Infantil Temprano; Desarrollo Psicomotor; TADI; Validez

Correspondencia:
Luisa Schonhaut
lschonhaut@alemana.cl

Editado por:
Francisco Cano Schuffenegger

Cómo citar este artículo: Andes pediater. 2024;95(4):353-363. DOI: 10.32641/andespediatr.v95i4.5149

Aprendizaje y Desarrollo Infantil (TADI). Este fue actualizado y robustecido entre 2018 y 2023 como TADI-2. **Objetivo:** Describir los principales requerimientos metodológicos considerados en la construcción de un instrumento de medición en salud y, en ese contexto, analizar la validez del TADI-2. **Desarrollo:** Destaca el riguroso proceso de validación del TADI-2, cumpliendo con los estándares de selección de contenido, estandarización, confiabilidad, validez en distintos entornos socioculturales y étnicos, bajo costo, fácil aplicación y uso abierto. La sensibilidad para el puntaje global en los niños/as de 4 y 5 años fue muy buena y, en los menores, moderada. **Conclusiones:** La evidencia recabada muestra que el TADI-2 puede ser implementado para el tamizaje del DIT a través de aplicación directa, de los niños/as menores de 6 años de Chile y resto del continente. Se recomienda continuar estudiando el instrumento en muestras más amplias para conocer la consistencia de los resultados en estudios complementarios y poblacionales.

Abstract

In Latin America, scales have been built to assess Early Child Development (EChD), most of them for direct evaluations. In Chile, the Psychomotor Development Evaluation Scale (EEDP) and the Test of Psychomotor Development (TEPSI) were developed in the 70's and 80's and are still used in different Spanish-speaking countries, but have not been updated or revalidated. In response to the need for a new instrument built in Chile for the evaluation of EChD on a large scale, the Test of Early Learning and Development (TADI-2, for its acronym in Spanish) was built and standardized between 2009 and 2012. This was updated and strengthened between 2018 and 2023 as TADI-2. **Objective:** To describe the main methodological requirements considered in the construction of a health measurement instrument and, in this context, to analyze the validity of the TADI-2. **Development:** The rigorous validation process of the TADI-2 stands out, complying with the standards of content selection, standardization, reliability, validity in different sociocultural and ethnic environments, low cost, easy application and open use. Sensitivity for the global score in 4- and 5-year-olds was very good and, in younger children, moderate. **Conclusions:** The gathered evidence shows that the TADI-2 may be implemented for EChD screening through direct assessment of children under 6 years of age in Chile and the rest of the continent. It is recommended to continue studying the instrument in larger samples to know the consistency of the results in complementary and population-based studies.

Keywords:

Developmental Screening; Psychometrics; Developmental Assessment; Early Childhood Development; Psychomotor Development; TADI; Validity

Introducción

La transición epidemiológica y demográfica ha modificado las prioridades inherentes a la atención sanitaria infantil. Así, por ejemplo, cuando la mortalidad infantil era elevada, el objetivo primordial era superar esa amenaza y vencer las enfermedades infectocontagiosas y la desnutrición. En la medida en que aumentó la sobrevivencia frente a condiciones anteriormente de mal pronóstico, el desarrollo infantil temprano (DIT) fue adquiriendo prioridad, y, cuando avanzamos en los determinantes de desarrollo, se sumaron nuevos desafíos, relacionados con el desarrollo socioemocional y bienestar infantil^{1,2}.

El DIT es el proceso mediante el cual los niños/as van cambiando a lo largo de sus primeros años de vida, en sus habilidades motoras, cognitivas, lingüísticas y socioemocionales, producto de la interacción compleja de múltiples factores biológicos, psicológicos y ambientales. Las distintas dimensiones ejercen gran influencia mutua, lo que otorga a este proceso su carácter holístico.

Los primeros 5 años de vida corresponden al periodo crítico para la fundación del desarrollo, la autonomía y la capacidad cognitiva. A nivel individual se sientan las bases de los logros educativos y profesionales, mientras que a nivel poblacional es uno de los determinantes del capital humano. Se ha demostrado que la detección temprana junto con la intervención oportuna de los niños/as con rezagos en el desarrollo, tiene gran impacto para el futuro del infante, su familia y la sociedad³.

Se estima que globalmente en 1 de cada 6 niños/as se diagnostica alguna discapacidad relacionada con el DIT^{4,5}. Al aplicar pruebas de tamizaje, cuyo objetivo es identificar a los niños/as en riesgo de dificultades, esta tasa aumenta a 1 en cada 4 a 5 niños/as⁶. En Chile, en la última década, al emplear escalas basadas en el reporte de los padres o cuidadores principales en una muestra normativa de lactantes de distintos niveles socioeconómicos se reportó una tasa de 17%⁷, ascendiendo a 28% al aplicar encuestas nacionales⁸. Similares son los resultados al utilizar pruebas de aplicación directa, que, además, dan muestra de una disparidad socioeco-

nómica, en que 14% de los preescolares atendidos en el sector privado y 30% de los niños/as atendidos en el sector público tienen déficit del desarrollo⁹.

En Chile existen políticas de evaluación y estimulación del DIT desde hace más de 50 años. Con el objeto de optimizar la pesquisa e implementar medidas efectivas de promoción del desarrollo y la salud mental infantil, el año 2008 se creó un Sistema de Protección Integral a la Infancia “Chile Crece Contigo” y se revisó la pertinencia de las escalas en uso en Atención Primaria de Salud (APS). En la Estrategia Nacional de Salud (ENS) para la década 2021-2030 se definió como meta “disminuir la prevalencia de niños/as que no alcanzan su desarrollo integral acorde a su potencialidad”, para lo cual es fundamental contar con instrumentos válidos y confiables¹⁰. En ese contexto, entre los años 2009 y 2012 se construyó, validó y estandarizó el Test de Aprendizaje y Desarrollo Infantil (TADI), que luego se robusteció en sus propiedades psicométricas y se actualizó entre los años 2018 y 2023 como TADI-2, gracias al apoyo de dos proyectos FONDEF (D07I1029 e IT18I0005)^{11,12}.

El objetivo del presente manuscrito es describir los principales requerimientos metodológicos que deben ser considerados en la construcción de un instrumento de medición en salud y, en ese contexto, revisar el proceso de construcción del TADI-2.

