

Evidencia de Confiabilidad y Validez del Test de Aprendizaje y Desarrollo Infantil (TADI) para Evaluación de Niñas y Niños Chilenos con Síndrome de Down

Reliability and Validity Evidence for Test de Aprendizaje y Desarrollo Infantil (TADI) in a Chilean Sample of Children with Down Syndrome

Marcela Tenorio, Josefina Bunster y Paulina Sofía Arango
Universidad de los Andes

Andrés David Aparicio, Ricardo Rosas y Katherine Strasser
Pontificia Universidad Católica de Chile

Este estudio ofrece un análisis de la evidencia de validez y confiabilidad para el uso de TADI, según los estándares internacionales, en infantes y niños chilenos con síndrome de Down (SD). Participaron 48 niños de Santiago, La Calera y Rancagua, incorporados tras muestreo no-probabilístico, con diagnóstico médico confirmado de SD y sus pares con desarrollo típico, pareados por edad cronológica ($M = 4,79$, $DE = 1,7$ años), seleccionados de manera aleatoria desde la base de datos ELPI (2013). Se analiza la evidencia de confiabilidad, vía análisis de consistencia de interna, así como la evidencia de validez de contenido, de estructura y por relación con otras variables. Para la medición fueron utilizados el Test de Aprendizaje y Desarrollo Infantil (TADI), la Escala Wechsler de Inteligencia, tercera edición-versión chilena (WISC-IIIv.ch.), el Leiter International Performance Scale, cuestionario socioeconómico y cuestionario sobre características de la condición médica. Los resultados sugieren una consistencia interna apropiada, la evidencia de validez indica que el contenido de la prueba es apropiado, no se registran efectos techo ni piso, hay diferencias en la estructura interna registrada en el grupo de SD y una adecuada relación con variables de criterio externo. Los profesionales requieren criterio para el uso de esta prueba en la muestra de interés y futuros estudios deberán ampliar los resultados presentados.

Palabras clave: síndrome de Down, escala de desarrollo, evaluación, TADI

This study presents the evidence of reliability and validity for TADI in a sample of Chilean infants and children with Down syndrome (DS), accordingly to international standards. Forty-eight children with confirmed medical diagnosis of DS and their typical peers were matched by chronological age ($M = 4,79$, $DE = 1,7$ years). Children with DS were selected using a non-probabilistic procedure and, children with typical development were randomly selected from ELPI's data base (2013). Data was obtained from Test de Aprendizaje y Desarrollo Infantil (TADI), Wechsler Intelligence Scale, third edition-Chilean version (WISC-IIIv.ch.), Leiter International Performance Scale, socioeconomic questionnaire and a survey about the medical condition. The evidence of reliability was established via Cronbach's Alpha, the evidence of validity was based on the analysis of content, internal structure and relation with other variables. Results showed good levels of reliability. The evidence of validity shows the content is adequate to explore development, there are not floor or ceiling effects, the internal structure for the group with DS is different from the one reported in the original test and there is strong relation with the external criteria. The professionals should be careful in the interpretation process and future studies must complete these results.

Keywords: Down syndrome, developmental scale, assessment, TADI

Marcela Tenorio D., Josefina Bunster y Paulina Sofía Arango, Escuela de Psicología, Universidad de los Andes, Santiago, Chile; Andrés David Aparicio, Facultad de Artes, Pontificia Universidad Católica de Chile, Santiago, Chile; Ricardo Rosas y Katherine Strasser, Escuela de Psicología, Pontificia Universidad Católica de Chile, Santiago, Chile.

Esta investigación ha sido financiada gracias al aporte de la Comisión Nacional de Ciencia y Tecnología del Gobierno de Chile, Proyecto FONDECYT Iniciación 11150800. Los autores agradecen especialmente a Marta Edwards y al equipo de CEDEP por su generoso acceso al TADI. Reconocen también al equipo de evaluación y apoyo, que hizo posible el trabajo con los niños, especialmente a Javiera Donoso, Ana Karen Fernández, María Ignacia Heusser, Jaime Lubascher, Camila Sabat, Consuelo Reyes, Valeria Valdivieso y Catalina A. Vidal.

La correspondencia relativa a este artículo debe ser dirigida a Marcela Tenorio, Escuela de Psicología, Universidad de los Andes, Avda. Monseñor Álvaro del Portillo 12.455, Edificio Mecano, piso 2, Las Condes, Santiago, Chile. E-mail: mtenorio@uandes.cl

La evaluación del desarrollo es una tarea que se realiza en contextos clínicos en el área de salud con infantes y niños en edad preescolar. Es una evaluación relevante que permite comparar el rendimiento individual contra el grupo normativo, otorgando criterios para establecer las áreas de fortalezas y de riesgo (Bellman, Byrne & Sege, 2013). Es considerada un procedimiento estándar para comprensión diagnóstica y el diseño de intervenciones durante la primera infancia (Akshoomoff, 2006; Long, Blackman, Farrell, Smolkin & Conaway, 2005).

Hay estándares internacionales disponibles que fijan la calidad deseable de los instrumentos que se utilizan para la medición en psicología y, dentro del área, para las mediciones del desarrollo (e.g., American Educational Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 2014; International Commission Test, 2018). Actualmente hay varias escalas de medición del desarrollo cuya evidencia de confiabilidad y validez es robusta a la luz de estos estándares, destacándose la Escala Bayley de Desarrollo Infantil, tercera edición (BSID-III; Bayley, 2005) y la Escala Mullen de Aprendizaje Temprano (Mullen, 1995), consideradas *gold estándar* en muchos países (Ringwalt, 2008). Sin embargo, hasta donde los autores de este artículo tienen conocimiento, no existen adaptaciones de estas escalas para los países latinoamericanos, así como tampoco estudios regionales sobre el uso de escalas de medición de desarrollo en poblaciones clínicas.

En Chile está disponible el Test de Aprendizaje y Desarrollo Infantil (TADI; Edwards & Pardo, 2013), que es una herramienta compleja en la que se conjugan la estructura de la Escala Bayley de Desarrollo Infantil, segunda edición (BSID-II; Bayley, 1993) con algunos principios del Inventario MacArthur para evaluación de la comunicación (Fenson et al., 1993). TADI fue desarrollado y validado con 3200 niños chilenos con desarrollo típico (DT), ofrece normas para 13 grupos de edad y presenta evidencia de confiabilidad y validez para el grupo poblacional sobre el que se basó su baremación. De acuerdo con la información disponible en el Manual TADI (Edwards & Pardo, 2013), la prueba presenta coeficientes de consistencia interna explorado vía Alfa de Cronbach que oscilan entre 0,8 y 0,9. En la evidencia de validez, los autores ofrecen información basada en el contenido del test a partir del juicio de expertos, evidencia de validez basada en la estructura interna de la prueba vía Análisis Factorial Confirmatorio con valores aceptables por edad y escala en los indicadores presentados según parámetro internacional (Schumacker & Lomax, 2016). Aparece también reportado análisis de progresión del desarrollo con el test de Sidak, con resultados en parámetros adecuados. Se presenta evidencia de validez cultural del contenido con resultados que descartan presencia de sesgo. Hay informe de evidencia de validez por relación con otras variables, comparándose TADI con Behavior rating inventory of executive function-preschool (BRIEF-P) (Gioia, Andrwes & Isquith, 1996), Escala de evaluación de desarrollo psicomotor para niños entre 0 y 2 años (Rodríguez, Arancibia & Undurraga, 1976) y TEPSI (Haeussler & Marchant, 2003).

Sin duda, las escalas tradicionales son ideales cuando hay una trayectoria de desarrollo típica; sin embargo, su utilidad disminuye al ser aplicadas en presencia de condiciones de riesgo, como es el caso del síndrome de Down (SD; Moore, Goodwin & Oates, 2008), debido a las particulares características y específicas necesidades que tienen estos grupos. En presencia de estas condiciones de riesgo, la evaluación de desarrollo cobra particular relevancia. Su función va más allá de establecer si hay un retraso general, buscando comprender las asociaciones y disociaciones de rendimiento respecto a las trayectorias, con el fin de comprender el perfil de funcionamiento y facilitar el diseño y monitorización de planes de intervención, tanto clínicas como educativas (Moore et al., 2008).

A pesar de la reconocida importancia que tiene el uso de escalas de desarrollo en el seguimiento de condiciones de riesgo, se han descrito algunas dificultades en su uso para la exploración y trazado del perfil de fortalezas y debilidades (Anderson et al., 2010). Estas dificultades se han relacionado, especialmente, con las dificultades en el desarrollo de cognición en general y lenguaje en particular que suele acompañar a estos grupos (Moore et al., 2008). Esto resulta ser una situación paradójica, pues, mientras es poco probable que la familia de un niño con DT consulte a especialistas para que evalúen su ritmo de crecimiento y desarrollo cognitivo, los niños con SD (o que portan cualquier otro factor de riesgo para el neurodesarrollo) deben ser monitoreados con regularidad.

