

**SÍNDROME DE DOWN
CUADERNOS DE ACTUALIDAD, 2**

**La medición de los resultados
en los estudios de investigación
sobre personas con
síndrome de Down**



FUNDACION
IBEROAMERICANA
DOWN21

Colección Cuadernos de Actualidad
Fundación Iberoamericana Down21
www.down21.org
Director: Jesús Flórez, M.D., Ph.D.

**La medición de los resultados en los estudios
de investigación sobre personas con síndrome de Down**

© Fundación Iberoamericana Down21

Edita: Fundación Iberoamericana Down21
Santander 2023
Realización: Consultores Initier S.L., Santander

PRESENTACIÓN

A nadie se le oculta que el avance definitivo en el conocimiento de lo que es el síndrome de Down, y sobre todo, de cómo son las personas que lo presentan y de su trayectoria a lo largo de la vida, pasa por realizar los estudios de investigación directamente en las personas con síndrome de Down. Los estudios en modelos animales, en las células aisladas o en organoides facilitan el camino para conocer en profundidad los mecanismos biológicos que están en la base de la alteración provocada por la presencia de un cromosoma extra. Pero cuando queremos profundizar en la persona, en el conocimiento de su real capacidad y de su limitación, de su potencialidad humana, de su recorrido vital, de sus comorbilidades, o de su respuesta a nuestras intervenciones para promover, potenciar, reparar o corregir, necesariamente tenemos que recurrir al estudio y análisis personal. Y ello requiere una metodología precisa que considere muy estrictamente la diversidad e individualidad que caracterizan a la población con síndrome de Down.

Esa metodología se encuentra recogida en las reglas de lo que denominamos ensayo clínico, en su más amplio sentido. El ensayo clínico mide los resultados de cualquier tipo de intervención sobre la salud, la cognición, la conducta, el bienestar, sea cual sea la modalidad de intervención que se aplique: educativa, conductual, farmacológica. Para ello es imprescindible acertar a seleccionar las técnicas que miden con precisión los resultados, es decir, las mediciones, utilizando los criterios que caracterizan al síndrome.

Los ensayos clínicos fallan si las mediciones seleccionadas no son apropiadas para las personas con síndrome de Down, o no tienen en consideración el fenotipo conductual que va asociado frecuentemente a él en una etapa determinada. El estudio que presentamos es una revisión sistemática que se centra en los problemas que plantea la medición de los resultados en el síndrome de Down, con un enfoque específico sobre las consideraciones que han de hacerse para evaluar los resultados relativos a dimensiones sustanciales de la persona, como son la cognición, el lenguaje y la conducta/psicopatología. Esta revisión también proporciona un resumen de las recomendaciones sobre la valoración de los resultados en estos dominios, así como recomendaciones para la futura investigación. Al mismo tiempo, analiza el impacto que tienen la salud física y la psicométrica de la evaluación sobre las mediciones de los resultados.

Las autoras son investigadoras muy acreditadas que han aportado un gran beneficio en el conocimiento sobre aspectos esenciales de la

personalidad de los individuos con síndrome de Down. Conscientes de la dificultad para llevar a cabo un análisis objetivo, han tenido en cuenta la problemática con la que ellas mismas se han encontrado, y ofrecen su experiencia para asegurar la medición de los resultados atendiendo no sólo a la necesidad de precisar qué se quiere medir y cómo, sino cómo elegir bien la muestra a analizar.

El presente estudio complementa el presentado en nuestro anterior Cuaderno. A la vista de la necesidad de seguir profundizando en las investigaciones, se hace preciso clarificar con objetividad la sistemática a seguir en cada proyecto, con el fin de que los resultados sean bien recogidos y las conclusiones resulten fiables.

Consideraciones sobre cómo medir los resultados individuales en los diversos contextos del síndrome de Down: Implicaciones en la investigación y en los ensayos clínicos

Considerations for measuring individual outcomes across contexts in Down syndrome: Implications for research and clinical trials

International Review of Research in Developmental Disabilities, Vol. 62: 192-225. 2022.
<https://doi.org/10-1016/bs.irrdd.2022.05.005>

AUTORES

Anna J. Esbensen^{a,b}, Emily K. Schworer^a, Deborah J. Fidler^c, Angela John Thurman^d

AFILIACIONES

^a Division of Developmental and Behavioral Pediatrics, Cincinnati Children's Hospital Medical Center,

Cincinnati, OH, United States

^b Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States

^c Human Development and Family Studies, Colorado State University, Fort Collins, CO, United States

^d University of California Davis Health, MIND Institute and Department of Psychiatry and Behavioral Sciences, Sacramento, CA, United States

SUMARIO

Resumen

Introducción

Problemas en la medición de resultados individuales en el síndrome de Down

Consideraciones sobre la evaluación basada en el perfil fenotípico asociado al síndrome de Down

1. Cognición

- 1.1 Selección de las pruebas de medición
- 1.2 Interpretar los resultados de la prueba de medición
- 1.3 Captar el cambio longitudinal
- 1.4 Recomendaciones para medir la cognición
- 1.5 Direcciones futuras para medir la cognición

2. Lenguaje

- 2.1 Selección de las pruebas de medición
- 2.2 Interpretar los resultados de la prueba de medición
- 2.3 Captar el cambio longitudinal
- 2.4 Recomendaciones para las medidas y evaluación del lenguaje
- 2.5 Direcciones futuras para medir el lenguaje

3. Conducta y psicopatología

- 3.1 Selección de las pruebas de medición
- 3.2 Interpretar los resultados de la prueba de medición
- 3.3 Captar el cambio longitudinal
- 3.4 Recomendaciones para la evaluación de la conducta y la psicopatología
- 3.5 Direcciones futuras para medir la conducta y la psicopatología

El impacto de la salud física sobre la evaluación de los resultados

1 Situaciones médicas

2 Habilidades motoras

3 Visión y audición

4 Recomendaciones y direcciones futuras para medir el impacto de la salud física a la hora de medir los resultados

Consideraciones complementarias sobre la evaluación en el síndrome de Down

1. El perfil fenotípico asociado al síndrome de Down: impacto sobre el funcionamiento

2. Interacción entre la ejecución y el contexto en el que es evaluada
3. Propiedades psicométricas de la evaluación
 - 3.1 Efectos suelo
 - 3.2 Validez
 - 3.3 Fiabilidad y efectos debidos a la repetición de la práctica
4. Dictamen del informante
5. Elegibilidad de los participantes en el estudio

Resumen de directrices futuras para medir resultados

Referencias

RESUMEN

Con creciente frecuencia las personas con síndrome de Down (SD) se ven implicadas en ensayos clínicos dirigidos a conocer los resultados en su desarrollo, como son la cognición y la conducta. El creciente interés por aplicar tratamientos a las personas con SD nos ha llevado a discutir actualmente sobre cómo seleccionar las técnicas que miden con precisión los resultados (mediciones), utilizando los criterios que caracterizan el síndrome. Su consideración está justificada ya que los ensayos clínicos fallan si las mediciones seleccionadas no son apropiadas para las personas con SD o no tienen en consideración el fenotipo conductual que va asociado frecuentemente a él en una etapa determinada.

Esta revisión se centra en los problemas que plantea la medición de los resultados en el SD, con un enfoque específico sobre las consideraciones que han de hacerse para evaluar los resultados relativos a la cognición, el lenguaje y la conducta/psicopatología. Esta revisión también proporciona un resumen de las recomendaciones sobre la valoración de los resultados en estos dominios, así como recomendaciones para la futura investigación. Se resumen también el impacto que tienen la salud física y la psicométrica de la evaluación sobre las mediciones de los resultados.

INTRODUCCIÓN

Las personas con discapacidades del desarrollo, incluido el síndrome de Down (SD), están siendo incorporadas cada vez más frecuentemente en la investigación, en los estudios de intervención y en los ensayos clínicos. Se han iniciado recientemente esfuerzos por parte de los *National Institutes of Health* de Estados Unidos (Proyecto INCLUDE) para incrementar y mejorar los apoyos en la atención y la calidad de vida de la personas con SD, promoviendo la investigación sobre nuevas formas de tratamiento, estudios en cohortes más amplias, y ensayos clínicos. Además, se han renovado propuestas para mejorar los resultados en el desarrollo, vida social y empleo a lo largo de su vida, con el interés específico de comprender mejor su fenotipo conductual y sanitario, y el impacto de las comorbilidades sobre la adaptación y las consecuencias que de ello se derivan (Esbensen et al., 2017; Gardiner et al., 2010; McCabe et al., 2011). Impulsados por estas nuevas iniciativas en la investigación sobre el SD, cobra aún mayor importancia la comprensión de la heterogeneidad con que se desarrollan los fenotipos conductuales y sanitarios en estos individuos, ya que esta información ha de tener sustanciales implicaciones a la hora de evaluar las consecuencias individuales y el impacto de las intervenciones.

Los hallazgos en la investigación con modelos animales de la trisomía 21y los estudios que se derivan de compuestos nutricionales (p. ej., el ácido fólico) han multiplicado el análisis de compuestos farmacológicos con potencial para mejorar la cognición y el lenguaje de las personas con SD (Heller et al., 2003, 2006; Kishnani et al., 2010; Mircher et al., 2020). El número de ensayos clínicos sobre fármacos en el SD se ha duplicado en la última década, con 17 ensayos entre 2000 y 2010, y 36 entre 2011 y 2021. Hasta la fecha, muchos ensayos clínicos no han mostrado resultados positivos. Es difícil identificar las razones por las que han fallado. Por ejemplo, los fármacos investigados pueden resultar ineficaces para el objetivo concreto, o para el tramo de edad elegido, o porque la duración del ensayo puede no haber sido suficiente para detectar los efectos. Se ha identificado como dificultad clave en los ensayos clínicos el método de medir un resultado (la medición), con problemas específicos en relación con las evaluaciones propensas a mostrar efectos suelo, o porque son impropias para el nivel de capacidad de los participantes, o son insensibles, o son poco fiables, o están limitadas debido al solapamiento de las exigencias de la medición (p. ej., mediciones de la cognición y del lenguaje) que confunden la interpretación de la ejecución.