Medición del desarrollo infantil temprano

La evaluación del desarrollo tiene distintas etapas, en ambientes clínicos se realiza una vigilancia del desarrollo, la que ha demostrado ser poco precisa, especialmente para la identificación de aquellos niños/as con déficits más sutiles¹³. Para detectar en forma objetiva las desviaciones del DIT y poder establecer políticas públicas acordes, se han validado escalas de *screening*, tamizaje o pesquisa. Estas escalas permiten comparar el rendimiento individual contra el grupo normativo, para identificar posibles retrasos, es decir, aquellos niños/as que no cumplen los hitos esperados en cada una de las funciones evaluadas y que, por lo tanto, se beneficiarán de una evaluación diagnóstica y de una intervención¹⁴.

Las pruebas de tamizaje son instrumentos de administración sencilla y fácil interpretación. A nivel mundial se han construido distintas pruebas de *screening* del DIT, algunas son de observación directa, es decir, que deben ser aplicados por profesionales que requieren de capacitación; mientras que otras son de reporte de padres o cuidadores principales, que observan al niño/a en su ambiente natural¹⁵. Estas últimas han ganado cada vez más espacio a nivel internacional¹⁶⁻¹⁸, destacando entre sus ventajas el empoderamiento e

involucramiento en el desarrollo de sus hijos/as^{7,19}. Si bien la Academia Americana de Pediatría (AAP) se ha inclinado cada vez más por el uso de estas últimas¹⁷, una de las limitantes para su aplicación masiva en países de medianos y bajos recursos es que su precisión podría estar sesgada por la menor alfabetización y/o comprensión lectora de los padres²⁰⁻²². Como contraparte, depender de tiempo profesional entrenado podría restringir la posibilidad de tamizaje masivo¹⁸, especialmente en tiempos de epidemias, en que los escasos recursos de salud deben ser redestinados.

Otras consideraciones al momento de implementar una prueba de tamizaje son la adecuación cultural de la escala a utilizar, además de la edad a la que está dirigida y los focos de la evaluación^{14,23}. Por otro lado, se debe demostrar que los resultados de los test son confiables y permiten guiar las decisiones clínicas²⁴.

Elementos metodológicos para la construcción y análisis de los instrumentos de medición en salud

Cuando se construye un instrumento de medición en salud, lo primero que se hace es seleccionar los ítems, para lo cual se realiza una revisión de la literatura disponible, luego se recurre a metodologías mixtas, para valorar su relevancia. Posteriormente, se realiza la estandarización, que busca disminuir la variabilidad determinada por distintos observadores o situaciones, mediante un proceso en que se establece cómo debe aplicarse la batería, qué materiales son necesarios y cuáles son los puntos de corte más precisos. Para ello, las respuestas a los ítems evaluados son llevados a un puntaje que luego es convertido en unidades de desviación estándar (DS). De este modo, se interpreta que un niño/a está en riesgo de retraso del desarrollo, cuando su puntuación se encuentra bajo la norma poblacional¹⁴.

Pero no sólo es importante contar con puntos de referencia, es necesario además comprobar la validez o exactitud con que se mide el fenómeno en estudio, la confiabilidad o estabilidad de la valoración independiente de que cambien las condiciones y la pertinencia para los sujetos a los que está destinada^{25,26}. En la tabla 1 se resumen los principales conceptos utilizados en la validación de un instrumento de medición en salud.

La validez engloba diferentes dimensiones, como la validez de contenido o aparente, que determina la relevancia y comprensibilidad integral del instrumento. La validez de constructo evalúa si la escala en realidad mide el propósito para el que fue construida, para ello se evalúa su estructura (validez interna) y se compara con otros cotejos o se prueba la hipótesis que supone que los puntajes se relacionan a factores que uno esperaría que se relacionen (validez externa)²⁷ (tabla 1).

Tabla 1. Resumen de los principales requerimientos metodológicos que deben ser considerados en la construcción de un instrumento de medición en salud

Propiedades medibles y analizables	Objetivo	Metodología y análisis	Criterios interpretación
Validez			
Validez de contenido	Determina la capacidad del instrumento de medir el constructo para el cual fue creado.	Cualitativa. Juicio de expertos. Ejemplo: Método Delphi	Consenso y criterio de saturación, que se logra cuando los expertos consideran que todos los elementos son pertinentes.
Validez de constructo (interna) estructural	Ayuda a examinar la estructura interna de un test.	Análisis factorial confirmatorio (AFC) (i) Root Mean Square Error (RMSEA) (ii) Standardized Root Mean Square Residual (SRMR) (iii) Index of comparative adjustment (CFI) (iv) Tucker Lewis index (TLI) Análisis Factorial Exploratorio (AFE) (i) Eigenvalues (Valores Propios) (ii) Porcentaje acumulado de varianza explicada	Análisis factorial confirmatorio (AFC) (i) RMSEA entre 0,05 y 0,08 es aceptable, si es $\leq 0,05$ indica un buen ajuste de los datos. (ii) SRMR entre 0,05 y 0,1 es aceptable, si es $< 0,05$ indica buen ajuste. (iii) CFI $> 0,9$ es aceptable, si es $\geq 0,95$ indica que se identificó el modelo. (iv) TLI $> 0,95$ es aceptable, si es $\geq 0,97$ indica que se identificó el modelo. Análisis Factorial Exploratorio (AFE) (i) Eigenvalues: Se suelen retener factores con valores propios mayores a 1. (ii) Porcentaje acumulado de varianza explicada: Se busca un porcentaje acumulado elevado, 70-80% o más, para garantizar que se capture una cantidad sustancial de la varianza total.
Validez de constructo (externa) convergente	Evalúa en qué medida un instrumento está midiendo de manera consistente el mismo fenómeno o constructo que otros instrumentos o medidas que se supone evalúan aspectos similares.	(i) Regresión logística. Odds Ratio (OR) o Riesgo relativo (RR) dependiendo del diseño del estudio (ii) Regresión lineal (Coeficiente Beta) (iii) Varianza	(i) Factores de riesgo: OR o RR > 1; Factor protector: $= R$ o RR < 1. En ambos casos, el IC que no incluye el valor de referencia (1 para regresión logística, 0 para regresión múltiple) sugiere que la asociación es estadísticamente significativa. (ii) $p < 0,05$ rechaza la hipótesis nula de que el coeficiente Beta = 0, o que no existe pendiente en la línea. (iii) Los factores seleccionados explican al menos el 50% de la varianza.
Validez de constructo (externa) divergente	Examina las relaciones entre un instrumento y varios constructos teóricos diferentes para determinar si el instrumento puede distinguir claramente entre ellos. Se busca establecer que el instrumento mide de manera única y clara el fenómeno de interés, sin confusión con otros conceptos afines.		
Validez de criterio Concurrente	Analiza en qué medida los resultados de una escala son comparables o se correlacionan con pruebas de referencia, confeccionadas con el mismo objetivo, y medidas simultáneamente en lo posible.	(i) Sensibilidad (S), especificidad (E), valores predictivos positivos (VPP) y negativos (VPN), valores de verosimilitud (LHR) (ii) Área bajo la curva (AUC) (iii) Coeficiente de Correlación de Pearson/Spearman (para variables continuas o categóricas) (r de Pearson o p (rho) de Spearman) (iv) Análisis de Regresión (v) Índices de Concordancia: Índice Kappa de Cohen o Bland-Altman	(i) En escalas de screening de desarrollo se aceptan valores de S y E entre 70 y 80%. - Se considera que la utilidad del test es buena si LHR positivo es entre 5 y 10 y LHR negativo es entre 0,1 y 0,2, y excelente si los valores son $> 0,1$ y $< 0,1$ respectivamente. (ii) AUC $> 0,60$ es aceptable, $> 0,7$ es buena. (iii) r o p (rho) 0,3-0,5 son débilmente positivas, 0,5-0,7 moderadamente positivas y $> 0,7$ altamente positivas. (iv) Kappa $> 0,6$ es aceptable. Para el método de Bland-Altman 95 % de los puntos de datos deben estar dentro de 2 desviaciones estándar de la diferencia media.
Validez de criterio Predictiva	Analiza en qué grado los resultados de una escala predicen los resultados de una prueba de referencia a mediano/ largo plazo.		