Chile es un país que tiene una tasa elevada de nacimientos de niños con SD, si se compara con las tasas mundiales (Nazer & Cifuentes, 2011) y, en consecuencia, requiere un sistema público robusto que permita dar atención a las necesidades de este grupo. Sin embargo, la práctica diaria de evaluación del desarrollo está basada en (y a menudo limitada a) el uso de instrumentos cortos de tamizaje, cuyos alcances comprensivos son limitados (Bedregal, 2008).

Los instrumentos usados con mayor frecuencia para la evaluación de desarrollo en Chile son la Escala de Evaluación del Desarrollo Psicomotor (EEDP), que cubre el rango de niños de ocho a 18 meses (Rodríguez, Arancibia & Undurraga, 1976) y el Test de Desarrollo Psicomotor (TEPSI), que cubre el rango de niños de 2 a 5 años (Haeussler & Marchant, 2003). Su alto uso está asociado a su incorporación como parte de los procedimientos basales en el programa nacional de protección a la infancia conocido como Chile Crece Contigo, dependiente del Ministerio de Desarrollo Social. Es posible observar la incorporación de las Escalas Bayley (en su segunda y tercera edición) en algunas instituciones de servicio privado; sin embargo, estas escalas no tienen pruebas para grupos clínicos en el país y, en algunas ocasiones, incluso carecen de versiones en español y verificación apropiada de sus cualidades psicométricas en el grupo poblacional de interés.

Más allá de estas barreras, pueden observarse algunos problemas técnicos que requieren atención. Primero, los instrumentos disponibles suelen producir *efecto piso* al ser considerada la edad cronológica de los niños, ya que el resultado estandarizado que se obtiene con la evaluación aparece con frecuencia a dos desviaciones estándar bajo la media. Esto hace difícil observar una varianza en las puntuaciones estandarizadas que permita documentar fortalezas y debilidades del niño, así como observar los cambios posibles secundarios a las intervenciones ofrecidas (Tenorio, Campos & Karmiloff-Smith, 2014).

Segundo, no están disponibles en el país instrumentos que permitan establecer la edad mental y, por lo tanto, no es posible contrarrestar, al menos para efectos de medición y seguimiento, el efecto piso que produce la edad cronológica.

Tercero, no hay estudios respecto a las características psicométricas de las escalas, pruebas y baterías disponibles para estas evaluaciones, en poblaciones clínicas en general y para SD en particular. Si se atiende a los estándares internacionales, su uso no está indicado en esta población (AERA, APA & NMCE, 2014).

Cuarto, tal y como sucede en otros lugares del mundo, la carencia de datos normativos asociados a esta población en los instrumentos de mayor uso es un problema, pues siempre se está comparando al niño contra pares típicos, sin lograr una caracterización propia de su varianza de desempeño y sus avances (Byrne, MacDonald & Buckley, 2002). La solución a este problema tampoco parece ser la generación de normas para cada grupo clínico de interés, sino la verificación de las cualidades psicométricas de la prueba ante su uso en grupos de desarrollo atípico.

Considerado todo lo anterior, el objetivo de este estudio fue ofrecer información respecto a la evidencia de confiabilidad y validez de TADI para su uso clínico con niñas y niños chilenos con SD. El método de análisis y la presentación de resultados se han organizado de acuerdo con los estándares internacionales para evaluación en psicología y educación del consorcio AERA, APA y NMCE (2014).

Vale la pena recordar que dichos estándares establecen que las pruebas no serán confiables o válidas en términos absolutos, sino que la evidencia deberá ser integrada por el profesional en función de las necesidades que ha fijado para cada evaluación. En este sentido, este estudio busca brindar evidencia acerca de las posibilidades del uso de TADI en la práctica clínica con niños con SD, más allá de su uso diagnóstico para determinar la presencia o no de riesgo o retraso en el desarrollo. En este estudio se analizan las posibilidades de uso de TADI como un instrumento que ayude a resolver la paradoja que se anunció en párrafos anteriores sobre el uso clínico de la prueba, para trazar un perfil de fortalezas y debilidades.

Método

Se presenta un estudio transversal en el que se aplicaron mediciones estandarizadas de desarrollo a niños con SD y a niños con DT.

Participantes

En este estudio se analizaron resultados de 96 niños con edades cronológicas ubicadas entre 1:2 y 7:0 años. La muestra se dividió en dos grupos, uno de niños con diagnóstico médico confirmado de SD y otro con niños con DT, cada grupo con 48 sujetos. Los grupos fueron pareados por edad cronológica ($M = 4,79$, $DE = 1,7$ años para ambos grupos) y nivel socioeconómico (NSE) con una estrategia uno a uno.

Para seleccionar los niños con SD, se consideró la presencia de diagnóstico médico confirmado vía cariotipo para su inclusión. Se realizó un muestreo no probabilístico, los participantes se contactaron a través de instituciones especializadas en atención de estos niños en las ciudades de Santiago, La Calera y Rancagua.

Los niños con DT fueron seleccionados de manera aleatoria a través de un algoritmo de rastreo diseñado para este fin específico, desde la base disponible de la Encuesta Longitudinal de la Infancia (ELPI; Bravo, 2013).

La decisión de realizar el pareo entre grupos en función de la edad cronológica y no a partir de la edad mental, criterio que suele privilegiarse al estudiar niños con SD, obedece a dos razones. En primera instancia, en la práctica actual, cuando un profesional hace uso de una escala exhaustiva de desarrollo con un niño que tiene un curso atípico de desarrollo, utiliza los mismos principios y normas para la aplicación e interpretación de resultados que se usan con niños con DT. Este estudio busca hacer evidentes las dificultades y barreras intrínsecas a estas acciones y decisiones. La segunda razón es que en este momento en Chile no está disponible una prueba adecuada para el cálculo de la edad mental. La única práctica posible es hacer una aproximación desde la medición del coeficiente intelectual, lo que, a su vez, supone que el niño ha podido enfrentar alguna Escala de Wechsler de Inteligencia para Niños (tercera o quinta edición, actualmente disponibles en el país; Ramírez & Rosas, 2010; Rosas & Pizarro, 2018). Esta acción solo puede ser ejecutada en una porción de niños con SD y solo desde los 6:0 años, edad superior a la edad mínima de los niños incorporados en este estudio.

Respecto a la estimación del NSE, se utilizó la dependencia escolar como proxy de esta variable, siguiendo las recomendaciones y métodos utilizados en estudios previos en Chile (Arango, Aparicio & Tenorio, 2018; García-Huidobro & Bellei, 2003; González, Mizala & Romaguera, 2004; Rojas-Barahona et al., 2017; Rosas & Santa Cruz, 2013; Santos & Elacqua, 2016). El grupo de NSE bajo está compuesto por los niños que asisten a colegios municipales (públicos), el NSE medio, por los que asisten a colegios particulares subvencionados (privados que reciben fondos públicos) y el NSE alto, por los que asisten a colegios particulares pagados (privados).

Como criterios de inclusión para el grupo de niños con SD se consideró tener diagnóstico médico confirmado, presentar agudeza visual y auditiva adecuada o corregida y contar con autorización de su apoderado para participar en el estudio. Fueron criterios de exclusión haber estado hospitalizado en los tres meses anteriores a la evaluación y estar con manejo farmacológico con psicotrópicos. Se presenta una descripción de la muestra de niños con SD en la Tabla 1.

Tabla 1

Información Descriptiva sobre Educación, Factores Médicos Asociados e Hitos de Desarrollo en la Muestra de Niñas y Niños con Síndrome de Down Participantes del Estudio

Dimensión	Variable	Categoría	Valor
Educación	Tipo de educación	Especial (%)	40
		Regular (%)	47
Factores de riesgo	Cardiopatías (%)		47
	Tiempo de gestación (media en semanas)		37,1 (DE = 2,1)
	Tipo de parto	Vaginal (%)	37
		Cesárea (%)	50
		Conocimiento prenatal diagnóstico (%)	2
	Edad materna en el nacimiento (media en años)		37,0 (DE = 6,6)
	Peso perinatal (media en gramos)		3028 (DE = 628)
Hitos del desarrollo	Gateo (media en meses)		13 (DE = 5,9)
	Bipedestación (media en meses)		24,8 (DE = 8,7)
	Primeras palabras (media en meses)		22,8 (DE = 13,8)

En el grupo de niños con DT fueron incluidos niños sin factores de riesgo asociado, que tenían mediciones de crecimiento dentro del rango esperado, según la evaluación del programa Chile Crece Contigo. Se verificó

que los niños que formaron este grupo alcanzaran rendimientos totales en TADI ubicados dentro del rango promedio.