Hay otros factores que complican más la interpretación de los resultados de un ensayo clínico. Están demostrando que los fármacos muestran su eficacia en sólo algunos de los individuos con SD, como ocurrió al suspender el ensayo de Roche (Moyer, 2021). Además, el problema se exagera si muchos posibles participantes son excluidos del ensayo por sus limitaciones en la cognición o en las habilidades comunicativas, o por sus condiciones conductuales o psicológicas, o por los rasgos de su condición física (d'Ardhuy et al., 2015; Kishnani, 2010). Aunque la exclusión de participantes con limitadas habilidades cognitivas y comunicativas puede asegurar que todos completen el protocolo del estudio, lo cierto es quedan excluidas personas que quizá fueran más indicadas para mejorar con un determinado fármaco. Y además, se sabe que la salud física (p. ej., las habilidades motrices, la audición, la visión) influyen sobre la expresión de las habilidades cognitivas y de lenguaje. Por eso los diseños de estudios con amplios criterios de exclusión pueden resultar problemáticos, ya que eso nos lleva a que las intervenciones farmacológicas aparezcan como exitosas, pero sólo sean eficaces para un subgrupo de personas.

Para evitar estos peligros, los investigadores clínicos deben prever los problemas de una evaluación y fiarse de expertos en el SD que les asesoren sobre la selección de unas mediciones que no restrinjan los criterios de elegibilidad. El asesoramiento experto debe basarse en nuestra comprensión más atinada de las características fenotípicas propias del SD en cada etapa de la vida, que expliquen el amplio espectro de su representación cognitiva y la posibilidad de que existan conjuntamente otras comorbilidades. Al hacerlo así, se asegurará la equidad en la investigación y el desarrollo de intervenciones dentro de la diversidad de los individuos con SD. Preparar los ensayos clínicos teniendo en cuenta estas premisas, incrementará también la probabilidad de que aparezcan resultados más certeros en los ensayos clínicos y reducirá el riesgo de falsos negativos/error tipo 2 (p. ej., se consiguieron efectos como resultado del tratamiento pero no se detectaron ni fueron informados).

Esta revisión: a) destaca los actuales problemas que están presentes en la medición de los resultados clínicos en los individuos con SD que pueden contribuir a falsos negativos en un ensayo, b) resume las consideraciones clave para medir estos resultados destacando el perfil cognitivo en el SD, y c) ofrece recomendaciones a seguir en la futura investigación. Resumiremos también los avances en la evaluación de los resultados en el SD y ofreceremos directrices futuras para progresar en esta área, que complementen los avances en los ensayos clínicos que se realicen en esta población.

PROBLEMAS EN LA MEDICIÓN DE RESULTADOS INDIVIDUALES EN EL SÍNDROME DE DOWN

El SD se encuentra asociado probabilísticamente a un fenotipo conductual que consiste en la presencia de relativos problemas en la cognición, en el lenguaje, y en aspectos específicos de la conducta y la salud física (Chapman y Hesketh, 2000; Daunhauer et al., 2014; Dykens et al., 2000; Esbensen y Mac Lean, 2017; Pulina et al., 2019). A pesar de que se hayan descrito patrones de retrasos en cada una de estas áreas a nivel de grupo, existe en los individuos con SD una variabilidad sustancial en los resultados dentro del mismo grupo (Karmiloff-Smith et al., 2016). La heterogeneidad dentro de los fenotipos cognitivos y conductuales, así como la de sus coexistentes problemas de salud, presentan todo un complejo conjunto de dificultades a la hora de medir de modo fiable y válido los resultados observados en las personas con SD. Por eso, es importante que las mediciones de los resultados sean evaluadas de manera específica para su utilización en esta población. Los investigadores han identificado muy pocas mediciones de resultados que hayan sido evaluadas empíricamente y sean psicométricamente seguras, para su utilización en los ensayos clínicos con personas con SD (Esbensen et al., 2017) o con poblaciones que tengan limitaciones significativas de comunicación y las correspondientes necesidades de ayuda.

En una investigación y una práctica que estén basadas en la evidencia, se exigen mediciones de los resultados que acrediten el actual estado del individuo, a menudo con el objetivo de definir si una determinada intervención del tipo que sea ha provocado un cambio objetivo. Las evaluaciones pueden estar basadas en la ejecución del individuo, en sus propios informes, en los informes del observador o del cuidador, o en los informes de un clínico. Para reducir los sesgos, las mediciones de los resultados necesitan ser fiables, válidas, sensibles al cambio con el tiempo, y mostrar variabilidad entre los individuos (Barry-Kravis et al., 2013; Esbensen et al., 2017). Han de ser relevantes para las cuestiones de investigación que se proponen, factibles y aceptables para el individuo o para el diseño del estudio, replicables, y elegidas antes de iniciar la intervención. Siempre que sea posible, las mediciones han de ser diseñadas de modo que permitan obtener respuestas mediante la utilización de formas diferentes de comunicación y que sean congruentes con los niveles cognitivos y habilidades adaptativas del individuo con SD.

En general, la selección de las mediciones para evaluar los resultados individuales depende del objetivo de la evaluación. En los con-

textos clínicos, los organizadores disponen de cierta flexibilidad a la hora de seleccionar las mediciones específicas para un determinado individuo y su presentación clínica. Sin embargo, en contextos de investigación, las decisiones sobre las mediciones han de basarse en la muestra en su conjunto que va a ser evaluada. Por eso es importante que los instrumentos que han de utilizarse sean apropiados y válidos para un margen más amplio de presentaciones fenotípicas. Debido a la escasez de mediciones de resultados psicométricamente sólidas para las personas con SD, se han utilizado criterios de inclusión y exclusión para un determinado estudio con el fin de alinear la población en estudio con los instrumentos disponibles. Y así será crítico el identificar y desarrollar las mediciones de resultados que sean psicométricamente apropiadas para utilizarlas en las personas con SD, con el fin de asegurar la equidad en la investigación.

En la selección de las mediciones de resultados, es también importante considerar de qué modo las mediciones elegidas se alinean con los mecanismos de la intervención. A menudo, se eligen estas mediciones para captar las dimensiones que se supone que son afectadas directamente por la intervención. Sin embargo, resulta beneficioso incluir mediciones de resultados que se supone que ejercen impactos secundarios, consecuencias, debidos a la intervención, o que su impacto puede moderar al impacto primario de la intervención. Por eso, la selección de determinadas mediciones en un estudio de tratamiento para el SD debería no sólo estar informada por la mayor probabilidad de que un fenotipo específico mejore o empeore, sino también por una visión más amplia que contemple la calidad de la vida diaria, y la posibilidad de que un tratamiento dirigido a un fin específico ofrezca efectos más vastos, en dimensiones más extensas del funcionamiento.

CONSIDERACIONES SOBRE LA EVALUACIÓN BASADA EN EL PERFIL FENOTÍPICO ASOCIADO AL SÍNDROME DE DOWN

La investigación desarrollada durante varias décadas sobre los resultados fenotípicos asociados al SD ha hecho posible seleccionar los sistemas de medición de resultados en formas más avanzadas e informativas del fenotipo. En esta sección resumimos las recomendaciones que son clave para la selección de los sistemas que miden la cognición, la comunicación y la conducta (Chapman y Hesketh, 2000; Dykens et al., 2000; Esbensen y MacLean, 2017; Pulina et al., 2019). Cada recomendación está adaptada o guiada para reducir el

error de la medición, eliminar elementos que confundan la interpretación, y capten los efectos del tratamiento con precisión. En la tabla 1 se expone un resumen de ejemplos representativos de las diversas mediciones alineadas en los dominios de la cognición, el lenguaje y la conducta/psicopatología.

Tabla 1. Ejemplos representativos de mediciones dentro de los dominios de cognición, lenguaje y conducta/psicopatología

Dominio	Medición	Edades
Cognición (CI)	KBIT-2	4-90 años
	SB-5	2-85+ años
	Leiter-3	3-75+ años
Cognición (subtests)	DAS-II Reconocimiento de imágenes)	2,5-17 años
	NEPSY-II (Precisión visomotora, Generación de palabras)	VP 5-12 años WG 5-16 años
	RBANS Aprendizaje de lista	12-89 años
	WISC-V extensión de imágenes	6-16 años
Cognición (prueba con computadora)	CANTAB (Aprendizaje de asociados emparejados, Cambio de conjunto intra-extra dimensional, Spatial span, Tiempo de reacción)	4 años y más
	NIH-TCB	3-85 años
Cognición (informe de padres)	BRIEF/BRIEF-2	Preescolar 2-5 años Escolar 5-18 años Adulto 18-90 años
	OMQPF	5-16 años
Lenguaje	Muestra de lenguaje expresivo	6-23 años
	CELF-P	3-6 años
	RBANS Nombrar imágenes y fluencia semántica	12-89 años
Conducta/Psicopatología	CBCL	Preescolar 2,5-4,5 años

SCQ
SRS/SRS-2

Escolar 4-18 años
Adulto 19 años y más
4 años y más
Preescolar 2,5-4,5
Escolar 4-18 años
Adulto 19 años y más

Nota: KBIT-2 = Kaufman Brief Intelligence Test, 2ª ed. SB-5 = Stanford-Binet, 5ª ed. CANTAB = Cambridge Neuropsychological Testing Automated Battery. NIH-TCB = National Institutes of Health-Toolbox Cognitive Battery. BRIEF = Behavior Rating Inventory of Executive Function. OMQPF = Observer Memory Questionnaire -Parent Form. DAS-II Differential Ability Scales, 2ª ed. RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; WISD-V = Weschler Intelligence Scale for Children, 5ª ed. CELF-P = Clinical Evaluation of Language Fundamentals Preschool; CBCL = Child Behavior Checklist. SCQ = Social Communication Questionnaire; SRS = Social Responsiveness Scale.