Validez transcultural	Refiere a la capacidad del instrumento de medir de manera válida y equitativa el desarrollo de niños en diferentes culturas o contextos culturales. Especifica la pertinencia y fiabilidad de las mediciones del instrumento cuando se aplica a niños de distintas culturas, evitando sesgos culturales que podrían afectar los resultados.	(i) Relevancia cultural (ii) Equivalencia Semántica (iii) Adaptación Cultural (iv) Equidad en la Evaluación (v) Validez de Contenido Cultural (vi) Consistencia Psicométrica	Consenso y criterio de saturación obtenido en paneles de participantes representativos de las culturas en las que se adaptó/ utilizó el instrumento. Análisis de resultados de pilotaje del instrumento en muestra local y posterior adaptación cultural del instrumento. Análisis de resultados de estudios cualitativos que den voz y agencia a los pueblos o culturas participantes en la traducción y posterior adaptación cultural del instrumento.
Confiabilidad			
Consistencia interna	Determina el grado de interrelación entre los distintos ítems que conforman el instrumento. Además, mide el grado en que los ítems de la escala miden de manera conjunta la misma característica o constructo	(i) Alfa de Cronbach para puntajes continuos (ii) Coeficientes de fiabilidad de Kuder-Richardson (KR) para resultados dicotómicos (iii) Coeficiente omega de McDonald (ω) (iv) Correlación entre el puntaje estandarizado de cada ítem con el puntaje estandarizado total. Coeficiente de Correlación de Pearson /Spearman (para variables continuas o categóricas) (r)	(i) Alfa de Cronbach $> 0,7$ es aceptable. (ii) KR $> 0,5$ es aceptable. (iii) ω entre 0,7 y 0,9 es aceptable. (iv) r 0,3-0,5 son débilmente positivas, 0,5-0,7 moderadamente positivas y $> 0,7$ altamente positivas.
Estabilidad o reproducibilidad	Evalúa si las puntuaciones para sujetos que no han experimentado cambios son consistentes en mediciones repetidas bajo diversas condiciones.	Coeficiente test-retest o coeficiente de estabilidad Coeficiente de formas paralelas o de equivalencia Correlación inter-juez e intra-juez Confiabilidad inter-juez	(i) ICC $> 0,7$ es adecuado (ii) r 0,3-0,5 son débilmente positivas, 0,5-0,7 moderadamente positivas y $\geq 0,7$ altamente positivas
Confiabilidad inter-juez	Mide la consistencia de los resultados al comparar la evaluación realizada simultáneamente por al menos 2 evaluadores diferentes	Error Típico de Medida (SEM) Teoría de Respuesta al Ítem (Modelo Rasch)	(iii) Kappa $> 0,6$ es aceptable. Para el método de Bland-Altman 95 % de los puntos de datos deben estar dentro de 2 desviaciones estándar de la diferencia media.
Error de medición	Evalúa el error sistemático y aleatorio de la puntuación de un sujeto que no se atribuye a cambios reales en el constructo a medir.	(i) Coeficiente intercalase (ICC) (ii) Correlación de Pearson (r) (iii) Índices de Concordancia: Índice Kappa de Cohen o Bland-Altman	
Factibilidad			
Aplicabilidad en terreno	Mide la posibilidad real de aplicación de la escala en el entorno previsto, considerando el la facilidad, tiempo y recursos necesarios para la aplicación. También evalúa la aceptabilidad de los sujetos y la adaptabilidad a diversos contextos.	Entrevistas y encuestas a evaluadores y a sujetos evaluados; observación directa; registros de tiempo, estudio piloto, análisis de costos y recursos.	Resultados de Informes cualitativos o mixtos que permitan valorar las fortalezas, riesgos, oportunidades, alternativas y los diferentes escenarios de costo beneficio.

La validez de criterio permite conocer si los resultados son comparables con los obtenidos al aplicar un *gold standard*, aunque la elección de la referencia suele representar un verdadero desafío. A través de este análisis, se obtiene la sensibilidad, especificidad y los valores predictivos. La sensibilidad nos indica en qué medida es probable que un niño/a que tiene un rezago en su desarrollo sea pesquisado en el tamizaje, mientras que la especificidad se refiere a la probabilidad de que un infante con desarrollo típico sea clasificado correctamente como tal. Considerando que cada test tiene un error de medición propio, además de las limitaciones del patrón de referencia, en la evaluación del DIT son aceptados niveles de sensibilidad y especificidad en el orden de 0,7 a 0,8^{14,23}.

Se ha descrito un efecto espejo entre sensibilidad y especificidad, debiendo buscar el punto de mejor rendimiento del test a través de curvas de ROC (acrónimo de Receiver Operating Characteristic, o Característica Operativa del Receptor)²⁸. Se ha demostrado que a mayor sensibilidad aumenta la proporción de niños/as categorizados falsamente como positivos, lo que se ha asociado a sobrecarga emocional para la familia y sobredemanda asistencial para el sistema de salud. Es por esto que para mejorar la precisión de la detección de las dificultades del DIT, especialmente de aquellos casos más sutiles, se recomienda aplicar escalas de tamizaje a distintas edades y combinando diferentes instrumentos y modalidades de evaluación¹⁷.