Instrumentos

Test de Aprendizaje y Desarrollo Infantil (TADI; Edwards & Pardo, 2013). TADI es una escala de desarrollo para niños desde 3 meses hasta 6 años que evalúa las dimensiones de Cognición (52 ítems, por ejemplo, el evaluador observa si el niño logra mostrar su edad con los dedos), Motricidad (47 ítems, por ejemplo, el evaluador entrega tres cubos al niño para que construya una torre), Lenguaje (55 ítems, por ejemplo, el evaluador observa si el niño sigue una orden compuesta) y Socioemocional (56 ítems, por ejemplo, el evaluador observa la conducta del niño al separarse de sus padres). La calificación se realiza a nivel de ítem, otorgando 1 punto si el niño *logra la tarea* presentada y 0 puntos si *no la logra*. La escala incluye, además, algunas preguntas para los cuidadores respecto a la presencia o ausencia de determinadas conductas (por ejemplo, se pregunta al cuidador si el niño elige qué quiere jugar o dónde ir como parte de la escala Socioemocional). TADI entrega resultados brutos y posición del niño en rangos de edad cronológica. La transformación de datos brutos genera puntajes T por dimensión y para la escala total. Su aplicación estandarizada en niños con DT toma aproximadamente 25 minutos; se inicia en los ítems esperados para la edad y hay suspensión después de tres fallas consecutivas.

Escala Wechsler de Inteligencia para Niños, tercera edición-versión chilena (WISC-III v.ch.; Wechsler, 1991; Ramírez & Rosas, 2007). Se procuró la aplicación del WISC-III v.ch a todos los niños con SD mayores de 6 años. El WISC-IIIv.ch. tiene trece subpruebas organizadas en dos escalas y cuatro índices factoriales. Se administra a niños de 6:0 a 16:11 años. Arroja puntajes para cada subprueba, para los índices factoriales y un valor integrado que corresponde al coeficiente intelectual. La consistencia interna para la escala total es de 0,87 y posee apropiada evidencia de validez para población chilena (Ramírez & Rosas, 2007)

Leiter International Performance Scale-Revised (Roid & Miller, 1997). Se incorporó como criterio externo una lista de chequeo basada en observación de conducta que fue adaptada de esta escala y que ha sido previamente utilizada en dos muestras de niños chilenos en edad preescolar (Strasser & Del Río, 2014; Yoshikawa et al., 2015). Este instrumento explora dos dimensiones de la conducta que son control atencional (observando, por ejemplo, la conducta de atención hacia las instrucciones y demostraciones con cuatro niveles posibles que van desde el mejor rendimiento descrito como: “mira los materiales con atención para comprender, pone atención y sigue las instrucciones del evaluador”, hasta el rendimiento de menor calidad descrito como: “pasa la mayor parte del tiempo sin prestar atención”) y conducta inhibitoria (explorado, por ejemplo, desde la observación de la modulación y regulación de sus emociones con el mejor desempeño asociado al descriptor: “es altamente regulado, nunca se pone triste ni frustrado” hasta el rendimiento más pobre descrito como: “se altera mucho y tiene dificultad en lograr el autocontrol), cada una a través de cinco ítems. Todos los ítems tienen cuatro alternativas para tipificar la conducta observada y se califica otorgando cuatro puntos a la mejor conducta y un punto a la conducta menos deseable. Las habilidades exploradas incluyen atención focalizada y sostenida, control del impulso, cooperación, interés hacia las actividades, control emocional y actividad motora. Se calcularon tres variables de interés, incluyendo control inhibitorio (mín. = 5, máx. = 20), control atencional (mín. = 5, máx. = 20) y puntaje total (mín. = 10, máx. = 40). La consistencia interna de la escala, evaluada con alfa de Cronbach fue 0,902, ubicada en un rango excelente.

Cuestionario sobre características de la condición médica para el grupo de niños con SD. Este cuestionario se aplicó a los padres e incluyó preguntas sobre el tipo e intensidad de la estimulación recibida, presencia de factores de riesgo pre, peri y postnatales, hitos del desarrollo, condiciones médicas asociadas, tratamiento farmacológico recibido y antecedentes familiares.

Cuestionario socioeconómico. Este cuestionario aplicado a los padres incluyó preguntas sobre el nivel educativo de los padres, la comuna de residencia del niño y la dependencia del colegio al que asiste (municipal, particular subvencionado o particular pagado). La información socioeconómica sobre las familias de origen de los niños con DT, así como datos sobre su dependencia escolar, están disponibles en los registros ELPI.

Procedimiento

El protocolo de esta investigación fue visado por el Comité de Ética y Bioética de la Pontificia Universidad Católica de Chile, cumpliendo los estándares internacionales.

Se explicó a los padres contactados los objetivos, beneficios y riesgos de la investigación. Todos los padres invitados en los centros de Santiago y Rancagua aceptaron la invitación, mientras que en La Calera tres padres declinaron. Todos los niños de 3 años o más pasaron por un asentimiento con testigo.

La evaluación se realizó en dos sesiones. En la primera sesión se aplicaron los dos cuestionarios (socioeconómico y sobre la condición médica) a los padres y se inició la aplicación de TADI, y en la segunda sesión se terminó la aplicación de TADI. El 50% de las evaluaciones realizadas a niños con SD fueron grabadas en video. Debido a la dificultad para predecir el nivel inicial de los niños con SD a partir de su edad cronológica, se aplicó TADI desde el primer ítem a todos los niños y se continuó hasta que se alcanzó el criterio de suspensión, fijado para la muestra de desarrollo atípico en cinco errores consecutivos. El tiempo promedio de aplicación fue de 90 minutos, distribuidos en dos sesiones. Al finalizar el proceso de aplicación de TADI, el evaluador debió escribir una síntesis del comportamiento observado en cada niño, incluyendo comentarios que dieran cuenta de situaciones que hubieran llamado particularmente su atención. Los evaluadores completaron también las listas de chequeo para la evaluación de inhibición conductual y control atencional.

Para todos los niños con SD de 6 años o más se intentó la administración de WISC-IIIv.ch., para lo que se siguió el protocolo estandarizado de aplicación y suspensión. En todos los casos en los que el niño obtuvo un puntaje bruto de cero puntos en tres de las subpruebas, se suspendió la administración de la prueba, según el criterio indicado en el manual (Ramírez & Rosas, 2010; Wechsler, 1991).

Los niños con SD fueron evaluados por el equipo de investigación. La información sobre rendimiento de los niños con DT se extrajo desde la ELPI (2013).

En un curso no superior a tres meses se entregó a cada padre un informe de rendimiento del niño en TADI, retroalimentación que fue acordada como agradecimiento a la participación. En todos los casos se pagó el transporte de padres e hijos al lugar donde se realizaron las evaluaciones y se ofreció una colación. Las evaluaciones fueron realizadas en campus universitario, en escuelas y en fundaciones.

Análisis de Datos

Para efectos de análisis y presentación de datos, la muestra se organizó en tres subgrupos, de acuerdo con la edad cronológica. El Grupo 1 lo conforman niños entre 1:2 y 3:11 años, el Grupo 2 incluye niños de 4:0 a 5:11 años y el Grupo 3 considera a los niños entre 6:0 y 7:0 años. Se presentan análisis descriptivos de las variables para los grupos de niños con SD y con DT. Se calculó la consistencia interna de cada prueba y sus subescalas utilizando alfa de Cronbach para el grupo de niños con SD, con una descripción de rango basada en los criterios planteados por Nunnally & Bernstein (1994).

La evidencia de validez para los datos de niños con SD fue establecida a partir del análisis de datos descriptivos y matriz de correlaciones bivariadas, en la que se controló la edad de los sujetos.

El tamaño de muestra no permitió realizar un análisis factorial confirmatorio (AFC; Ullman & Bentler, 2013), proceso adecuado para verificar si la estructura teórica que se ve representada en los datos es similar entre niños con DT y sus pares con SD, evidencia relevante. Se realizó, entonces, un análisis factorial exploratorio (AFE) con los datos provenientes de los niños con SD, con el fin de verificar la carga de los ítems en las dimensiones propuestas y las relaciones entre estas. El AFE siguió el criterio de estructura simple (*very simple structure criterion*; Revelle & Rocklin, 1979). Se usó rotación varimax y fueron considerados como índices de ajuste χ^2/gl y su valor p correspondiente, la raíz del error cuadrático medio de aproximación (RMSEA) y la raíz del residuo cuadrático medio estandarizado (SRMR). Para verificar el nivel de adecuación del modelo, se consideró un índice χ^2/gl menor a 2, y un RMSEA y SRMR con valores de 0,08 o menos (Kline, 2011; Schumacker & Lomax, 2016).