1. COGNICIÓN

El SD en general muestra relativamente buenas habilidades en el procesamiento visual, la memoria a corto plazo visuoespacial, las habilidades visuoespaciales y capacidad de imitación; todo ello coexiste con una clara discapacidad intelectual general (Gathercole y Alloway, 2006; Hodapp et al., 1992; Kanno e Ikeda, 2002; Klein y Mervis, 1999; Zoia et al., 2004). Sus relativos puntos débiles se aprecian en su memoria operativa y su memoria a largo plazo (Fidler et al., 2008; Hodapp y Dickens, 2004). Pero hay un alto grado de heterogeneidad entre los individuos en términos de presentación del fenotipo. Aunque la mayoría de los niños muestran retrasos cognitivos ya en la niñez más temprana, el grado de retraso varía mucho de un niño a otro (Grieco et al., 2015). Se observa esta heterogeneidad a todo lo largo de la vida, y eso tiene importantes implicaciones a la hora de seleccionar las mediciones de los resultados. De modo que para que haya una representación completa de personas con SD en un concreto estudio sobre tratamientos, las mediciones que se elijan deben captar las ejecuciones que ellas realicen teniendo en cuenta ese amplio rango en sus niveles de habilidades cognitivas.

Además, puesto que la velocidad de aprendizaje es más lenta que la de la población general, el diferencial de la capacidad de ejecución entre personas con y sin SD aumenta conforme avanza la edad cronológica. No obstante, las habilidades cognitivas indudablemente

progresan a lo largo de la niñez y adolescencia hasta entrada la vida adulta (Couzens et al., 2011). Después este desarrollo va seguido de un periodo de estabilización en las capacidades cognitivas y funcionales hasta la mitad de la adultez. Y pasados los 50 el riesgo de la demencia aumenta de forma dramática (Hawkins et al., 2003). Antes de los 40 años el riesgo de demencia está alrededor del 4% (Strydom et al., 2010), aumenta al 17% entre los 50 y 54, al 32 % entre los 55 y 59, y del 50 al 60% por encima de los 60 años (Coppus et al., 2006; Holland et al., 2000; Zigman et al., 1996). Puesto que hay aspectos relacionados con la etapa de la vida que afectan al funcionamiento cognitivo en el SD, la selección de una determinada medición debe estar también influenciada por la fase de desarrollo, la edad cronológica y el riesgo de declive por causa de la edad.

1.1 Selección de las pruebas de medición

Para captar adecuadamente teniendo en cuenta la heterogeneidad de habilidades cognitivas dentro del propio SD, se han de seleccionar las pruebas de medición que evalúen un rango amplio de niveles de habilidad (Karmiloff-Smith, 2016). Por ejemplo, aunque los signos físicos de demencia (placas y ovillos) aparecen en las etapas tempranas de la adultez y están presentes en la mayoría de los adultos con SD, no todos exhiben la presentación clínica de demencia, lo que confirma que las mediciones realizadas en la adultez deben ser capaces de cubrir y medir una amplia variedad de funcionamientos (Jack et al., 2013). Esta heterogeneidad tiene también implicaciones a la hora de medir los resultados individuales conforme los adultos pasan del estado de no-demencia al comienzo de los síntomas propios de la demencia. Por tanto, es esencial que se elijan las mediciones con una amplia escala de posibles puntuaciones capaces de evaluar ese rango amplio de actuaciones dentro de la misma prueba de medición.

También es importante saber elegir las técnicas de medición de resultados que sean apropiados para el tramo de edad cronológica de los participantes en un determinado estudio. Mediciones que son diseñadas para su uso en adolescentes y adultos en la población general pueden evaluar habilidades complejas del proceso informativo y, por tanto, llevarnos a efectos suelo si se usan en personas con SD. Al mismo tiempo, es importante no olvidar sus realidades madurativas. Por ejemplo, si se está trabajando con adultos, aunque las pruebas de medición propias del niño/adolescente pueden resultar apropiadas en algunos casos, es importante considerar hasta qué punto las vidas de

los adultos difieren de las de los niños en edad escolar, por lo que las pruebas habrán de adaptarse adecuadamente. Por ejemplo, se pueden adaptar las preguntas relacionadas con la vida escolar a las de la vida laboral. Saber establecer un equilibrio entre cuestiones que capten el nivel apropiado de desarrollo cognitivo, al tiempo que se incluyen también factores adaptados a la edad cronológica, es un desafío importante para establecer las mejores prácticas en la selección de una prueba de medición en el SD.

1.2 Interpretar los resultados de la prueba de medición

Al interpretar los resultados de la evaluación, sucede a menudo que las evaluaciones con propiedades psicométricas basadas en individuos de la misma edad cronológica que la de una persona con SD incluirán ítems demasiado difíciles para esta persona. Al mismo tiempo, tests con ítems más adecuados al grado de desarrollo pueden tener una información normativa ajustada a los individuos más jóvenes e impedirán la capacidad para derivar puntuaciones estandarizadas. Se han implementado también sistemas de puntuación basados en la edad de desarrollo, pero no se ha completado todavía una evaluación rigurosa de este método de puntuación.

1.3 Captar el cambio longitudinal

Cuando se evalúan las habilidades cognitivas a lo largo del tiempo, los métodos de medición tienen que ser sensibles tanto al crecimiento como al declive. En los jóvenes se esperan mejorías en el desarrollo. Sin embargo, dado que la velocidad de crecimiento en los individuos con SD es generalmente más lenta que en los que tienen desarrollo típico, una alta proporción de ellos pueden tener una puntuación suelo (es decir la puntuación más baja de acuerdo con las normas de la técnica de medición), lo que hace imposible describir la heterogeneidad y/o las mejorías en el desarrollo de estas personas. En el *Kaufman Brief Intelligence Test, 2ª edición (KBIT-2)*, el 37% de una muestra de niños con SD mostró puntuaciones CI estándar en el suelo del test (Edgin et al., 2017; Kauffman y Kauffman, 2004). Datos en bruto mostraban tasas menores de efectos suelo (5-9% de los participantes) pero no aportaban información sobre las actuaciones relativas a lo esperado según su edad cronológica. Estos problemas en relación con los efectos suelo en el KBIT-2 han sido replicados y destacan la necesidad de elegir mediciones del CI que sean más capaces de captar la heterogeneidad que

observamos en las personas con SD (Schworer, Esbensen et al., 2021). Dada la coexistencia de demencia en los adultos, también es importante que las técnicas de medición consideren el potencial declive del funcionamiento cognitivo. Si no se seleccionan bien los sistemas de medición de los resultados, pueden ser difíciles de interpretar los resultados conseguidos en los estudios que se dediquen a la intervención o a ensayos clínicos, y aparecerían resultados negativos debido a la inadecuada medición de resultados y no a la ineficacia del tratamiento.

1.4 Recomendaciones para medir la cognición

A causa del grado de problemas que la mayoría de las personas con SD experimentan en los dominios cognitivos, que son objeto lógico de las intervenciones farmacológicas y conductuales, las técnicas que miden la cognición aparecen en cabeza de los esfuerzos por validar las mediciones de los resultados. Se ha conseguido progresar significativamente en la medición del CI global mediante el empleo de las puntuaciones de desviación en el *Stanford-Binet, 5ª edición (SB-5)* (Roid, 2003; Sansone et al., 2014). La mayoría de las evaluaciones de la inteligencia, como es el SB-5, dan puntuaciones de CI que tradicionalmente tienen su puntuación más baja en 40. Puesto que un sustancial número de personas con SD funcionarán a, o por debajo de, esta cifra, el uso de los métodos tradicionales para generar la puntuación del CI enmascarará una variabilidad significativa del funcionamiento cognitivo. Un modo de resolver estos problemas utiliza los métodos propios de la desviación de la puntuación Z. Estos métodos utilizan la estadística descriptiva de los datos en bruto del subtest para cada banda de edad dentro de la muestra estandarizada, para generar una desviación de puntuación Z que ofrece una comparación más precisa del funcionamiento del participante en relación con el funcionamiento de la muestra estandarizada (Sansone et al., 2014). Se ha desarrollado una desviación en la puntuación Z para el SB-5, así como específicamente para los individuos con SD para los que se use la *Wechsler Intelligence Scale for Children, 4ª edición (WISC-IV)* (Pezzutti et al., 2018; Sansone et al., 2014). Esta técnica de puntuar la desviación elimina los efectos suelo propios de las mediciones CI tradicionales y, a diferencia de los datos en bruto, ofrece un índice de funcionamiento cognitivo en relación con la edad cronológica (Roid, 2013; Sansone et al., 2014). Por ejemplo, las puntuaciones de escala tradicionales en el SB-5 son propensas a mostrar un perfil plano en los individuos con discapacidad intelectual (con puntuaciones de 1 a lo largo de los sub-

dominios cognitivos). En lugar de agrupar todos los datos en bruto de 0-13 en una categoría de puntuación de escala de 1, los métodos de puntuación de la desviación captan las fortalezas y debilidades en los subdominios cognitivos. Un niño con puntuaciones de escala de 1 tanto en *Nonverbal Fluid Reasoning* como en *Verbal Knowledge* podría tener puntuaciones de desviación de -2,33 en *Nonverbal Fluid Reasoning* y -5,45 en *Verbal Knowledge*, demostrando una relativa fortaleza en habilidades no verbales y problemas en las habilidades verbales. Sin los efectos suelo en las mediciones del CI típicamente encontradas en las poblaciones con discapacidad intelectual, es posible describir las habilidades cognitivas más afinadamente en los participantes de un estudio.