La confiabilidad se refiere a la medida en que las puntuaciones están libres de error y se relaciona con la estabilidad de las medidas^{25,29}. Evalúa la consistencia de un instrumento, la homogeneidad de los distintos ítems, la relación entre ellos y su aporte al total, lo que se conoce como consistencia interna. Mide también, la variabilidad propia del sujeto al aplicar la prueba en distintas circunstancias (*test-retest*), la variabilidad entre evaluadores (inter-juez) y el error propio de la medición²⁹. Por lo general, al evaluar DIT se acepta un margen de confiabilidad de 0,7 y 0,9 (tabla 1).

La validez transcultural se refiere a la capacidad de la escala de medir de manera válida y equitativa el desarrollo de niños/as en diferentes contextos culturales. Especifica la pertinencia y fiabilidad de las mediciones cuando se aplica a individuos de distintas culturas, evitando sesgos que podrían afectar los resultados^{30,31}.

Finalmente, se analiza la factibilidad de aplicar en condiciones de la vida real. Algunos aspectos que habitualmente se evalúa son la aceptación, el tiempo de aplicación y los costos relacionados. La consistencia del instrumento se conoce luego de que se replica su uso en distintas muestras, ya sean clínicas o de investigación³², pero también es importante conocer su funcionamiento en el "mundo real"¹⁶. En relación

a los instrumentos de medición en salud, es sabido que es un proceso constante y se recomienda revisar y revalidarlos cada 10 años, debido a las transformaciones demográficas y culturales que van ocurriendo en la población; así también, se debe considerar el efecto Flynn, referido a los cambios en las normas a través del tiempo, según los cuales los puntajes globales se incrementan anualmente³³.

Escalas de tamizaje utilizadas en Chile y Latinoamérica

En Latinoamérica se emplean diversos instrumentos, incluyendo escalas desarrolladas en otros países y otros construidos para una población determinada, sin que exista consenso respecto a cuál es el "mejor"^{15,34}.

La importancia de contar con un instrumento de *screening* de DIT construido para el país radica en su apropiación cultural y representatividad de niños/as de distintas realidades; así, por ejemplo, debemos considerar que en Chile se estima que 13% de los niños/as pertenecen a pueblos indígenas y 5% provienen de familias migrantes^{35,36}. Sin embargo, hay que tener en cuenta la ventaja de utilizar pruebas previamente validados en otros países, pues permiten comparar distintas realidades, teniendo como desventaja el eventual cobro de licencia y/o derechos de uso. Pero, cualquiera sea el caso, deben pasar por procesos de validación local^{37,38}.

En la región se han construido instrumentos para la evaluación del DIT, la mayoría de los cuales son de observación directa³⁸⁻⁴⁰. Existe cierta homogeneidad en las dimensiones que evalúan, incluyendo, en su mayoría motricidad gruesa y fina/ coordinación, lenguaje/ comunicación, cognición. Cada vez más se ha optado por incluir ítems que evalúan el desarrollo socioemocional, socio-afectivo y/o personal-social, especialmente relevantes en el contexto de las recientes políticas de detección de Trastorno del Espectro Autista (TEA).

En Chile, en los años 70-80 se creó la Escala de Evaluación del Desarrollo Psicomotor (EEDP) y el Test de Desarrollo Psicomotor (TEPSI), las que se utilizan hasta la actualidad en distintos países hispanoparlantes⁴¹⁻⁴⁴, a pesar de que no han sido actualizadas ni revalidadas, carecen de un enfoque intercultural y de pertinencia étnica^{19,45}. Además, estudios cualitativos señalan que el TEPSI es extenso y monótono^{41,45}. Por otro lado, se han validado en Chile instrumentos de reporte de padres o cuidadores principales, como son el Cuestionario de Detección del Autismo en Niños y Niñas Pequeños (M-CHAT) para la pesquisa de TEA y el Ages and Stages Questionnaire (ASQ) para el tamizaje del desarrollo psicomotor^{7,46,47}. El M-CHAT es

aplicado con asistencia profesional de APS⁴⁸, y el ASQ, está en uso en algunas instituciones privadas, debido a que requiere pago de licencia para su utilización⁴⁷.

En respuesta a una largamente sentida necesidad de contar con un nuevo instrumento construido en Chile para la evaluación del DIT a gran escala, entre los años 2009 y 2012 se desarrolló el TADI, en su primera edición¹¹. En un análisis de una muestra amplia de niños/as, Bedregal y cols.⁴⁹ reportó que la sensibilidad del EEDP, TEPSI y de la primera edición del TADI era baja. En relación al TADI, en un análisis secundario, al incorporar la categoría que incluía promedio de puntuación normal, pero con una dimensión en déficit (Normal con Rezago), el test alcanzaba una sensibilidad adecuada. En tanto, la especificidad era adecuada para el EEDP y excelente para el TEPSI y TADI. Estos análisis, entre otros, motivaron a robustecer las propiedades psicométricas del TADI, cuya actualización es conocida como TADI-2¹².

Historia, criterios de validez y confiabilidad del TADI-2

El TADI-2 fue estandarizado sobre una muestra representativa de la población nacional compuesta por 882 niños/as de entre 3 y 72 meses de edad. Está integrado por cuatro dimensiones: Cognición, Motricidad, Lenguaje y Socioemocionalidad; cada una de ellas constituye una dimensión independiente.

La aplicación del TADI-2 es individual y toma entre 20 y 40 minutos, y se basa principalmente en observación directa de tareas solicitadas al niño/a y preguntas al adulto acompañante, para lo cual es necesario contar con una batería de materiales y con un profesional para su aplicación e interpretación. Al respecto es importante señalar que no requiere entrenamiento ni certificación, ya que todas las instrucciones se encuentran en el manual.

El TADI-2 conserva las dimensiones y la mayoría de los ítems de la primera edición del instrumento; así también varias de sus propiedades psicométricas fueron conservadas. El TADI-2 ofrece puntajes estandarizados hasta los 6 años y puntajes T con un promedio de 50 y una desviación estándar de 10. Los puntajes obtenidos pueden ser interpretados en categorías que clasifican al niño/a según su rendimiento en el test.

El contenido del TADI-2 fue validado mediante juicio experto. La validez de constructo se determinó mediante análisis factorial; los resultados indican la plausibilidad de que en esta edición del test las cuatro dimensiones se mantengan separadas factorialmente en todas las edades evaluadas. La validez de constructo también se determinó analizando la progresión

del desarrollo por edad mediante análisis de Teoría de Respuesta a Ítem (IRT). Otra señal de validez de constructo radica en las diferencias estadísticamente significativas entre los puntajes de los niños/as según la escolaridad de sus madres.