Se realizó un análisis de varianza para la comparación de medias entre el grupo de niños con SD contra el grupo de niños con DT.

A fin de ofrecer información sobre la evidencia de validez basada en el proceso de respuesta, la información reportada por los evaluadores fue codificada y aquellas conductas que fueron reportadas por al menos el 75% de los evaluadores se consideraron significativas y se presentan en los resultados.

Resultados

En este apartado se presentan los resultados asociados a la evidencia de confiabilidad y validez de TADI obtenida en este estudio.

Respecto a la evidencia de confiabilidad de TADI para la muestra de niños con SD, se evaluó la consistencia internada basada en estimación de Alfa de Cronbach, procedimiento que se ajusta a los datos disponibles. Se registraron valores en rango *Excelente* para todas las dimensiones (Cognición = 0,919, Motor = 0,926, Lenguaje = 0,912 y Socioemocional = 0,863). Al analizar la consistencia por grupos de edad, los valores se ubican en términos generales en rango *Excelente* (Tabla 2).

Tabla 2
Alfa de Cronbach para cada Dimensión de TADI y Escala Total en Niños con Síndrome de Down, según Grupo de Edad

Dimensión	Grupo 1	Grupo 2	Grupo 3
Cognición	0,92	0,90	0,89
Motor	0,95	0,85	0,71
Lenguaje	0,83	0,90	0,90
Socioemocional	0,80	0,89	0,79
Total	0,96	0,96	0,93

Nota. Grupo 1 = 1:2 a 3:11 años, Grupo 2 = 4:0 a 5:11 años, Grupo 3 = 6:0 a 7:0 años.

En relación con la validez, se ofrece un análisis en función de evidencia basada en el contenido de la prueba, en la estructura interna y en la relación con otras variables (Lezak, Howieson, Bigler & Tranel, 2012; Mitrushina, Boone, Razani & D'Elia, 2005).

En cuanto a la evidencia basada en el contenido de la prueba, TADI se corresponde con lo esperado para una escala de desarrollo diseñada para la evaluación de niños con DT. Los autores de TADI reportan haberse basado en las Escalas Bayley y MacArthur para su diseño, y está documentada la cercanía general de varias escalas del desarrollo en términos de su estructura y contenido (Cirelli, Bickle Graz & Tolsa, 2015). La evidencia de contenido ha sido relacionada con la evidencia aparente. Este es un concepto que, si bien no se incluye dentro de los estándares del consorcio AERA, APA y NMCE (2014), se considera relevante para el uso clínico de los instrumentos (Nevo, 1985). La validez aparente se refiere a la correspondencia entre la impresión clínica y los resultados de la prueba, estando en estrecha relación con la evidencia de validez de contenido. En principio, TADI parece presentar una apropiada validez aparente, lo que implica el reporte de una consistencia general entre los resultados obtenidos por los niños evaluados y la impresión clínica de los evaluadores.

La evidencia basada en el proceso de respuesta se desprende del análisis de las sesiones de evaluación grabadas. En todos los casos de niños con SD evaluados se requirió más del doble del tiempo de aplicación de la escala de desarrollo, necesitándose realizar constantes maniobras clínicas para atraer la atención de los niños y facilitar su enganche con las tareas. La hipotonía que está presente en los niños con SD hizo que se requirieran algunas adecuaciones ambientales para conseguir que estos pudieran interactuar de manera apropiada con los materiales (e.g., uso de almohadillas de apoyo para mantener a los niños sentados). Fueron observadas con alta frecuencia conductas de hiperoralidad, dificultades para el seguimiento y sostenimiento de contacto ocular, sialorrea y necesidad de múltiples repeticiones y modelamiento de las acciones esperadas. Si bien fue necesario realizar estas maniobras clínicas, la estructura del ítem a ítem no sufrió ninguna modificación.

Si bien la evidencia con DT es adecuada, estudios previos han discutido la pertinencia de algunos ítems presentes en escalas de desarrollo cuando se utiliza el material con niños con SD, básicamente por ser ítems que exigen habilidades atencionales o lingüísticas que suelen estar comprometidas en estos niños (Moore et al., 2008). A fin de explorar empíricamente cuán adecuado es el contenido para niños chilenos con SD, en la Tabla 3 se ofrece la descripción en porcentaje de la cantidad de niños con SD de la muestra que, por grupo de

edad, pudieron enfrentar y responder correctamente cada uno de los ítems que forman la prueba. Los datos sugieren que la prueba no presenta efecto piso ni efecto techo en la muestra de niños con SD, para ninguna de las escalas de TADI, el porcentaje de logro va disminuyendo a medida que se incrementa la edad y ninguno de los niños de la muestra consiguió responder a la totalidad de los ítems presentados.

Tabla 3
Porcentaje de Logro del Total de la Muestra de Niños con SD a cada Ítem que Forma las Cuatro Dimensiones Evaluadas de TADI

Cognición		Motor		Lenguaje		Socioemocional					
Ítem	%	Ítem	%	Ítem	%	Ítem	%				
Grupo 1	1	100	Grupo 1	1	100	Grupo 1	1	100	Grupo 1	1	100
	2	100		2	100		2	100		2	100
	3	100		3	100		3	100		3	100
	4	100		4	100		4	100		4	100
	5	100		5	100		5	100		5	95
	6	100		6	100		6	100		6	100
	7	100		7	100		7	100		7	84
	8	95		8	100		8	100		8	98
	9	98		9	95		9	98		9	95
	10	98		10	100		10	98		10	98
	11	98		11	95		11	100		11	98
	12	100		12	98		12	95		12	82
	13	95		13	95		13	98		13	100
	14	77		14	91		14	93		14	98
	15	86		15	95		15	93		15	98
	16	89		16	91		16	89		16	95
	17	80		17	89		17	89		17	100
	18	80		18	91		18	91		18	93
	19	70		19	84		19	80		19	93
	20	77		20	82		20	68		20	100
	21	66		21	70		21	66		21	93
	22	70		22	80		22	77		22	98
	23	52		23	68		23	66		23	84
	24	57		24	52		24	55		24	89
	25	48		25	61		25	41		25	82
	26	14		26	52		26	41		26	70
	27	41		27	70		27	39		27	84
	28	45		28	20		28	45		28	86
	29	41		29	41		29	30		29	73
	30	16		30	23		30	18		30	55
	31	36		31	27		31	2		31	57
	32	20		32	20		32	14		32	32
	33	9		33	18		33	9		33	34
	34	7	Grupo 2	34	16		34	2		34	16
	35	5		35	23		35	2		35	36
Grupo 2	36	5		36	23		36	5		36	23
	37	2		37	0	Grupo 2	37	7		37	23
	38	0		38	7		38	7		38	9
	39	5		39	5		39	0	Grupo 2	39	9
	40	2		40	0		40	0		40	7
	41	2		41	2		41	2		41	5
	42	2		42	0		42	0		42	0

(continúa)

Tabla 3 (conclusión)
Porcentaje de Logro del Total de la Muestra de Niños con SD a cada Ítem que Forma las Cuatro Dimensiones Evaluadas de TADI

Cognición		Motor		Lenguaje		Socioemocional		
Ítem	%	Ítem	%	Ítem	%	Ítem	%	
43	2	43	0	43	2	43	5	
44	0	Grupo 3	44	0	44	2	44	0
45	2	45	0	45	2	45	2	
46	0	46	0	46	0	46	0	
47	0	47	0	47	0	47	0	
48	2			48	2	48	0	
Grupo 3	49			49	0	49	0	
	50			50	0	50	0	
	51			Grupo 3	51	0	51	0
	52			52	0	Grupo 3	52	0
				53	0	53	0	
				54	0	54	0	
				55	0	55	0	
						56	n/a	

Nota. El grupo 1 está compuesto por niños entre 1:2 y 3:11 años, el grupo 2 está compuesto por niños de 4:0 a 5:11 y el grupo 3 está compuesto por niños entre 6:0 y 7:0 años.

Como parte de la evidencia basada en el contenido de la prueba, los estándares actuales solicitan la referencia explícita al fundamento de imparcialidad de la prueba. A este respecto, la observación cualitativa realizada durante el uso clínico de la escala, indica que las dificultades en el desarrollo de lenguaje, tanto comprensivo como expresivo, observadas en algunos niños con SD pueden afectar su rendimiento en esta escala. En este estudio se observó que en los ítems iniciales que forman las cuatro escalas hay mayor imparcialidad, dada la simpleza en las instrucciones, la presencia de mejores posibilidades de modelamiento y la posibilidad de ofrecer respuestas basadas en pautas motoras no orales. Al avanzar en la aplicación de la prueba, se observa que la demanda sobre el lenguaje comprensivo es superior, pues las instrucciones son más complejas, así como también se requieren mayores habilidades expresivas para conseguir ofrecer las repuestas esperadas. Esta interferencia del lenguaje sobre el rendimiento se observa especialmente en la escala Socioemocional.