Otra aproximación a más largo plazo para resolver el tema de los efectos suelo sería conseguir de los editores que extiendan el límite más bajo de las puntuaciones tradicionales del CI al establecer las normas del test. Por ejemplo, el *Leiter-3* tiene un suelo normativo más bajo que otras evaluaciones de CI (el más bajo es de 30), y muestra mínimos efectos suelo en una muestra de adolescentes y adultos con SD (1% en el suelo) (d'Ardhuy et al., 2015). Basándose en estas características, se ha recomendado para diseños de estudios que requieran evaluaciones de CI no verbal.

Otra importante y reciente innovación en el campo de las mediciones de resultados en el SD es la de incorporar la evaluación psicométrica de mediciones específicas que se propongan como candidatas, para determinar si son apropiadas para grupos específicos de edad en personas con SD. Se ha realizado recientemente un trabajo de validación para el *Arizona Cognitive Test Battery (ACTB)*, que ha sido desarrollado por Edgin et al. (2010) para evaluar la capacidad de las regiones hipocámpica, prefrontal y cerebelar en el SD (Edgin et al., 2010; Edgin et al., 2017). Se examinaron las mediciones a lo largo de un periodo de 3 meses de test-retest en niños y adultos (Edgin et al., 2017). Los resultados de este trabajo demostraron que ciertas tareas ACTB eran prometedoras, especialmente las que evaluaban el funcionamiento del cerebelo [el *Cambridge Neuropsychological Testing Automated Battery (CANTAB) Simple Reaction Time*, y la *NEPSY Visuomotor Precisión Task* (Edgin et al., 2017)]. Con todo, las tareas que evalúan el funcionamiento del hipocampo mostraron resultados ambiguos: el *CANTAB Paired Associates Learning* generó un funcionamiento variable en los efectos de fiabilidad test-retest y práctica, y el *CANTAB Intra-Extra Dimensional Set-Shift* presentó problemas en los efectos suelo (53%) y una fiabilidad test-retest inadecuada (Coeficiente de co-

rrelación intraclase [ICC] $<0,50$). Otra medición de resultados a partir de ACTB es el *CANTAB Spatial Span forward* que demuestra una buena fiabilidad test-retest $ICC > 0,70$ para la puntuación total, pero pobres propiedades psicométricas para la medición de errores (Edgin et al., 2017). La adecuada fiabilidad test-retest fue replicada en adolescentes con SD ($ICC = 0,67$; d'Ardhuy et al., 2015); sin embargo los índices de fiabilidad test-retest fueron menores para los adultos ($ICC = 0,55$; d'Ardhuy et al., 2015). Hubo también una tendencia hacia la identificación de los efectos por práctica (Edgin et al., 2017), lo que sugiere la necesidad de versiones alternativas para esta medición.

Más allá de la ACTB, se han hechos otros intentos para identificar las mediciones psicométricas más seguras para evaluar la cognición en el SD. En un estudio, se estudiaron las mediciones de los resultados del candidato en estudios de tratamientos en el SD a lo largo de un periodo de 6 meses en adolescentes y adultos. Este estudio mostró una buena fiabilidad test-retest ($ICC = 0,64 - 0,69$) en el *Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) List Learning* en adolescentes y adultos con SD (d'Ardhuy et al., 2015). En otro estudio, se evaluó un conjunto de evaluaciones clínicas de la memoria operativa en niños y adolescentes con SD (Schworer, Esbensen et al., 2021). Las mediciones específicas que cumplieron con la mayoría de los criterios psicométricos (p. ej., medir la terminación de $\geq 80\%$, fiabilidad test-retest $ICC \geq 0,75$, y efectos “práctica” no significativos) fueron la *SB-5 Working Memory* (verbal y no verbal), la *Differential Ability Scales*, 2ª edición (DAS-II) Reconocimiento de Imágenes, y el *WISC-V Picture Span*. En cada una de estas mediciones, las puntuaciones en bruto o habilidades generalmente mostraron mejor calidad psicométrica que las puntuaciones estándar, y los efectos suelo estuvieron por debajo del 20% de la muestra (Schworer, Esbensen et al., 2021).

Las evaluaciones de las mediciones cognitivas han incorporado también programas computarizados. La *National Institutes of Health Toolbox Cognitive Battery (NIH,TCB)* incluye mediciones de la memoria operativa, memoria episódica, rapidez de procesamiento, control de la inhibición, flexibilidad cognitiva, lectura oral y vocabulario receptivo, y ofrece ampliaciones relacionadas con el desarrollo que facilitan la viabilidad en la evaluación de las personas con discapacidad intelectual (Hessl et al., 2016; Shields et al., 2020). Se entiende la viabilidad como el porcentaje de niños examinados que pueden terminar la tarea para conseguir una puntuación válida. En recientes evaluaciones del conjunto de mediciones, esta viabilidad osciló entre el 66 y el 92% en niños y adolescentes con SD y síndrome X-frágil (Hessl

et al., 2016) y entre el 61 y 99% específicamente en el SDS (Shields et al., 2020). En conjunto, los resultados de la fiabilidad test-retest, validez del constructo y validez ecológica fueron prometedores de cara a utilizar mediciones computarizadas en futuros ensayos clínicos (Hessl et al., 2016; Shields et al., 2020).

Los cuestionarios a padres son ampliamente utilizables y las mediciones por informantes sobre las funciones ejecutivas y la memoria han servido de evaluación para niños y adultos con SD (d'Ardhuy et al., 2015; Esbensen et al., 2019). Se ha comprobado que el *Behavior Rating Inventory of Executive Function (BRIEF)* y el *Observer Memory Questionnaire-Parent Form (OMQPF)* son psicométricamente viables en niños y adolescentes con SD (d'Ardhuy et al., 2015; Esbensen et al., 2019; Schworer, Voth et al., 2022). Concretamente, el BRIEF generalmente demuestra una consistencia interna entre aceptable y excelente ($\alpha > 0,70$), y el OMQPF tiene una fiabilidad test-retest aceptable (ICC = 0,70; 0,76). El BRIEF-preescolar fue utilizado por d'Ardhuy et al. (2015) y se vio que no era apropiado para los adultos con SD, pero fue recomendado para los adolescentes.

1.5 Direcciones futuras para medir la cognición

Para ayudar a seleccionar de forma apropiada las técnicas que miden los resultados en el área cognitiva de las personas con SD, se necesita una mayor comprensión de la historia natural de su desarrollo cognitivo. Si bien este campo ha hecho avances para evaluar la diversidad de técnicas de medición asociadas a diversos constructos cognitivos, se necesita evaluar adicionales constructos cognitivos con mayor detalle, para garantizar la selección de las mediciones de resultados en futuros ensayos clínicos, particularmente en las primeras fases de la vida (Esbensen et al., 2017). Si bien los estudios anteriormente comentados han abordado las propiedades psicométricas de las técnicas de medición que valoran la cognición en niños mayores, adolescentes y adultos con SD, hay muy poco trabajo empírico que examine las propiedades psicométricas de las técnicas de medición de la cognición temprana en esta población. Además, es importante investigar los mejores abordajes para puntuar las evaluaciones estandarizadas cuando se utilizan para fuera del tramo normativo de edad. Se necesita investigar para: (a) comprender cómo la puntuación no estandarizada impacta sobre la interpretación, y (b) desarrollar datos normativos basados en los individuos con SD en donde resulte apropiado. Fuera del rango o intervalo, se puede calcular la puntuación basándose en el nivel de

desarrollo, pero tal como observó d'Ardhuy, no siempre esto puede ser apropiado. Por eso, necesitamos una mayor comprensión de cuándo, y para qué tramos de edad, es conveniente puntuar las mediciones estandarizadas basadas en el nivel de desarrollo del individuo.

2. LENGUAJE

Los niños con SD muestran retrasos en la adquisición de los hitos del lenguaje, en la mayoría de los dominios del lenguaje: fonología, vocabulario, gramática y pragmática (Martin et al., 2009; McDuffie y Abbeduto, 2009; Thurman y del Hoyo Soriano, 2021). Además, el grado de retraso en las habilidades lingüísticas es con frecuencia mayor en las personas con SD que en las que muestran otras formas de discapacidad intelectual (Chapman, 2006; del Hoyo Soriano et al., 2018; Finestak y Abbeduto, 2010; Finestack et al., 2013; Thurman et al., 2021).