La evidencia de validez del TADI-2 se obtuvo al aplicarlo concurrentemente con las escalas Bayley de Evaluación del Desarrollo Infantil 3ª edición y Inventario de Desarrollo Battelle II en su versión *screening*, obteniéndose una correlación positiva estadísticamente significativa. La sensibilidad para el puntaje global en los niños/as de 4 y 5 años fue muy buena y, en los menores, moderada. Estos resultados deben ser interpretados en forma crítica, considerando las limitaciones relacionadas con las pruebas usadas como referencia, el tamaño muestral, y la definición de los puntos de corte⁵⁰.

Esta evidencia se suma a la recogida para la versión original del TADI, donde se reportó correlación significativa con las cuatro dimensiones de EEDP y con TEPSI. Adicionalmente, mostró correlación estadísticamente significativa con el Behavior Rating Inventory of Executive Function – Preschool Version (BRIEF-P).

El TADI-2 también cuenta con evidencia de validez cultural, en relación con población inmigrante de habla hispana residente en Chile (por ejemplo, bolivianas, venezolanas, colombianas y peruanas). Ésta se suma a la recogida en la primera edición del TADI para la población mapuche, aimara y atacameña. En ambos casos, los análisis realizados para comparar el desempeño de niños/as de la población general del país con similares características sociodemográficas indican que niños/as de distinto origen pueden desplegar sus habilidades sin sesgo, al ser evaluados con el TADI-2.

Respecto a la confiabilidad, se demostró la consistencia interna de las dimensiones del TADI-2, a lo que se suma a la recogida para la edición original del TADI, para la cual las cuatro dimensiones tienen adecuada consistencia interna, lo cual es consistente a través del tiempo, en que según los valores obtenidos en la correlación test-retest en todos los ítems, con independencia del examinador, dado por la correlación interjuez.

La consistencia del TADI se desprende de las múltiples aplicaciones en terreno y sus publicaciones. Estudios que han usado datos secundarios de la ELPI 2012 han descrito menor desempeño en el TADI asociado a factores socioeconómicos, lo que le otorga validez de constructo discriminante⁵¹⁻⁵⁴. También se ha reportado la validez del instrumento en grupos de riesgo como son los infantes con Síndrome de Down^{55,56}, así como su validez cultural en población Mapuche⁵⁷. La tabla 2 sintetiza las propiedades psicométricas del TADI en sus dos versiones, las que además se presentan en la figura 1.

Tabla 2. Comparación de los análisis de validación y características psicométricas del Test de Aprendizaje y Desarrollo Infantil "TADI" en su primera y segunda versión

Características psicométricas	TADI (1ª edición)	TADI-2 (2ª edición)
Año Publicación	2012	2023
Tamaño muestral	2.862	882
Poblaciones especiales	311 niños/as de los pueblos originarios	109 niños/as inmigrantes desde países hispanoparlantes
Validez de contenido	Revisión del estado del arte en desarrollo infantil. Consulta a expertos nacionales e internacionales	Revisión de expertos de contenido y orden de los ítems.
Validez de constructo (interna) estructural	Análisis factorial exploratorio y confirmatorio. El análisis mostró que las distintas dimensiones logran reconstruirse factorialmente en todos los tramos de edad, explicando entre el 38% y el 60% de la varianza.	Análisis factorial confirmatorio en base a los resultados de la aplicación a la muestra de estandarización, confirmó que las dimensiones mantienen casi la totalidad de los ítems de la versión del test original.
Validez de constructo (externa)	Correlación significativa con EEDP ($r=0,4$ a $0,6$ según dominio) TEPSI ($r=0,7$)	Correlación significativa con Bayley ($r=0,3$ a $0,5$ según dominio) BDI-2 ($r=0,7$)
Validez de criterio Concurrente	EEDP: $r=0,4$ a $0,6$ TEPSI: $r=0,7$ BRIEF-P: r cercano a $0,3$ En el estudio de profundización con BDI-2 $S=0,95$ $E=0,60$ cuando se considera como déficit la categoría Normal con rezago	<3 años ($n=93$) tomando como referencia Bayley: Cognición: $S=0,82$ $E=0,48$ Motricidad: $S=0,42$ $E=0,61$ Lenguaje: $S=0,79$ $E=0,57$ Socioemocional: $S=0,58$ $E=0,48$ >4 años ($n=80$) tomando como referencia BDI-2 screening: $S=0,92$ $E=0,53$
Validez transcultural	Análisis de Funcionamiento Diferencial de Ítems (DIF) descarta sesgo cultural para submuestra de población atacameña, aimara y mapuche	Comparación de población boliviana, colombiana, peruana y venezolana con población general del país de similares características sociodemográficas descarta sesgo contra poblaciones inmigrantes.
Consistencia interna (Alfa de Cronbach)	Análisis por edad y dimensión, todos mayores a $0,88$	Todos mayores a $0,86$
Confiabilidad test-retest	Mayor a igual a $0,97$	No se hizo.
Confiabilidad interjuez	Mayor o igual a $0,99$	No se hizo
Confiabilidad según Teoría de Respuesta al Ítem (IRT)	No se hizo	Estimación Esperada a Posteriori (EAP) Cognición $0,978$ Motricidad $0,076$ Lenguaje $0,977$ Socioemocional $0,975$
Interpretabilidad	Interpretación resultados: Avanzado: $\geq T 60$ ó $\geq +1$ DS Normal: entre $T40$ y $T59$ ó entre -1 DS y $+1$ DS Riesgo: entre $T30$ y $T39$ ó ≤ -1 DS y >-2 DS Retraso: $\leq T29$ ó ≤ -2 DS	Interpretación resultados: Avanzado: $\geq T 60$ ó $\geq +1$ DS Normal: entre $T50$ y $T59$ ó entre promedio y $+1$ DS Alerta: entre $T40$ y $T49$ ó entre -1 DS bajo el promedio y promedio. Riesgo: entre $T30$ y $T39$ ó entre $+1$ y -2 DS Retraso: $\leq T29$ ó ≤ -2 DS
	Puntaje global: Se agrega Normal con rezago cuando el promedio es normal y hay una dimensión en déficit.	Puntaje global: Mismas categorías de las dimensiones.

S: Sensibilidad; E: especificidad; DS: Desviación Estándar. EEDP: Escala de Evaluación del Desarrollo Psicomotor; TEPSI: Test de Desarrollo Psicomotor; BDI-2: Battelle Developmental Inventory 2ª edición; Bayley: Bayley Scales of Infant and Toddler Development; Behavior Rating Inventory of Executive Function - Preschool Version.