Respecto a la evidencia basada en la estructura de la prueba, se probó la organización de los ítems en las dimensiones sugeridas, además de la relación entre dichas dimensiones, entendiéndose que, teóricamente, la prueba permite agrupar los ítems en cuatro dimensiones que deberían tener baja relación entre sí. A fin de controlar el efecto de la edad sobre las relaciones de constructo, en la Tabla 4 se presenta la matriz de correlaciones bivariadas, controlando esta variable, con información separada para el grupo de niños con SD y para el grupo de niños con DT. El análisis de matrices sugiere que hay patrones de relación diferenciados entre las dimensiones en los dos grupos de interés. En la matriz observada para los datos de niños con SD aparece una relación alta, positiva y significativa entre las dimensiones Cognición y Lenguaje, como también entre Lenguaje y Socioemocional. Por el contrario, los resultados de la dimensión Motora no correlacionan con los otros componentes de la matriz. En los resultados de niños con DT aparecen relaciones significativas entre todas las variables, aunque la fuerza de estas es menor.

En la matriz rotada del AFE se observa que los ítems de la prueba se organizan en una estructura de un factor, sin diferencias en las cargas en los cuatro factores esperados, según la estructura de la prueba observada en la muestra de niños con DT. El índice de complejidad del criterio de estructura simple alcanza su valor máximo para la solución de tres factores (VSS = 0,93); los índices de ajuste para esta solución son: $\chi^2(19698, N = 48) = 0$, $\chi^2/g1 = 0$, $p = 1$; RMSEA = 0,038; SRMR = 0,079.

Tabla 4
Matriz de Correlaciones entre las Dimensiones de TADI para la Muestra de Niños con SD y con DT

	Motor	Lenguaje	Socioemocional
<i>Niños con SD</i>			
Cognición	0,24	0,83*	0,64*
Motor		0,13	0,10
Lenguaje			0,70*
<i>Niños con DT</i>			
Cognición	0,63*	0,54*	0,48*
Motor		0,56*	0,49*
Lenguaje			0,59*

* $p < 0,01$

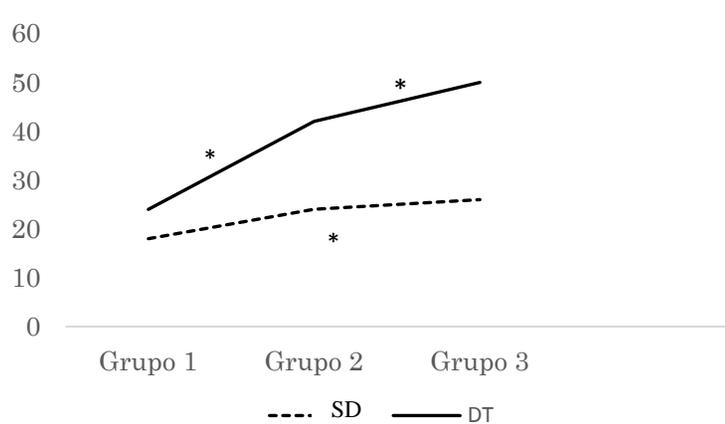
Se ha establecido también, desde el nivel teórico, que el rendimiento de los niños en este tipo de escalas debe mejorar con el paso del tiempo, pues van avanzando en los logros de su desarrollo. En el caso de las escalas de desarrollo, dicha progresión debería verse reflejada en un aumento progresivo en puntajes brutos. Los resultados observados en este estudio muestran un avance en dichas puntuaciones, con cambios estadísticamente significativos entre cada grupo de edad, tanto para SD como para DT (Tabla 5). En el grupo de niños con DT se registran diferencias estadísticamente significativas entre los grupos de edad en cognición, $F(2, 48) = 92,484$, $p < 0,001$, $\eta^2 = 0,804$, en desarrollo motor $F(2, 48) = 71,216$, $p < 0,001$, $\eta^2 = 0,760$, en lenguaje, $F(2, 48) = 84,670$, $p < 0,001$, $\eta^2 = 0,794$, y en desarrollo socioemocional, $F(2, 48) = 69,945$, $p < 0,001$, $\eta^2 = 0,757$. Por su parte, la comparación de medias en el grupo de niños con DS arrojó diferencias estadísticamente significativas entre los grupos de edad en cognición, $F(2, 48) = 5,002$, $p = 0,011$, $\eta^2 = 0,182$, en desarrollo motor, $F(2, 48) = 4,08$, $p = 0,023$, $\eta^2 = 0,154$, en lenguaje, $F(2, 48) = 4,41$, $p = 0,018$, $\eta^2 = 0,164$, y en desarrollo socioemocional, $F(2, 48) = 3,66$, $p = 0,034$, $\eta^2 = 0,143$.

Para las cuatro dimensiones analizadas, se producen interacciones estadísticamente significativas entre grupo y edad, con un mejor rendimiento en niños con DT en los rangos de edad mayores. La información estadística de este resultado se presenta incorporada en la figura 1.

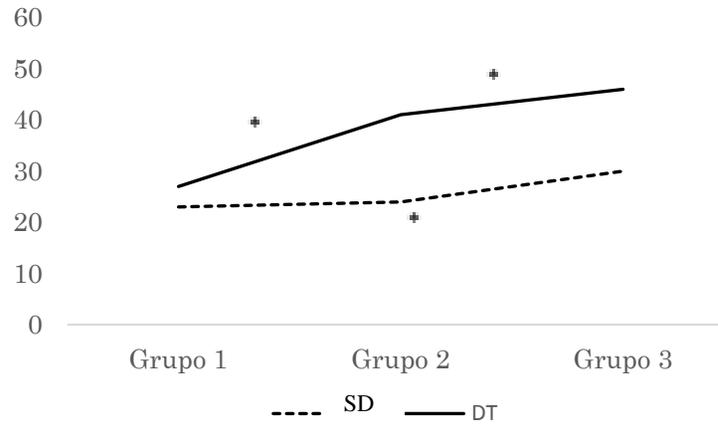
Tabla 5
Promedio y Desviación Estándar en las Cuatro Dimensiones de TADI en Niños con SD y con DT, por Grupo de Edad

	Cognición	Motor	Lenguaje	Socioemocional
<i>Niños con SD</i>				
Grupo 1	18,000 (5,3) [14,442-21,558]	22,727 (15,3) [18,917-26,538]	19,000 (3,3) [15,481-22,519]	23,300 (8,7) [19,274-27,326]
Grupo 2	23,563 (5,3) [20,612-26,513]	24,313(3,9) [21,153-27,472]	23,250 (4,8) [20,332-26,168]	29,125 (4,9) [25,943-32,307]
Grupo 3	25,714 (7,8) [23,139-28,289]	30,143(3,2) [27,385-32,901]	25,857 (7,9) [23,310-28,404]	30,86 (8,0) [28,079-33,635]
<i>Niños con DT</i>				
Grupo 1	24,000 (7,6) [20,442-27,558]	27,091 (7,3) [23,281-30,901]	25,800 (7,6) [22,109-29,491]	30,000 (6,2) [26,162-33,838]
Grupo 2	41,563 (5,3) [38,612-44,513]	40,625 (4,0) [37,466-43,784]	42,125 (6,8) [39,207-45,043]	45,500 (6,3) [42,318-48,682]
Grupo 3	50,476 (3,2) [47,901-53,051]	46,048 (1,3) [43,290-48,805]	53,238 (2,3) [50,691-55,785]	53,667 (3,8) [50,889-56,445]

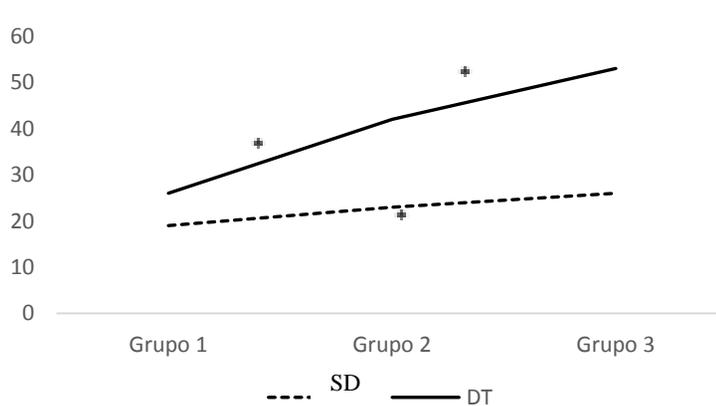
Nota. El grupo 1 está compuesto por niños entre 1:2 y 3:11 años, el grupo 2 está compuesto por niños de 4:0 a 5:11 y el grupo 3 está compuesto por niños entre 6:0 y 7:0 años. Los intervalos de confianza están calculados al 95%.