Es importante señalar que con frecuencia se observan áreas de relativa fortaleza o de debilidad en todas las dimensiones del lenguaje. Las habilidades receptivas con frecuencia son superiores a las expresivas, las pragmáticas también suelen ser superiores a las de otras habilidades estructurales (p. ej., vocabulario y sintaxis), de modo que las de vocabulario suelen ser superiores a las sintácticas (Abbeduto et al., 2016). Sin embargo, incluso si se considera una de estas áreas, (p. ej., el vocabulario receptivo, la comunicación no verbal), pueden observarse patrones de fortaleza y de debilidad (Fidler et al., 2005; Finestack et al., 2013; Hahn et al., 2018; Phillips et al., 2014). No obstante, existe también una enorme variabilidad en los niveles absolutos de habilidad lingüística entre los diversos individuos. Las habilidades de lenguaje expresivo de los niños mayores y adultos varían desde un habla limitada hasta frases y conversaciones sintácticamente complejas (Finestack y Abbeduto, 2010; Klein y Marvis, 1999). Estos retrasos en las habilidades del lenguaje constituyen una enorme barrera para la independencia de los individuos; por eso, tienen alta prioridad los tratamientos que promuevan su desarrollo.

2.1 Selección de las técnicas de medición

Al medir específicamente las habilidades de lenguaje, una vez más se hace preciso considerar el posible rango de puntuaciones que se pueden alcanzar en una técnica de medición y el grado en que los efectos suelo y la restricción del tramo afectan al conjunto de los hallaz-

gos. Además, siempre que sea posible, habrán de utilizarse los tipos de medición que permitan utilizar el espectro de estilos propios de la comunicación, con el fin de asegurar que personas con SD con graves problemas en su habla queden también incluidas en el estudio.

En particular, las evaluaciones que se apoyan en instrucciones y respuestas verbales pueden ser de escasa utilidad para las personas que utilizan modos alternativos de comunicación. La evaluación de las habilidades lingüísticas en concreto, o de las habilidades en otros dominios (p. ej., el funcionamiento ejecutivo), puede verse apoyada si se permite al participante hacer uso de la Comunicación Aumentativa y Alternativa (CAA) o formas de comunicación diferentes del lenguaje verbal. En estos casos, tanto en las instrucciones como en las respuestas de los participantes, la CAA sin ayuda (p. ej., expresiones faciales, gestos o lenguaje de signos) o con ayuda (p. ej., figuras, objetos u otra aplicación) pueden reforzar la eficacia de la comunicación hasta que se adquiera el lenguaje hablado.

Aunque procedimientos fuera de lo estandarizado puedan aportar conocimiento para la utilización de apoyos que se puedan generalizar a otros contextos (p. ej., la casa, la escuela, centros laborales), desviarse de los procedimientos de evaluación supone un elemento de confusión si se usan datos normativos para realizar una comparación. Si se emplean, estos métodos de evaluación que no siguen lo estandarizado exigen estar documentados y ser publicados.

2.2 Interpretar los resultados de la evaluación

Para que la interpretación de las evaluaciones del lenguaje sean ajustadas, es de nuevo importante considerar tanto la edad como el nivel del desarrollo. Una evaluación puede ser apropiada al nivel de lenguaje desde el punto de vista del desarrollo, pero pueden surgir problemas relacionados con la edad cronológica. Aunque una evaluación normativizada sobre niños más pequeños sin discapacidad puede ofrecer la posibilidad de evaluar avances en el lenguaje en un estudio sobre intervención, el contenido de esa evaluación puede ser inapropiada desde el punto de vista del desarrollo para individuos de más edad, lo que prohibiría su uso.

2.3 Captar los cambios longitudinales

Las técnicas de medición para evaluar el lenguaje en el SD han de ser sensibles a su crecimiento y a su declive con el fin de detectar tanto los avances en su desarrollo como la posible regresión en la adolescencia o el declive de cognición/lenguaje en la edad adulta. Para captar el avance y el cambio en esta área debería también considerarse la etapa concreta de lenguaje (p. ej., preverbal, sólo palabras, combinación de palabras) y el nivel de desarrollo del niño o del adulto. Los indicadores del progreso en el desarrollo del lenguaje varían según sea la etapa del lenguaje, con aumentos en la frecuencia y tipos de las formas comunicativas prelingüísticas (p. ej., vocalización, gestos), que preceden a los aumentos en la frecuencia y tipos de las formas léxicas y gramaticales. Debido a su lentitud del desarrollo del lenguaje en comparación con sus pares, la etapa del lenguaje es particularmente importante a la hora de programar investigaciones sobre cambios a lo largo del tiempo. Además, el lenguaje es también sensible a los cambios que ocurren en los sistemas neurocognitivos que con él se relacionan y estas relaciones pueden cambiar también a lo largo del desarrollo. Por ejemplo, en las etapas tempranas del desarrollo, es probable que aumenten las faltas de fluidez conforme los niños empiezan a producir frases más largas y más complejas (Zachheim y Conture, 2003); como contraste, en posteriores etapas del desarrollo, los aumentos de faltas en la fluidez pueden indicar declives cognitivos (Faroqi-Shah et al., 2020).

2.4 Recomendaciones para las medidas y evaluación del lenguaje

Hasta la fecha, se han utilizado muchos sistemas de medición para evaluar el lenguaje de las personas con SD; sin embargo se necesitan más estudios todavía para comprender por completo las propiedades psicométricas de estas evaluaciones en la población SD (Esbensen et al., 2017). No obstante, se ha avanzado. Como ya hemos mencionado, el uso de las puntuaciones de desviación ha mejorado significativamente la medición del CI; con estos métodos, las puntuaciones del conocimiento verbal a partir de SB-5 ya no están sujetas a los efectos suelo, proporcionando así medidas afinadas de la ejecución y actuaciones en las personas con SD en comparación con sus pares con desarrollo ordinario, emparejados cronológicamente (Roid, 2003; Sansone et al., 2014).

Además, disponemos de nuevos datos que consideran la viabilidad y las propiedades psicométricas de la utilización del *Expressive Lan-*

guage Sampling (ELS) como instrumento para generar mediciones de los resultados en el lenguaje hablado en individuos con SD (Thurman et al., 2021). Más en concreto, al utilizar la metodología ELS para medir las habilidades lingüísticas en niños y jóvenes adultos con SD, los procedimientos fueron considerados factibles y produjeron variables para medir el vocabulario expresivo, la sintaxis y la inteligibilidad que fueron psicométricamente válidas (Thurman et al., 2021). Las tres mediciones mostraron una robusta validez de constructo. Es decir, estaban significativamente correlacionadas con otras variables diseñadas para medir el mismo constructo (demostrando validez convergente), y no estaban asociadas, o lo estaban en menor grado, con variables diseñadas para medir constructos distintos (demostrando validez de discriminación). Además, las tres puntuaciones demostraron la estabilidad test-retest, con mínimos efectos “práctica” y una poderosa fiabilidad test-retest (ICCs: 0,79-0,93) que fue apreciada a lo largo de un intervalo de 4 semanas. Los autores informaron que las tasas de no adhesión para estos procedimientos eran más probables para los jóvenes con habilidades cognitivas y lingüísticas más limitadas (jóvenes menores de 12 años que mostraron en el SB-5 un CI abreviado con puntuaciones equivalentes a 4 años, 9 meses, y un habla a nivel de frase o menos); para los jóvenes que no alcanzaban todos estos criterios, las tasas de no adhesión fueron del 5% o menos. Aunque quedan pendientes todavía algunos datos en relación con el grado en que las variables del lenguaje derivadas de los procedimientos ELS son sensibles al cambio en los individuos con SD, los datos preliminares que consideran la utilidad de tales procedimientos en los jóvenes sin discapacidad (Channell et al., 2018) y en pequeñas muestras de jóvenes con SD (Connors et al., 2018; del Hoyo Soriano et al., 2020) sugieren que las mediciones son sensibles al cambio a lo largo del tiempo.

D’Ardhuy et al. (2015) consideraron la utilidad de tres mediciones adicionales de la actividad lingüística en los adolescentes y adultos con SD: (1) *Clinical Evaluation of Language Fundamentals Preschool – 2 (CELF-P) subtest Word Classes* (Wiig et al., 2004), (2) *RBANS Picture Naming subtest* puntuación bruta, y (3) *RBANS Semantic Fluency subtest* puntuación bruta (d’Ardhuy et al., 2015). En relación con CELF-P, aunque se apreció una fiabilidad test-retest aceptable en adolescentes (ICC receptivo: 0,63; ICC expresivo: 0,71) y adultos con SD (ICC receptivo: 0,68; ICC expresivo: 0,69), además de una positiva asociación con el CI, no se observaron asociaciones significativas con la edad (d’Ardhuy et al., 2015). Se apreció además un cierto efecto techo para las puntuaciones receptivas en el CELF-P en personas

con SD. El *RBANS Picture Naming* no mostró fiabilidad test-retest adecuada ($ICC = 0,50, 0,53$) en adolescentes y adultos en el intervalo de 6 meses, aunque las puntuaciones mostraron asociación con el CI. Los resultados mostraron que el subtest de *RBANS Semantic Fluency* tenía una fiabilidad test-retest aceptable en el intervalo de 6 meses en adultos con SD, con un ICC de 0,73 y ningún efecto “práctica”; las puntuaciones mostraban también asociación con el CI. Se vio que las puntuaciones de *RBANS Semantic Fluency* declinaban con el tiempo. Además se ha demostrado que distingue a los pacientes con enfermedad de Alzheimer en una muestra comparada de adultos emparejada por edad, educación y sexo que no estaban diagnosticados de demencia (Duff et al., 2008). También se ha demostrado que una tarea similar de fluencia verbal semántica (*NEPSY II Word Generation*) tiene una fiabilidad adecuada (74-85%), moderada fiabilidad test-retest ($ICC = 0,72-0,75$) y mínimos efectos “práctica” en una muestra de jóvenes con SD (Smeyne et al., en prensa). Se necesitan más datos de esta puntuación sobre la trayectoria del desarrollo a lo largo de la adolescencia y la edad adulta en personas con SD, para validar la utilidad de esta técnica de medición como indicador de declive cognitivo en el SD.