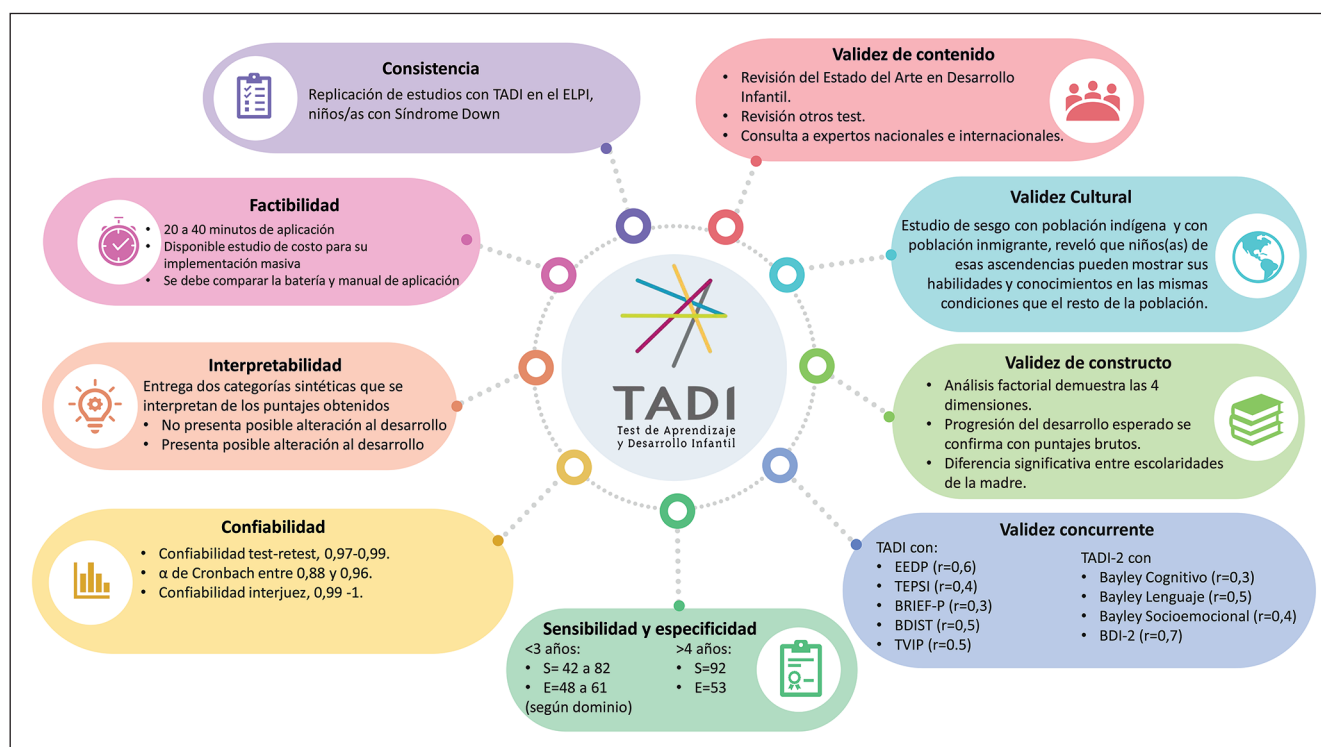


Figura 1. Resumen proceso validación TADI. EEDP Escala de Evaluación del Desarrollo Psicomotor; TEPSI: Test de Desarrollo Psicomotor; BRIEF-P: Behaviour Rating Inventory of Executive Function, Preschool Version; BDI-2: Battelle Developmental Inventory 2° edición. TVIP: Test de Vocabulario en imágenes Peabody; Bayley: Bayley Scales of Infant and Toddler Development, Bayale-III: Third Edition; S: Sensibilidad; E: Especificidad; ELPI: Encuesta Longitudinal Primera Infancia, Chile.

Conclusiones

El proceso de validación de un instrumento es dinámico y continuo. De acuerdo a la evidencia recabada, el TADI-2 ha sido validado a través de un riguroso proceso en población nacional, incluyendo grupos indígenas e inmigrantes, por lo que concluimos que puede ser implementado para el *screening* del DIT de los niños/as menores de 6 años que habitan en Chile y el resto del continente. Para mejorar la precisión, se recomienda la monitorización de desarrollo aplicando escalas de tamizaje a distintas edades y combinando distintos instrumentos y modalidades de evaluación. Sin duda se debe continuar estudiando el TADI-2 en muestras más amplias para conocer la consistencia de

los resultados en estudios complementarios y poblacionales.

Agradecimientos

Marcela Pardo y Marta Edwards agradecen el financiamiento otorgado por ANID/PIA/Fondos Basales para Centros de Excelencia FB0003.

Conflicto de intereses

Marta Edwards y Marcela Pardo son las autoras del Test de Aprendizaje y Desarrollo Infantil (TADI).