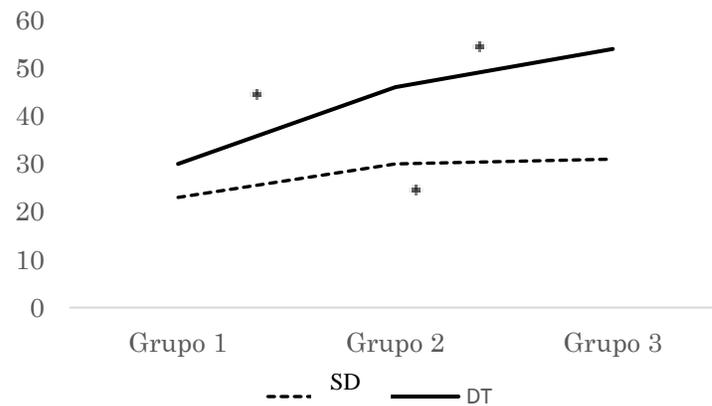
(a) Cognición ($F=18,013$, $p<,001$)



(b) Motor ($F=7,202$, $p<,001$)



(c) Lenguaje ($F=21,662$, $p<,001$)



(d) Socioemocional ($F=11,122$, $p<,001$)

Figura 1. Comparaciones de rendimiento en puntajes brutos de TADI entre el grupo de niños con SD y el grupo de niños con DT.

Para la evaluación de validez en relación con otras variables se intentó aplicar a los niños con SD mayores de 6 años el WISC-IIIv.ch., buscando la comparación con un instrumento estandarizado para la población chilena (con DT). Sin embargo, en los 21 niños con SD que tenían 6 o más años solo fue posible completar la evaluación de WISC-IIIv.ch., con estimación de CI total, en cuatro niños. En los casos restantes, 10 (47%) niños no pudieron seguir instrucciones básicas necesarias para ejecutar las tareas solicitadas y, en consecuencia, no fue posible realizar la aplicación; en siete (33%) niños fue posible aplicar algunas subpruebas, pero alcanzaron tres puntajes brutos con cero puntos y la aplicación debió ser suspendida por criterio clínico. Dado que no era parte del objetivo del estudio comparar a los niños en términos de su rendimiento en WISC-IIIv.ch., esta prueba no fue aplicada en este contexto a los niños con DT y se entiende que es de esperar que su ubicación se ordene en torno a la media, según datos poblacionales disponibles (Ramírez & Rosas, 2010).

El rendimiento promedio en CI total fue de 45 puntos y no se observaron diferencias entre las escalas e índices que componen la batería. Tal y como se reportó en una publicación anterior (Tenorio et al., 2014), los principales problemas observados al momento de realizar estas evaluaciones fueron dificultades en los niños para seguir instrucciones y limitaciones en el lenguaje en los niveles comprensivo y expresivo. Estos resultados impidieron hacer comparación de este instrumento con TADI.

Ante la falta de instrumentos de comparación para establecer evidencia de validez de relación con otras variables, este estudio incorporó el uso de una escala de observación que completó cada evaluador al finalizar la aplicación de TADI. Esta escala incluyó la evaluación de inhibición conductual y tono atencional. Al tratarse de dimensiones conductuales y no cognitivas, se esperó un bajo patrón de correlación entre el resultado de la escala (tanto total como en sus dimensiones de control atencional e inhibición conductual) y TADI. En los resultados se observa una matriz de correlaciones con relaciones significativas entre control atencional y la escala socioemocional de TADI, $r(48) = -0,512$, $p = 0,008$, y la escala total de criterio externo y la subescala socioemocional de TADI, $r(48) = -0,432$, $p = 0,027$.

Discusión

Este estudio ofrece un análisis de la evidencia de confiabilidad y validez, según el estándar internacional fijado por el consorcio AERA, APA y NMCE (2014) para TADI con una muestra de infantes y niños chilenos con SD.

En primera instancia, la evidencia sugiere que TADI es un instrumento con adecuada evidencia de confiabilidad para todos los rangos de edad incluidos en este estudio. Aunque los valores de consistencia interna tienen una disminución en el grupo de mayor edad, los mismos se mantienen en rangos *excelente* y *bueno*. Futuros estudios pueden ofrecer nueva evidencia de confiabilidad a partir de la aplicación test-retest, paso relevante para afinar esta evidencia.

Respecto a la validez, en los resultados se ha presentado evidencia basada en el contenido, en la estructura interna y en la relación con otras variables. En cuanto a la evidencia de validez basada en el contenido, la información disponible sugiere que TADI es una escala de desarrollo construida según las directrices internacionales y los ítems están formulados y organizados según lo esperado para un test como este. Para el caso específico del contenido en la muestra de niños con SD, los hallazgos de este estudio sugieren que se trata de una prueba más versátil para la muestra de niños más pequeños, en tanto permite mayor modelamiento y apoyo a la ejecución, mientras se complejiza en los más grandes, llegando a exigir un nivel de rendimiento que se ve interferido por dificultades en áreas cognitivas, como atención y lenguaje. En consecuencia, se sugiere a los profesionales proceder con precaución en la interpretación de resultados cuando se usa TADI en los rangos de edad mayor. Futuras versiones de las escalas nacionales de desarrollo deberán incluir formas de apoyo para una administración inclusiva.

En el análisis llama la atención que, en relación al porcentaje de éxito en la respuesta a nivel de ítem, ninguno de los niños con SD consigue enfrentar con éxito la totalidad de la escala. Esta observación tendría, al menos, dos interpretaciones posibles. Por un lado, puede ser una consecuencia observada en los resultados de un instrumento que no fue diseñado incorporando adecuaciones necesarias para esta población, creando barreras de acceso en los niños evaluados y viéndose con ello comprometido el rendimiento. Por otro lado, podría también interpretarse como un resultado vinculado a las dificultades intrínsecas del desarrollo asociadas al SD y presentes en la muestra evaluada. Más allá de esto, hay un punto positivo y es que, en ausencia de efectos piso y techo, la aplicación de TADI en estos niños permitió observar varianza y avanzar

en la generación de perfiles comprensivos individuales. Esta cualidad hace que TADI se profile como un instrumento clínicamente útil para la evaluación del desarrollo en niños con SD.

En la evidencia de validez basada en la estructura interna de la prueba, en términos generales los resultados obtenidos muestran diferencias en estructura al comparar el comportamiento que tiene TADI en niños con DT versus niños con SD. La presencia de matrices de correlación diferentes, así como también una organización estructural que parece ajustar mejor en tres y no en cuatro factores, podrían implicar que los supuestos teóricos que dan soporte a la prueba no se reflejan en su totalidad en estos niños. Sin embargo, es importante considerar que el análisis ideal para realizar comparaciones en este nivel es el AFC, el que no pudo ser realizado por el tamaño de muestra.

Sobre la presencia de diferencias en la matriz de correlaciones, el hallazgo de una relación positiva y significativa entre las cuatro dimensiones evaluadas por TADI en niños con DT versus una relación estrecha y positiva que excluye la dimensión de desarrollo motor en niños con SD resulta llamativa. Es un resultado de difícil interpretación, pero estudios previos en Chile han mostrado que las intervenciones tempranas para los niños con SD suelen enfocarse en el desarrollo motor, restando atención a las otras áreas, lo cual podría aportar a la interpretación de la disarmonía presente en el desarrollo y su reflejo y los resultados obtenidos (Arango et al., 2018).

Respecto a las diferencias en estructura factorial, características propias que parecen acompañar el desarrollo cognitivo y socioemocional de niños con SD y que han sido descritos en algunos estudios previos podrían ayudar a comprender las diferencias observadas. Aunque hay estudios que sugieren la presencia de un endofenotipo cognitivo propio para SD (e.g., Chapman & Hesketh, 2000; Fidler, Most & Philofsky, 2008), aún es necesario avanzar en el estudio desde la perspectiva dinámica que ofrece la teoría de trayectorias de desarrollo (Campos, 2018) y explicar la importante varianza intrasindromática presente en este grupo. En ese sentido, es posible sugerir que el resultado que se ha observado en este estudio puede ser consecuencia de un proceso de modularidad de la cognición que sigue un curso atípico en presencia del SD y, por ello, no es posible observar la misma estructura que en DT. Es posible que en niños con SD de mayor edad se llegue a una organización de los cuatro factores descritos para TADI.