2.5 Direcciones futuras para medir el lenguaje

Para caracterizar mejor la fiabilidad y utilidad de las maneras de medir los resultados del lenguaje, será útil en futuros estudios el considerar la adecuación psicométrica de tales métodos, con el fin de comprobar las influencias del nivel de lenguaje y la edad sobre las propiedades psicométricas. Ciertamente, conforme un individuo va pasando a lo largo de las etapas del lenguaje (p. ej., preverbal, primeras palabras, combinaciones de palabras), es probable que se observen diferencias en la velocidad e indicadores del crecimiento. Por ejemplo, una persona que está utilizando gestos y vocalizaciones, pero todavía no palabras, para comunicarse no tiene las habilidades fundamentales para participar en un test de habilidad semántica. Además, se necesita una mayor comprensión de las trayectorias del desarrollo por parte de las mediciones del lenguaje a lo largo de la vida para determinar el grado en que las actuales técnicas de medición son sensibles tanto al crecimiento propio del desarrollo como al declive cognitivo en la adultez tardía de muchas personas con SD. Así, pues, hay una gran necesidad de identificar formas de medición del lenguaje que aborden sus múltiples facetas y sean apropiadas a cada etapa de la vida.

3. CONDUCTA Y PSICOPATOLOGÍA

Como grupo, las personas con SD tienden a demostrar sus cualidades en específicas habilidades sociales (p. ej., la mirada, los gestos faciales, las interacciones sociales), y muestran en general menos problemas de conducta que sus pares con niveles similares de discapacidad intelectual (Fidler y Nadel, 2007). No obstante, si se comparan con la población general, o incluso con algunos otros individuos con discapacidad intelectual y del desarrollo, muestran tasas mayores de síntomas conductuales corrientes, incluidas la desobediencia o incumplimiento, la falta de atención, o el vagar errante (Dykens et al., 2000; Mantry et al., 2008). Además, pese a que hay áreas de relativa cualificación en la relación social, cognición social y comunicación, puede haber otras que son problemáticas para muchos (Abbeduto et al., 2016; Chanell, 2020). Está también elevado el riesgo de desarrollar el trastorno de déficit de atención con hiperactividad (TDAH) y el trastorno de espectro autista (TEA) en los niños con SD, en comparación con sus pares sin discapacidad intelectual o con otros tipos de discapacidad intelectual (Dekker y Koot, 2003; DiGuseppi et al., 2010; Ekstein et al., 2011; Froehlich et al., 2007).

3.1 Selección de las técnicas de medición

La mayoría de las evaluaciones sobre la conducta que se utilizan en el SD descansa sobre la notificación del informante, a menudo uno de los padres, el profesor, el cuidador. Al igual que en la selección de las técnicas de medición de la cognición y el lenguaje, las de la conducta exigen ser elegidas por su adecuación a la edad y al nivel de desarrollo de cada individuo. Debe considerarse también si la medición de la conducta ha sido desarrollada para individuos con desarrollo típico o con discapacidad intelectual, y si se necesitan datos normativos o disponibles para el estudio.

Al evaluar los síntomas de la psicopatología, las opciones de utilización son los informes de otra persona cercana o los auto-informes del propio individuo. Un problema con los informes del cuidador se debe a la dificultad de inferir pensamientos y sentimientos de otro individuo, como pueden ser las preocupaciones o ideas delirantes. Sin embargo, el método de los auto-informes también tiene sus problemas, incluida la necesidad de asegurar que la medición se ajusta al nivel de desarrollo tanto en el redactado del enunciado del ítem como en las opciones que se ofrezcan como respuesta. El auto-informe sobre estas dimensiones

depende de la teoría de la mente y de las habilidades en el vocabulario que hacen posible caracterizar el propio estado mental. Por ejemplo, un adulto con SD puede no comprender del todo lo que significa “estar deprimido” y puede ser necesario hacer preguntas más concretas sobre las conductas. En su lugar, los entrevistadores pueden preguntar a la persona con SD si se sienten tristes o la frecuencia con que han llorado en la pasada semana o mes. También puede ayudar a describir sus sentimientos y emociones el presentar imágenes de personas que muestran emociones diferentes. Además, si se usa un instrumento de CAA con sonido, será necesario que el interrogado tenga acceso a imágenes, botones o llaves que le sirvan para apoyar su expresión de emociones y sus estados mentales. Ciertamente, estos problemas en la evaluación de los síntomas de la psicopatología al utilizar auto-informes no son exclusivos del SD; son analizados con detalle en otro lugar (Finlay y Lyons, 2001).

3.2 Interpretación de los resultados de la evaluación

Es importante comprender la diversidad con que se presentan los síntomas que pueden aparecer en el SD y cómo influyen sobre la interpretación de los constructos conductuales y psicopatológicos que han de medirse. Es corriente que tengan soliloquios y amigos imaginarios y eso les dure toda la vida (Patel et al., 2018). Si bien es frecuente que tengan soliloquios, cambios en su nivel (volumen, intensidad) pueden indicar tensión o soledad (McGuire y Chicoine, 2006). Por eso, si se están midiendo las reacciones a un factor estresor, será importante tener en cuenta la intensidad del soliloquio (y no sólo la frecuencia o su presencia). En la población general, la medición del soliloquio se incluye con frecuencia a la hora de medir las psicosis y alucinaciones. Por eso, es obligado tener cuidado al evaluar las mediciones de los resultados en asegurarse de que una conducta propia del desarrollo en el SD sea calificada erróneamente como un síntoma psicopatológico.

3.3 Captar los cambios longitudinales

Se han desarrollado evaluaciones de la conducta y de la psicopatología con distintos propósitos. Las mediciones de la conducta se prestan a señalar las conductas que aparecen en ese momento concreto. Es de esperar que las mediciones realizadas para captar los cambios que van ocurriendo a lo largo del tiempo sean las adecuadas, pero esto debe todavía confirmarse con estudios empíricos. Sin embargo, las medicio-

nes apropiadas para personas con discapacidad intelectual y del desarrollo que se han elaborado van dirigidas más bien al diagnóstico y no a evaluar los cambios que acaezcan a lo largo del tiempo. Para hacer un diagnóstico clínico, estas formas de evaluación valoran los síntomas a lo largo de un periodo determinado, (p. ej., seis meses). En consecuencia, tanto el lenguaje de la enunciación del ítem que se ha empleado en la medición con fines diagnósticos como el espacio de tiempo evaluado por esa medición pueden contribuir a que sea insensible a los cambios temporales, si el espacio de tiempo entre evaluaciones o en un ensayo clínico es muy corto (p. ej., un cambio en la última semana/mes).

3.4 Recomendaciones para la evaluación de la conducta y la psicopatología

Se hace preciso prestar atención a la medición de la conducta y de la psicopatología en el SD si se quiere acertar en futuros ensayos clínicos que tengan como objetivo dichas dimensiones, o figuren como resultados primarios en otros dominios. La validación de las mediciones clínicas para usarlas en futuros ensayos clínicos será crítica para determinar si las mediciones utilizadas actualmente en la población general son las apropiadas para las personas con SD. Por ejemplo, la vigilancia temprana tipo TEA se ha convertido en el estándar de la atención a la población infantil general, con mediciones ampliamente utilizadas que seleccionan rasgos tempranos de comunicación social y conductas repetitivas. Sin embargo, muchas de las mediciones de las que disponemos actualmente incluyen ítems que en general se consideran parte del perfil fenotípico asociado al SD, como son los retrasos habituales en la comunicación. Un objetivo importante en el futuro para mejorar la detección precoz de riesgo de TEA en el SD, así como para medir la eficacia de los tratamientos del TEA, será el de distinguir entre esos rasgos que son específicos cuando coexisten SD y TEA, y qué rasgos son observados más generalmente en los que muestran SD, con independencia de su estado TEA. Tenemos el mismo problema en relación con el diagnóstico y detección del TDAH en el SD (Esbensen et al., 2021).

Las mediciones de conducta y psicopatología basadas en el informe de los padres son factibles y han demostrado su fiabilidad en esta población. Por ejemplo, se aceptó que el *Child Behavior Checklist (CBCL)* es un instrumento de cribado para los problemas globales de conducta en el SD (internalizadora, externalizadora, puntuación global), si bien la fiabilidad y validez fueron variables para las subescalas individuales, exigiendo cierta precaución si se utilizan las subescalas *Anxious/*

Depressed, Somatic Complaints y Thought Problems (Esbensen et al., 2018). Otra medición realizada mediante informe paterno, el *Social Communication Questionnaire (SCQ)* se utiliza con frecuencia como instrumento de cribado en TEA, y en datos piloto parece apropiado para niños y adolescentes con SD (Magyar et al., 2012). El SCQ mostró buena sensibilidad y especificidad para detectar el TEA, junto con alta validez convergente con el *Autism Diagnostic Interview-Revised* (Magyar et al., 2012). También se ha descrito el *Social Responsiveness Scale* como apropiado para describir los síntomas TEA en niños, adolescentes y adultos con SD (Channell, 2020; Channell et al., 2015; Schworer, Hoffman et al., 2021). Sin embargo, no se ha apreciado convergencia entre el SRS-2 y las evaluaciones directas de cognición social (Schworer, Hoffman et al., 2021). Por tanto, la información paterna puede estar evaluando una aplicación diferente de habilidades sociales en comparación con lo que se mide en el laboratorio. Para todas las mediciones de TEA, se necesitan más estudios para comprender qué modificaciones son exigibles a la hora de diagnosticar el TEA en las personas con SD.