Referencias

- Schonhaut L, Zepeda A, Rojas P. One hundred years of health in a demographic, epidemiological and public policy context: the transition of childhood morbidity and its challenges. *Andes Pediatrica*. 2022;93(6):799-806. doi: 10.32641/andespediatr.v93i6.4539
- Wasserman RC, Fiks AG. The Future(s) of Pediatric Primary Care. *Acad Pediatr*. 2021;(3):414-24. doi: 10.1016/j.acap.2020.10.015.
- Guralnick MJ. Effectiveness of early intervention for vulnerable children: a developmental perspective. *Am J Ment Retard*. 1998;102(4):319-45. doi: 10.1352/0895-8017(1998)102<0319:eoefv>2.0.co;2.
- Olusanya BO, Davis AC, Wertlieb D, et al. Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Glob Health*. 2018;6(10):e1100-21. doi: 10.1016/S2214-109X(18)30309-7.
- Wondmagegn T, Girma B, Habtemariam Y. Prevalence and determinants of developmental delay among children in low- and middle-income countries: a systematic review and meta-analysis. *Front Public Health*. 2024;12:1301524. doi: 10.3389/fpubh.2024.1301524.
- Gil JDC, Ewerling F, Ferreira LZ, et al. Early childhood suspected developmental delay in 63 low-and middle-income countries: Large within- and between-country inequalities documented using national health surveys. *J Glob Health*. 2020;10(1):010427. doi: 10.7189/jogh.10.010427.
- Armijo I, Schonhaut L, Cordero M. Validation of the Chilean version of the Ages and Stages Questionnaire (ASQ-CL) in Community Health Settings. *Early Hum Dev*. 2015;91(12):671-6. doi: 10.1016/j.earlhumdev.2015.10.001.
- Ministerio de Salud Chile Informe Encuesta Nacional de Salud 2016-2017. Desarrollo Funcional Infantil. 2018. [citado 05.01.24] Disponible en: <https://goo.gl/oe2iVt>
- Bedregal P, Hernández V, Mingo MV, et al. Desigualdades en desarrollo infantil temprano entre prestadores públicos y privados de salud y factores asociados en la Región Metropolitana de Chile. *Rev Chil Pediatr*. 2016;87(5):351-8. <http://dx.doi.org/10.1016/j.rchipe.2016.02.008>
- Ministerio de Salud Chile. Estrategia Nacional de Salud para los Objetivos Sanitarios al 2030. 2022. [citado 05.01.24] Disponible en: <https://www.minsal.cl/wp-content/uploads/2022/03/Estrategia-Nacional-de-Salud-2022-MINSAL-V8.pdf>
- Pardo M, Gomez M, Edwards M. Test de Aprendizaje y Desarrollo Infantil (TADI). 2012. Santiago, Chile: Editorial Universitaria.
- Edwards M, Pardo M. TADI-2: Test de Aprendizaje y Desarrollo Infantil, segunda edición. Manual de Aplicación 2024. Santiago, Chile: Editorial Universitaria.
- Sheldrick RC, Merchant S, Perrin EC. Identification of developmental-behavioral problems in primary care: A systematic review. *Pediatrics*. 2011;128(2):356-63. doi: 10.1542/peds.2010-3261.
- Council on Children With Disabilities; Section on Developmental Behavioral Pediatrics; Bright Futures Steering Committee; Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. *Pediatrics*. 2006;118(1):405-20. doi: 10.1542/peds.2006-1231. Erratum in: *Pediatrics*. 2006;118(4):1808-9.
- Romo-Pardo B, Liendo-Vallejos S, Vargas-López G, et al. Pruebas de tamizaje de neurodesarrollo global para niños menores de 5 años de edad validadas en Estados Unidos y Latinoamérica: revisión sistemática y análisis comparativo. *Bol Med Hosp Infant Mex*. 2012;69(6):450-62.
- Rah SS, Jung M, Lee K, et al. Systematic Review and Meta-analysis: Real-World Accuracy of Children's Developmental Screening Tests. *J Am Acad Child Adolesc Psychiatry*. 2023;62(10):1095-109. doi: 10.1016/j.jaac.2022.12.014.
- Lipkin PH, Macias MM. Promoting Optimal Development: Identifying Infants and Young Children With Developmental Disorders Through Developmental Surveillance and Screening. *Pediatrics*. 2020;145(1):e20193449. doi: 10.1542/peds.2019-3449.
- Lipkin PH, Macias MM, Baer Chen B, et al. Trends in Pediatricians' Developmental Screening: 2002-2016. *Pediatrics*. 2020;145(4):e20190851. doi: 10.1542/peds.2019-0851.
- Schonhaut L, Armijo I, Millan T, et al. Comparación de la Evaluación Tradicional del Desarrollo Psicomotor versus una Prueba Autoadministrada. *Rev Chil Pediatr*. 2010;81(6):498-505. <http://dx.doi.org/10.4067/S0370-41062010000600003>
- Shrestha M, Schwinger C, Hysing M, et al. Agreement Between Mothers and Fieldworkers While Assessing Child Development Using Ages and Stages Questionnaires, Third Edition in Nepal. *Front Psychol*. 2020;11:579412. doi: 10.3389/fpsyg.2020.579412. eCollection 2020.
- Eugenin MI, Moore R, Martinez-Gutierrez J, et al. Screening for autism in Santiago Chile: Community perspectives. *Int J Child Adolesc Health*. 2015;8(4):439-48.
- Lockhart M, Chauv R, Chevin M, et al. Classification Performance of the Ages and Stages Questionnaire: Influence of Maternal Education Level. *Children*. 2023;10(3):449. doi: 10.3390/children10030449.
- Faruk T, King C, Muhit M, et al. Screening tools for early identification of children with developmental delay in low- and middle-income countries: a systematic review. *BMJ Open*. 2020;10(11):e038182. doi: 10.1136/bmjopen-2020-038182. Erratum in: *BMJ Open*. 2021;11(11):e038182corr1.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). Standards for educational and psychological testing. 2014.
- Alarcón AM, Muñoz S. Medición en salud: Algunas consideraciones metodológicas. *Rev Méd Chile*. 2008;136:125-30. <http://dx.doi.org/10.4067/S0034-98872008000100016>
- Aiken LR. Tests psicológicos y evaluación. 11° edition. Mexico: Pearson Education; 2003.
- Bean GJ, Bowen NK. Item Response Theory and Confirmatory Factor Analysis: Complementary Approaches for Scale Development. *J Evid-Based Soc Wor*. 2021;18(6):597-618. <https://doi.org/10.1080/26408066.2021.1906813>
- Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*. 2007;115(5):654-7. doi: 10.1161/CIRCULATIONAHA.105.594929.
- Manterola C, Grande L, Otzen T, et al. Confiabilidad, precisión o reproducibilidad de las mediciones. Métodos de valoración, utilidad y aplicaciones en la práctica clínica. *Rev Chilena Infectol*. 2018;35(6):680-8. <http://dx.doi.org/10.4067/S0716-10182018000600680>
- McHenry MS, Oyungu E, Yang Z, et al. Cultural adaptation of the Bayley Scales of Infant and Toddler Development, 3rd Edition for use in Kenyan children aged 18-36 months: A psychometric study. *Res Dev Disabil*. 2021;110:103837. doi: 10.1016/j.ridd.2020.103837.
- Arestad KE, MacPhee D, Lim CY, et al. Cultural adaptation of a pediatric