Sobre las diferencias observadas entre los grupos de edad, el hallazgo sugiere que la prueba es sensible a las diferencias esperables por desarrollo, pero deben realizarse estudios longitudinales a futuro que den soporte a este hallazgo y que permitan documentar su uso para trazo de trayectorias clínicas individuales.

En cuanto a la validez basada en la relación con otras variables, la carencia de otras pruebas de evaluación diseñadas y validadas en niños chilenos con SD representa una barrera para este análisis. Fue una decisión no incorporar la aplicación de pruebas cortas de tamizaje de desarrollo que tienen un uso ampliado en Chile, pues se trata de instrumentos con baremaciones realizadas hace más de 20 años. Tampoco están disponibles protocolos o pruebas para la evaluación específica de las dimensiones consideradas en TADI. Aunque hay algunos desarrollos de instrumentos orientados a la evaluación específica de cognición en SD (de Sola et al., 2015; Edgin et al., 2010; Hessel et al., 2016) y algunas pruebas de instrumentos tradicionales en esta población (Galeote, Sebastián, Checa, Rey & Soto, 2011), estos han sido probados preferencialmente en adultos y/o no han sido utilizados en población latinoamericana. En consecuencia, estos instrumentos no son ideales para el análisis de relación con TADI en la muestra de interés. Hasta donde se registra en las bases de datos electrónicas con palabras claves (*síndrome de Down, evaluación, cognición*, en español e inglés), no hay otros instrumentos disponibles para evaluación de niños con SD en el mundo hispano. La carencia de otros instrumentos que permitan realizar una comparación apropiada y estimar el nivel de convergencia de la medición constituye no solo una de las más importantes limitaciones de este estudio, sino especialmente una de las limitaciones más importantes que enfrentan los profesionales que trabajan con esta población en Latinoamérica.

A nivel de resultados propios de este estudio, se destaca la evidencia divergente entre la medición de variables conductuales y los resultados de TADI, lo que sugiere que el resultado en la batería no está directamente influido por el nivel de desarrollo del control atencional y la inhibición conductual.

Más allá de las propiedades psicométricas específicas de TADI como prueba de medición del desarrollo, los resultados mismos llaman la atención. En todos los casos los niños con DT superan a niños con SD, lo que puede ser entendido como un hallazgo consistente con la investigación previa y esperable por la presencia misma de la condición médica. Sin embargo, además de esta diferencia observada en todos los casos, es llamativo que la distancia entre grupos se acentúe al aumentar la edad de los niños evaluados, al punto en que para varios casos no hay diferencias significativas entre SD y DT en el grupo de edad que reúne a los

más pequeños, mientras que hay diferencias entre todas las dimensiones en el grupo de edad que contiene a los niños mayores. Este estudio no puede explicar este resultado, pero es posible hipotetizar que proviene de una compleja interacción de factores que incluyen características propias del proceso educativo ofrecido a niños con SD en el país: expectativas parentales, oportunidades de inclusión y participación social y cívica, estigma social hacia personas con SD, entre otros. Estudios previos en otros países han demostrado que estos son factores que pueden impactar negativamente el curso de desarrollo en presencia de un desarrollo atípico (e.g., Kock et al., 2012). Futuros estudios deberán ampliar el estudio de estos factores sobre el desarrollo, bajo condiciones culturales propias del país.

Es necesario señalar que este estudio tiene limitaciones que se deben tomar en consideración al incorporar sus resultados en tomas de decisiones. La primera es el tamaño de la muestra, tanto en el número de niños participantes como en el amplio rango de edad incluido. El tamaño de muestra impidió conducir análisis de estructura interna que son relevantes y deberán ser resueltos en futuras investigaciones. Además de esto, contar con datos de niños con DT provenientes de la ELPI representó una ventaja en términos de diseño y costos del estudio, pero impidió el análisis a nivel de ítem. Dado que este estudio buscó ofrecer una aproximación metodológica frente al uso de un instrumento particular en una muestra determinada, es de esperar que futuros estudios incorporen las recomendaciones aquí planteadas y consigan muestras adecuadamente diseñadas para responder preguntas propiamente del desarrollo.

Hay una limitación evidente que tiene que ver con la falta de estudios de validez concurrente a partir de la comparación entre instrumentos. América Latina en general y Chile en particular tienen la tarea de avanzar en la construcción de una cultura de evaluación. Para ello, es necesario que equipos de investigación y desarrollo aporten con instrumentos modernos de evaluación y pruebas apropiadas en muestras especiales.

Es también una limitación del estudio la falta de análisis de validez prospectiva. Si bien el examen en un punto estático del tiempo permite tomar decisiones puntuales para el uso de TADI, son necesarios estudios de predicción cuyos resultados puedan ser incorporados en tomas de decisiones, por ejemplo, en el ámbito educacional.

Conclusión

Los resultados aquí presentados buscan contribuir al área de la evaluación psicológica en Chile, aportando nuevo conocimiento que permita avanzar hacia mejores prácticas profesionales. Situaciones como el uso de instrumentos obsoletos, las adaptaciones personales, las traducciones sin verificación, el uso de normas extranjeras o el uso de normas locales con niños inmigrantes deberán ir desapareciendo para mejorar la calidad de la tarea profesional del psicólogo.

La incorporación de conocimientos respecto a los alcances que tiene TADI en la evaluación de niños chilenos con SD que se ofrecen en este estudio, puede facilitar la toma de decisiones en profesionales médicos y no médicos, además de permitir monitorear el impacto de los programas de intervención que deberían orientarse desde la evidencia.

Referencias

- Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment of young children with autism spectrum disorders. *Child Neuropsychology*, *12*, 269-277. <https://doi.org/10.1080/09297040500473714>
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2014). *The standards for educational and psychological testing* (8^a ed.). Washington, DC: American Educational Research Association.
- Anderson, P. J., De Luca, C. R., Hutchinson, E., Roberts, G., Doyle, L. W., & Victorian Infant Collaborative Group (2010). Underestimation of developmental delay by the New Bayley-III Scale. *Archives of Pediatrics & Adolescent Medicine*, *164*, 352-356. <https://doi.org/10.1001/archpediatrics.2010.20>
- Arango, P. S., Aparicio, A., & Tenorio, M. (2018). Developmental trajectories of children with Down syndrome by socio-economic status: The case of Latin America. *Journal of Intellectual Disability Research*, *62*, 759-774. <https://doi.org/10.1111/jir.12516>
- Bayley, N. (1993). *Bayley Scales of Infant and Toddler Development* (2nd ed.). San Antonio, TX: Psychological Corporation.
- Bayley, N. (2005). *Bayley Scales of Infant and Toddler Development* (3rd ed.) (*Bayley-III*). San Antonio, TX: Pearson Clinical.
- Bedregal, P. (2008). Instrumentos de medición del desarrollo en Chile. *Revista Chilena de Pediatría*, *79*, 32-36. <https://doi.org/10.4067/S0370-41062008000700006>
- Bellman, M., Byrne, O., & Sege, R. (2013). Developmental assessment of children. *British Medical Journal*, *346*, artículo e8687. <https://doi.org/10.1136/bmj.e8687>
- Bravo, D. (2013). *Encuesta longitudinal de la primera infancia: primeros resultados 2012*. Santiago, Chile: Universidad de Chile, Centro de Microdatos.