3.5 Direcciones futuras para medir la conducta y la psicopatología

Se necesitan nuevos instrumentos de medición para evaluar los problemas de conducta y el fenotipo conductual más probable en el SD, así como para comprender las trayectorias de los perfiles de síntomas a lo largo del desarrollo (Patel et al., 2018). Es frecuente que varios problemas clave de conducta relacionados con el SD no sean captados de manera adecuada en las actuales mediciones de conducta de que disponemos, como son el andar vagando, la desobediencia, las conductas de autoestimulación, la insistencia en la repetición, limitaciones sociales y exceso de soliloquios (Capone et al., 2006; Feeley y Jones, 2008; McGuire y Chicoine, 2006; Patel et al., 2018; Stein, 2016). También se necesita incrementar los intentos por desarrollar y validar las mediciones de los auto-informes sobre la psicopatología. Como ya se ha comentado, hay cuestiones específicas en la intersección entre las ciencias del desarrollo y las evaluaciones diagnósticas que deben ser contestadas con el fin de mejorar nuestra medición de la sintomatología propia de esas realidades coexistentes como son el TEA y el TADH. En la actualidad no existe una información comprehensiva sobre la aparición temprana del TEA en el contexto de las primeras etapas del desarrollo en un SD, y no hay medios para distinguir entre rasgos fenotípicos tempranos asociados al SD y los indicadores tempranos de

riesgo de TEA. Del mismo modo, estamos empezando a saber más sobre las formas tempranas del control cognitivo en los bebés con SD (Schworer, Fidler y Danhauer, 2021; Schworrer et al., 2022), pero a día de hoy, no se ha establecido enlace alguno entre la variabilidad en la presentación del control cognitivo del bebé y la posterior aparición de coexistencia del TDAH. Mientras no se resuelvan estos temas, las actuales mediciones de conducta y psicopatología pueden no detectar, o pueden hacerlo erróneamente, áreas de preocupación. Como resultado, no se identificará a las personas con SD que necesiten un tratamiento efectivo, o puede que reciban un tratamiento ineficaz.

EL IMPACTO DE LA SALUD FÍSICA SOBRE LA EVALUACIÓN DE LOS RESULTADOS

1. SITUACIONES MÉDICAS

La expresión de los fenotipos cognitivos y lingüísticos en el SD se ve influida fuertemente por los problemas de salud que le acompañan. Son varios los que existen el SD. Las cardiopatías congénitas y los problemas de sueño afectan aproximadamente a la mitad de los individuos, influyen negativamente sobre el funcionamiento cognitivo y lingüístico, y explican en parte la variabilidad intra-síndrome en los resultados del desarrollo y en la adaptación (Breslin et al., 2014; Bull y Genética, 2011; Shott, 2006; Visootsak et al., 2011). Los niños tienen mayor riesgo de desarrollar una leucemia linfoblástica aguda, y se ha demostrado que impacta sobre la trayectoria del desarrollo en los niños con desarrollo ordinario (Bull y Genética, 2011). Tienen también mayor riesgo de presentar problemas gastrointestinales, hipotiroidismo y otras situaciones que impactan al desarrollo y a los resultados individuales (Bull y Genética, 2011; McCarron et al., 2005). Por tanto, es preciso tomar en consideración las comorbilidades a la hora de seleccionar los sistemas de medición y de interpretar sus resultados. También se necesita controlar de forma apropiada los cambios que ocurran en las comorbilidades y sus síntomas cuando haya que interpretar las mediciones de los resultados en la conducta, en la cognición y en el lenguaje, especialmente en los estudios que analicen los cambios a lo largo del tiempo.

Se hace preciso considerar las comorbilidades debido a su impacto al medir los resultados y al interpretar los resultados de la medición,

especialmente si se desea evaluar un cambio en los resultados previamente revisados de salud mental. Desde un punto de vista sintomático, son varios los problemas médicos que se presentan con rasgos similares a los de una depresión, y eso puede influir sobre la medición de los síntomas depresivos (Esbensen, 2020). Los síntomas del hipotiroidismo pueden aparecer en forma de letargia y de una reducción de interés en las actividades. Los síntomas de apnea del sueño pueden manifestarse como fatiga, irritabilidad, pérdida de interés. Los síntomas de la menopausia o el dolor se pueden manifestar como disminución de la actividad. Al interpretar los resultados de la salud mental, es importante tener en consideración estas mismas preocupaciones en relación con el solapamiento de síntomas que puedan aparecer por causas médicas (tiroides, aparato gastrointestinal, aparato respiratorio) o mentales.

2. HABILIDADES MOTORAS

A la hora de decidir sobre las mediciones de resultados han de considerarse también las habilidades motoras y la coordinación para asegurar la evaluación acertada de una determinada dimensión cognitiva o conductual. Con frecuencia las habilidades motoras se encuentran alteradas en las personas con SD debido a su bajo tono muscular y a sus dificultades con el equilibrio y la coordinación; a ello se suman otras dificultades en la planificación motora, el control y la fuerza muscular (Frank y Esbensen, 2015; Winders, 2013). Concretamente en las primeras etapas del desarrollo, las respuestas de carácter motor pueden constituir la única modalidad para evaluar la capacidad cognitiva y comunicativa. Por ejemplo, para valorar las primeras capacidades para planificar y resolver problemas, puede que se exija a los bebés con SD tirar de una tela para obtener un objeto que está fuera de su alcance (Schworer et al., 2020), y a los niños a sacar manualmente un objeto contenido en una caja (Fidler et al., 2005). Más adelante, también la coordinación de motor fino impactará la valoración de la rapidez de procesamiento cuando la tarea requiera la utilización de papel y lápiz. Esto es aplicable también a las tareas de iPad que exigen por parte del participante seleccionar respuestas mediante la pulsación en la pantalla (Hessl et al., 2016; Shields et al., 2020). En cada una de estas situaciones, las dificultades motoras pueden interferir en la medición válida de la cognición relacionada con la resolución de problemas o con conducta dirigida hacia un objetivo. Evaluar el estado de la motricidad, antes de utilizar este tipo de mediciones, será esencial para

interpretar correctamente las valoraciones cognitivas y comunicativas que emplean instrumentos que han de ser manipulados o exigen respuestas motoras, a lo largo de las diversas etapas de la vida.

Pero no deja de ser interesante que, pese a estas dificultades relacionadas con el desarrollo motor, el uso de sistemas que requieren manipulación facilita el aprendizaje de las personas con SD, y eso facilita la ejecución en las mediciones de los resultados del individuo (Buckley y John-Glenberg, 2008; Wolpert, 2001). Esta realidad es aplicable a los sistemas de manipulación virtual y tangible, ya que se ha demostrado que ambas modalidades contribuyen de manera eficaz al aprendizaje en niños con otros problemas del desarrollo como son los TEA (Bouck et al., 2014). Por tanto, una importante estrategia en la evaluación puede ser la utilización de instrumentos de manipulación que incrementen el interés en la tarea, la motivación y la ejecución de los sistemas de medición de los resultados.

Aunque se ha insistido en el papel de las habilidades del motor fino en la evaluación de la cognición y la comunicación, también las de motor grueso pueden limitar la capacidad de los participantes para manejarse en el mundo de las evaluaciones. Por ejemplo, en el *Autism Diagnostic Observation Schedule, 2ª ed. (ADOS-2)*, la evaluación de las iniciaciones sociales y otras conductas sociales esenciales puede verse entorpecida por la capacidad del niño para deambular. Mostrar un objeto al cuidador o al examinador llevándose lo, por ejemplo, es una conducta importante propia de la iniciación de la atención compartida. Si bien es posible hacer adaptaciones para los participantes con problemas de motor fino o de motor grueso, será importante tenerlas en cuenta a la hora de diseñar un estudio, con el fin de reducir elementos no intencionados que puedan confundir la evaluación en la ejecución de la medición de resultados, por causa de la variabilidad en las habilidades motoras.

3. VISIÓN Y AUDICIÓN

Las personas con SD muestran con frecuencia problemas de visión y audición (Esbensen y MacLean, 2017). En la niñez, muchos necesitan lentes correctoras para la visión próxima o lejana, para el astigmatismo, o si muestran estrabismo; tienen mayor tendencia a desarrollar cataratas, riesgo que aumenta con la edad. Los trastornos de visión influyen directamente sobre la medición de resultados, ya que muchas de las evaluaciones tienen un componente visual para iniciarlas y pueden tener implicaciones más amplias en el funcionamiento cognitivo

de los adultos (Krinsky-McHale et al., 2014). Las infecciones del oído medio aumentan el riesgo de pérdida de la audición de tipo conductivo en la niñez y eso va a influir sobre el retraso en el lenguaje en comparación con otros niños con SD que no tengan pérdida de audición (Laws y Hall, 2014). En la adultez, los trastornos de audición, especialmente de tipo neurosensorial para sonidos de alta frecuencia, comienzan antes que en la población general (Buchanan, 1990). Son pocos los adultos con pérdida auditiva que llevan audífonos, y la mayoría de los cuidadores no son conscientes de esta pérdida auditiva, por lo que los trastornos auditivos tienen consecuencias importantes para comprender la trayectoria en el desarrollo del envejecimiento de los adultos con SD a la hora de medir los resultados individuales (Van Buggenhout et al., 1999).