- functional assessment for rehabilitation outcomes research. *BMC Health Serv Res.* 2017;17(1):658. doi: 10.1186/s12913-017-2592-6.
32. Ana Carvajal Valcárcel C. ¿Cómo validar un instrumento de medida de la salud? *An. Sist. Sanit. Navar.* 2011;34(1):63-72.
 33. Aylward GP. Developmental testing. *The Cambridge Encyclopedia of Child Development.* Hopkins B, Geangu E, Linkenauger S, editors. Cambridge: Cambridge University Press 2017;88-92
 34. Seibel BL, Lerner R, Bacelar RC, et al. Child development assessment in Latin American countries: A systematic review. *Rev Bras Aval.* 2024;13(1):e130224. <http://dx.doi.org/10.4322/rbaval202412002>
 35. Ministerio de Desarrollo Social y Familia. Encuesta CASEN 2017: niños, niñas y adolescentes. 2017. [Citado 05.01.24] Disponible en <https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2017>
 36. Instituto Nacional de Estadísticas (INE), Fondo de las Naciones Unidas para la Infancia Chile (UNICEF). Niñez y adolescencia migrante en Chile - Estimación de población Análisis a partir de la estimación de personas extranjeras residentes en Chile al 31 de diciembre de 2021. [Citado 05.01.24] Disponible en <https://serviciomigraciones.cl/wp-content/uploads/2023/07/informe-estimacion-ninez-y-adolescencia-migrante.pdf>
 37. Rousseau M, Dionne C, Savard RT, et al. Translation and Cultural Adaptation of the Ages and Stages Questionnaires (ASQ) Worldwide. *J Dev Behav Pediatr.* 2021;42(6):490-501. doi: 10.1097/DBP.0000000000000940.
 38. Jurado-Castro V, Carlos Rebolledo-Cobos R. Análisis de escalas para la evaluación del desarrollo infantil usadas en américa: una revisión de literatura. *Rev Mov Cient.* 2016;10(2):72-82. <https://doi.org/10.33881/2011-7191.mct.10206>
 39. UNICEF Argentina F. Instrumentos y metodologías de evaluación del desarrollo infantil temprano [Internet]. 2023. [Citado 15.01.24] Disponible en www.unicef.org.ar
 40. Romero Otalvaro C, María A, Argel M, et al. Instrumentos de evaluación de pesquisa de neurodesarrollo en la intervención temprana. *Tesis Psicológica.* 2016;11(2):54-71.
 41. Angulo-Ramos M, Merino-Soto C. TEPSI en cuestión: ¿usarlo mejor o reemplazarlo? *Rev enferm Herediana.* 2014;(7): 107. <https://doi.org/10.20453/renh.v7i2.2534>
 42. Espósito AVL, Korzeniowski CG, Santini Bertoldi M. Normas preliminares del Test de Desarrollo Psicomotor (TEPSI) para niños argentinos de 3 y 4 años. *Liber.* 2018;24(1):9-27. <http://dx.doi.org/https://doi.org/10.24265/liberabit.2018.v24n1.0>
 43. Zela-Coila F, Merma-Valero JM, Cevallos-Ariza F, et al. Necesidad de validación de los test de evaluación del crecimiento y desarrollo de niños menores de cinco años en Perú. *Investig Innov Clin Quir Pediatr.* 2023;1(1):90-2. doi:10.59594/iicqp.2023.v1n1.18.
 44. Gutierrez EL, Lazarte F, Alarcon G, et al. La importancia de la evaluación del neurodesarrollo en niños menores de treinta meses en el contexto peruano. *Acta Med Peru.* 2016;33(4):304-12.
 45. Rivas Riveros E, Catalán MY, Flores FN, et al. Vivencias y expectativas de los actores sociales participantes en evaluación del desarrollo psicomotor, en relación a las Escalas EEDP y TEPSI utilizadas en la Araucanía, Chile, durante los años 2017-2018. *Enfermería (Montevideo).* 2019;8(1):4-21.
 46. Coelho-Medeiros ME, Bronstein J, Aedo K, et al. M-CHAT-R/F Validation as a screening tool for early detection in children with autism spectrum disorder. *Rev Chil Pediatr.* 2019;90(5):492-9. doi: 10.32641/rchped.v90i5.703.
 47. Schonhaut L, Armijo I, Rojas P, et al. Reliability and acceptability of web-based administration of Spanish ages and stages questionnaires third edition ®. *Infant Child Dev.* 2023;32(4):e2425. <https://doi.org/10.1002/icd.2425>
 48. Ministerio de Salud de Chile. Norma Técnica para la Supervisión de salud integral de niños y niñas de 0 a 9 años en la Atención Primaria de Salud - 2ª edición. Actualización 2021. [Citado 04.01.24] Disponible en [https://www.minsal.cl/sites/default/files/files/2014_Norma%20T%c3%a9cnica%20para%20la%20supervisi%c3%b3n%20de%20ni%c3%b1os%20y%20ni%c3%b1as%20de%200%20a%209%20en%20APS_web\(1\).pdf](https://www.minsal.cl/sites/default/files/files/2014_Norma%20T%c3%a9cnica%20para%20la%20supervisi%c3%b3n%20de%20ni%c3%b1os%20y%20ni%c3%b1as%20de%200%20a%209%20en%20APS_web(1).pdf)
 49. Bedregal P, Villarreal L, Padilla O, Valenzuela P. Evaluación de instrumentos de medición del desarrollo infantil para niños y niñas de 0 a 5 años, para mejorar el seguimiento de la trayectoria de desarrollo en la primera infancia en Chile. 2017 (versión corregida marzo 2019). [Citado 04.01.24] Disponible en http://54.148.75.48/bitstream/handle/123456789/3876/INFORME-FINAL-EVALUACIO%cc%81N-INSTRUMENTO_marzo-2019.pdf?sequence=1
 50. Cunha ACB, Berkovits MD, Albuquerque KA. Developmental Assessment with Young Children: A Systematic Review of Battelle Studies. *Infants Young Child.* 2018;31(1):69-90.
 51. Rodríguez-Garcés C, Muñoz-Soto J. Delay in child development: The importance of education quality of family atmosphere. *Rev Int Investig Cienc Soc.* 2017;13(2):253-70. <https://doi.org/10.18004/riics.2017.diciembre.253-270>.
 52. Narea M, Arriagada V, Allel K. Center-Based Care in Toddlerhood and Child Cognitive Outcomes in Chile: The Moderating Role of Family Socio-Economic Status. *Early Educ Dev.* 2020;31(2):218-33. <https://doi.org/10.1080/10409289.2019.1626191>
 53. Reynolds SA. Center-Based Child Care and Differential Improvements in the Child Development Outcomes of Disadvantaged Children. *Child Youth Care Forum.* 2022;51(2):395-420. doi: 10.1007/s10566-021-09634-0.
 54. Soto-Ramírez P, Jiang H, Strasser K, et al. Early home learning environment profiles and their contribution to child development in Chile. *J Appl Dev Psychol.* 2022;82:101452. <https://doi.org/10.1016/j.appdev.2022.101452>.
 55. Tenorio M, Bunster J, Arango PS, et al. Reliability and validity evidence for test de aprendizaje y desarrollo infantil (TADI) in a Chilean sample of children with down syndrome. *Psykhé.* 2020;29(1):1-16.
 56. Tenorio M, Arango PS, Aparicio A. BENDI: Improving Cognitive Assessments in Toddlers and Children with Down Syndrome Using Stealth Assessment. *Children.* 2023;10(12):1923. <https://doi.org/10.3390/children10121923>
 57. Salinas V, Valenzuela EA, Aranis DP. Cognitive development and parenting during early childhood among Mapuche and non-indigenous Chileans. *Early Child Res Q.* 2021;55:165-78. <https://doi.org/10.1016/j.ecresq.2020.11.001>.