- Byrne, A., MacDonald, J., & Buckley, S. (2002). Reading, language and memory skills: A comparative longitudinal study of children with Down syndrome and their mainstream peers. *British Journal of Educational Psychology*, *72*, 513-529. <https://doi.org/10.1348/00070990260377497>
- Campos, R. (2018). If you want to get ahead, get a good master. Annette Karmiloff-Smith: The developmental perspective [Si quieres avanzar, ten una buena maestra. Annette Karmiloff-Smith: la mirada desde el desarrollo]. *Infancia y Aprendizaje*, *41*, 90-137 <https://doi.org/10.1080/02103702.2017.1401318>
- Chapman, R. S. & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*, 84-95. [https://doi.org/10.1002/1098-2779\(2000\)6:2<84::AID-MRDD2>3.0.CO;2-P](https://doi.org/10.1002/1098-2779(2000)6:2<84::AID-MRDD2>3.0.CO;2-P)
- Cirelli, I., Bickle Graz, M., & Tolsa, J. F. (2015). Comparison of Griffiths-II and Bayley-II tests for the developmental assessment of high-risk infants. *Infant Behavior & Development*, *41*, 17-25. <https://doi.org/10.1016/j.infbeh.2015.06.004>
- de Sola, S., de la Torre, R., Sánchez-Benavides, G., Benjam, B., Cuenca-Royo, A., Del Hoyo, L. ... The TESDAD Study Group (2015). A new cognitive evaluation battery for Down syndrome and its relevance for clinical trials. *Frontiers in Psychology*, *6*, artículo 708. <https://doi.org/10.3389/fpsyg.2015.00708>
- Edgin, J. O., Mason, G. M., Allman, M. J., Capone, G. T., DeLeon, I., Maslen, C. ... Nadel, L. (2010). Development and validation of the Arizona Cognitive Test Battery for Down syndrome. *Journal of Neurodevelopmental Disorders*, *2*, 149-164. <https://doi.org/10.1007/s11689-010-9054-3>
- Edwards, M. & Pardo, M. (2013). *Test de Aprendizaje y Desarrollo Infantil (TADI)*. Santiago, Chile: Editorial Universitaria.
- Evers, A., Hagemeister, C., & Hostmaelingen, A. (2013). *EFPA Review Model for the description and evaluation of psychological and educational tests*. Brussels: European Federation of Psychology Associations. Recuperado de <http://www.efpa.eu/professional-development/assessment>
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J. ... Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA: Singular Publishing Group.
- Fidler, D. J., Most, D. E., & Philofsky, A. D. (2008). The Down syndrome behavioural phenotype: Taking a developmental approach. *Down Syndrome Research and Practice*. Publicación anticipada en línea. <https://doi.org/10.3104/reviews/2069> Extraído de https://assets.cdn.down-syndrome.org/pubs/a/reviews-2069.pdf?_ga=2.7675186.1422344357.1568399319-187576648.1568399319
- Galeote, M., Sebastián, E., Checa, E., Rey, R., & Soto, P. (2011). The development of vocabulary in Spanish children with Down syndrome: Comprehension, production, and gestures. *Journal of Intellectual and Developmental Disability*, *36*, 184-196. <https://doi.org/10.3109/13668250.2011.599317>
- García-Huidobro, J. E. & Bellei, C. (2003). *Desigualdad educativa en Chile*. Santiago, Chile: Universidad Alberto Hurtado.
- Gioia, G. A., Andrus, K., & Isquith, P. K. (1996). *Behavior rating inventory of executive function-preschool version (BRIEF-P)*. Odessa, FL: Psychological Assessment Resources.
- González, P., Mizala, A., & Romaguera, P. (2004). *Vouchers, inequalities and the Chilean experience*. Documento de Trabajo. Centro de Economía Aplicada (CEA). Universidad de Chile.
- Haeussler, I. M. & Marchant, T. (2003). *TEPSI: Test de Desarrollo Psicomotor 2-5 Años* (10ª ed). Santiago, Chile: Ediciones Universidad Católica.
- Hessl, D., Sansone, S. M., Berry-Kravis, E., Riley, K., Widaman, K. F., Abbeduto, L. ... Gershon, R. C. (2016). The NIH Toolbox Cognitive Battery for intellectual disabilities: Three preliminary studies and future directions. *Journal of Neurodevelopmental Disorders*, *8*, artículo 35. <https://doi.org/10.1186/s11689-016-9167-4>
- Kline, R. B. (2011). *Principles and practice of structural equation modelling* (3ª ed.). New York, NY: Guilford Press.
- International Test Commission. (2018). *Directrices internacionales para el uso de los tests*. Rescatado el 7 de noviembre de 2019 desde https://www.intestcom.org/files/guideline_test_use_spanish.pdf
- Kock, E., Molteno, C., Mfiki, N., Kidd, M., Ali, A., King, M., & Strydom, A. (2012). Cross-cultural validation of a measure of felt stigma in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, *25*, 11-19. <https://doi.org/10.1111/j.1468-3148.2011.00644.x>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5ª ed). New York, NY: Oxford University Press.
- Long, C. E., Blackman, J. A., Farrell, W. J., Smolkin, M. E., & Conaway, M. R. (2005). A comparison of developmental versus functional assessment in the rehabilitation of young children. *Pediatric Rehabilitation*, *8*, 156-161. <https://doi.org/10.1080/13638490400022287>
- Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). Use of methodological concepts in neuropsychology practice. En M. Mitrushina, K. B. Boone, J. Razani, J. & L. F. D'Elia (Eds.). *Handbook of normative data for neuropsychological assessment* (2ª ed.; pp.12-31). New York, NY: Oxford University Press.
- Moore, D. G., Goodwin, J. E. & Oates, J. M. (2008). A modified version of the Bayley Scales of Infant Development-II for cognitive matching of infants with and without Down syndrome. *Journal of Intellectual Disability Research*, *52*, 554-561. <https://doi.org/10.1111/j.1365-2788.2008.01064.x>
- Mullen, E. M. (1995). *Mullen Scales of Early Learning (AGS ed.)*. Circle Pines, MN: American Guidance Services.
- Nazer, J. & Cifuentes, L. (2011). Estudio epidemiológico global del síndrome de Down. *Revista Chilena de Pediatría*, *82*, 105-112. <https://doi.org/10.4067/S0370-41062011000200004>
- Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, *22*, 287-293. <https://doi.org/10.1111/j.1745-3984.1985.tb01065.x>
- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory* (3ª ed.). New York, NY: McGraw-Hill.
- Ramírez, V. & Rosas, R. (2007). Estandarización del WISC-III en Chile: descripción del test, estructura factorial y consistencia interna de las escalas. *Psykhe*, *16*, 91-109. <https://doi.org/10.4067/S0718-22282007000100008>
- Ramírez, V. & Rosas, R. (2010). *Test de Inteligencia para Niños de Wechsler WISC-III. Manual de administración y puntuación. Normas de estandarización chilena* (5ª ed.). Santiago, Chile: Ediciones Universidad Católica.
- Revelle, W. & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, *14*, 403-414. https://doi.org/10.1207/s15327906mbr1404_2
- Ringwalt, S. (Comp.) (2008). *Developmental screening and assessment instruments with an emphasis on social and emotional development for young children ages birth through five*. Chapel Hill, NC: National Early Childhood Technical Assistance Center.
- Rodríguez, S., Arancibia, V., & Undurraga, C. (1976). *EEDP: Escala de Evaluación del Desarrollo Psicomotor de 0-24 Meses*. Santiago, Chile: Galdoc.
- Roid, G. H. & Miller, L. J. (1997). *Leiter International Performance Scale-Revised*. Wood Dale, IL : Stoelting.

- Rojas-Barahona, C. A., Gaete, J., Olivares, E., Förster, C. E., Chandía, E., & Chen, M. -Y. (2017). Psychometric Evaluation of the Adolescent Health Promotion Scale in Chile: Differences by Socioeconomic Status and Gender. *The Journal of Nursing Research, 25*, 471-480. <https://doi.org/10.1097/JNR.0000000000000196>
- Rosas, R. & Pizarro, M. (2018). *WISC-V: manual de administración y corrección*. Santiago, Chile: CEDETi UC.
- Rosas, R. & Santa Cruz, C. (2013). *Dime en qué colegio estudiaste y te diré qué CI tienes: radiografía al desigual acceso al capital cognitivo en Chile*. Santiago, Chile: Ediciones Universidad Católica.
- Santos, H. & Elacqua, G. (2016). Socioeconomic school segregation in Chile: Parental choice and a theoretical counterfactual analysis (CEPAL Review). Recuperado del sitio de internet de CEPAL: <https://repositorio.cepal.org/handle/11362/40792>
- Schumacker, R. E. & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling* (4ª ed.). New York, NY: Routledge.
- Strasser, K. & Del Río, F. D. (2014). The role of comprehension monitoring, theory of mind, and vocabulary depth in predicting story comprehension and recall of kindergarten children. *Reading Research Quarterly, 49*, 169-187. <https://doi.org/10.1002/rrq.68>
- Tenorio, M., Campos, R. & Karmiloff-Smith, A. (2014). What standardized tests ignore when assessing individuals with neurodevelopmental disorders. *Estudios de Psicología, 35*, 426-437. <https://doi.org/10.1080/02109395.2014.922264>
- Ullman, J. B. & Bentler, P. M. (2013). Structural equation modeling. En J.A. Schinka, W. F. Velicer, & I.B. Weiner (Eds). *Handbook of psychology: Research methods in psychology*, (Vol. 2, pp. 607-634). Hoboken, NJ: John Wiley & Sons. <https://doi.org/10.1002/0471264385.wei0224>
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M. C., Weiland, C. ... Arbour, M. C. (2015). Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Developmental Psychology, 51*, 309-322. <https://doi.org/10.1037/a0038785>

Fecha de recepción: Marzo de 2018.

Fecha de aceptación: Marzo de 2019.