4. RECOMENDACIONES Y DIRECCIONES FUTURAS PARA MEDIR EL IMPACTO DE LA SALUD FÍSICA A LA HORA DE MEDIR LOS RESULTADOS

Pese a los avances realizados en la evaluación de los sistemas de medición de resultados en la cognición y comunicación, presentados en las anteriores secciones, existen grandes lagunas en nuestra actual comprensión de cómo las comorbilidades físicas o mentales impactan sobre la selección de un determinado sistema de medida. En la validación de futuras medidas, los investigadores habrán de evaluar si las comorbilidades, las capacidades motoras y los trastornos de audición y visión influyen sobre las propiedades psicométricas de posibles nuevos sistemas. Esto significa recoger datos normativos o de la historia natural con el fin de perfilar las diferentes trayectorias del desarrollo en los niños con o sin una determinada entidad médica. Mientras no se disponga de estos datos, se recomienda que los diseños de ensayo clínico aleatoricen a los participantes en función de la entidad clínica o al menos la tengan en cuenta en sus análisis estadísticos.

En la evaluación habrán de tenerse en cuenta los problemas habituales de audición y visión para determinar si durante la evaluación han de presentarse procedimientos que impliquen estímulos auditivos o visuales. Puede que se necesiten estímulos más prolongados o imágenes más grandes. Pueden ser necesarias exploraciones visuales y auditivas antes de participar en un estudio de investigación, y los trabajos se pueden beneficiar de los procedimientos que aseguren si se han de utilizar gafas o audífonos durante todo el estudio. Puede que se necesite una habitación silenciosa para hacer la evaluación o tener opción a que se aumente el volumen. La medición de los resultados

individuales que no tengan en cuenta la visión y la audición puede originar datos erróneos o inadecuados. Puede que también se necesiten adaptaciones a las habilidades motoras. En situaciones en que los instrumentos de manipulación sean pequeños, el disponer de opciones más grandes permitirá que las mediciones de resultados sean accesibles para niños con un mayor rango de habilidades motoras.

CONSIDERACIONES COMPLEMENTARIAS SOBRE LA EVALUACIÓN EN EL SÍNDROME DE DOWN

1. EL PERFIL FENOTÍPICO ASOCIADO AL SÍNDROME DE DOWN: IMPACTO SOBRE EL FUNCIONAMIENTO

Los instrumentos a utilizar en los estudios sobre el SD para medir adecuadamente el alcance de un tratamiento deben considerar de manera fiable un amplio abanico de situaciones (comorbilidades físicas y psiquiátricas) así como los niveles de desarrollo y de lenguaje. Esto es necesario si se quiere captar con validez el rango completo de niveles que existen en sus capacidades. Además, una determinada habilidad que suscite interés puede verse influenciada por el nivel de ejecución en otro dominio distinto del desarrollo o por una comorbilidad. Por ejemplo, la mayoría de las personas con SD tienen dificultades en su memoria operativa (de trabajo); y así, las exigencias de la memoria operativa junto con todo un espectro de tareas cognitivas, lingüísticas y conductuales pueden limitar su capacidad de actuación y contribuir a los efectos suelo. Del mismo modo, las limitaciones en la atención pueden impactar de forma significativa las actuaciones que son objeto de medición de los resultados cognitivos. Y así, más allá de mediciones específicas, las visitas de evaluación en un ensayo clínico han de considerar acortar los tiempos de cada evaluación para reducir la carga del estudio, suscitar un mayor interés por la prueba, y en último término facilitar una mejor disposición y aceptación por parte del participante en el estudio.

Además, puesto que las habilidades del lenguaje tienen un papel especial en el desarrollo social, el desarrollo cognitivo y los logros académicos, es esencial que el lenguaje se considere también cuando se planifique la selección de los métodos para medir los resultados en otros dominios. La mayoría de las mediciones de resultados, sea cual fuera el dominio de interés, descansa sobre la capacidad de la persona para

comprender las instrucciones verbales de la tarea y de su capacidad para formular una respuesta verbal (Edgin et al., 2019). Por ejemplo, si un individuo con SD falla en su respuesta a un ítem porque carece de las habilidades lingüísticas necesarias para responder, será difícil para el investigador desencajar los problemas del lenguaje de los problemas propios del dominio que está analizando, con el consiguiente descenso y desacierto en la puntuación que trata de analizar el verdadero nivel del individuo en ese dominio. Es decir, la comprensión de la tarea y la capacidad de responder impactan sobre la aceptación y el cumplimiento de la tarea, cualidades esenciales para obtener mediciones válidas y fiables de la conducta que se está analizando.

2. INTERACCIÓN ENTRE LA EJECUCIÓN Y EL CONTEXTO EN EL QUE ES EVALUADA

Al evaluar los resultados que se han de medir, existe otro factor a considerar: el contexto y el ambiente en que se realiza la evaluación. Puede tener lugar en casa, en la escuela, en un laboratorio de investigación, en la clínica o cualquier otro ambiente. Cada uno de estos contextos puede influir sobre los resultados que se miden. En algunos contextos, la medición puede hacerse de forma más natural y sus resultados son más generalizables. Por ejemplo, un niño puede sentirse más confortable en su casa o en su escuela, o ser evaluado por un profesor o instructor que le resulte familiar. Un niño que utilice un instrumento de CAA es más probable que lo use en la escuela donde se le ha enseñado a emplearlo que en un laboratorio de investigación; o incluso en casa, donde sus padres pueden comprender mejor su lenguaje u otras formas de comunicación. En otros contextos, el control del estudio y su estandarización pueden ser mejores, pero pueden también introducir evaluadores nuevos que pueden impedir la generalización de los hallazgos o la disposición del participante para terminar adecuadamente su tarea. Puede ser beneficioso realizar un primer ensayo (pre-testing) para reducir una posible variación en el trabajo achacable a que la tarea o el examinador son nuevos, y para mejorar la familiarización del participante con las exigencias de la tarea. Dependiendo de cuáles sean los resultados que se han de medir, el contexto puede también influir sobre la selección de las mediciones de los resultados. Por ejemplo, si se estudia el lenguaje expresivo, es importante el uso de las mediciones estandarizadas del lenguaje; pero también el lenguaje producido en ambientes más naturales, ya que éstos captan componentes diferentes sobre cómo el lenguaje cambia con el tiempo.

Otro aspecto importante de la interacción entre la ejecución y el contexto es la vigilancia de los niveles de atención del participante y de la atención que presta a la tarea. Si bien una prueba computarizada ofrece la oportunidad de conseguir la puntuación de forma estandarizada y automática, este tipo de pruebas puede verse perjudicado por el relativamente corto tiempo que dura su atención y la naturaleza social de las personas con SD si las tareas son menos motivadoras o demasiado largas. Si se utiliza un test computarizado, se ha comentado que la utilización de una tableta es más motivadora y más sencilla que el uso de pantalla/ratón (Hessl et al., 2016). Además, el grado de familiarización que cada individuo con SD tiene con los ordenadores y la tecnología puede influir sobre la prueba si se usan estas tecnologías.

3. PROPIEDADES PSICOMÉTRICAS DE LA EVALUACIÓN

3.1 Efectos suelo

Como ya se ha indicado, los rasgos del perfil fenotípico asociado al SD contribuyen a los efectos suelo. Cuando se eligen sistemas de medición que habían sido desarrollados para la población general, es alta la probabilidad de que aparezcan muchos efectos suelo al aplicarlos a un individuo con SD. Las puntuaciones en el suelo son especialmente importantes para comprender que además se asegure una adecuada interpretación. Los individuos pueden puntuar en el suelo de una medición debido a la dificultad de la tarea, a no mostrar docilidad con la tarea, a la dificultad por comprender las instrucciones de la tarea, o por causa de otras exigencias de la tarea. Es importante vigilar si el niño está considerando sus opciones de respuesta antes de responder para rastrear su compromiso con la actividad.

Ciertas tareas terminan con puntuaciones poco fiables debido a simple tendencia a la conformidad o a la casualidad. Por ejemplo, en las tareas de memoria que se basan en el reconocimiento o colocación de fichas sobre una cuadrícula pueden darse respuestas correctas por azar (Schworer, Voth et al., 2022). Por eso es importante vigilar si el individuo con SD ha creado un patrón, o está eligiendo el mismo rincón en la siguiente presentación, con el fin de entender mejor si la puntuación es verdadero reflejo de su actuación correcta. Entender cómo la misma puntuación se puede obtener desde distintos escenarios es importante a la hora de decidir cuáles son los datos que se deben incluir en una in-

vestigación o en una evaluación clínica, y de interpretar las mediciones obtenidas en los resultados y describir los hallazgos.

La preocupación por los efectos suelo pueden ser especialmente importantes en relación con otras habilidades de interés en los individuos con SD, como son el lenguaje y la memoria operativa. Los efectos suelo debilitan dramáticamente la evaluación de una intervención, ya que reducen la capacidad para demostrar cambios en la particular actividad que se está estudiando (Hessl et al., 2016). Por eso es necesaria la selección de datos brutos o formas alternativas de calcular las puntuaciones estándar (Hessl et al., 2016).

3.2 Validez

La trayectoria del fenotipo en el SD tiene implicaciones a la hora de elaborar mediciones precisas y válidas en específicos constructos de interés. Medir la cognición, por ejemplo, a menudo requiere respuestas verbales. Frecuentemente, en estudios sobre adultos con SD, se han utilizado tareas de aprendizaje de listas, en las que el estímulo auditivo debe ser recordado verbalmente (recuerdo verbal). Sin embargo, los frecuentes problemas que presentan en el lenguaje expresivo pueden comprometer la validez de este tipo de evaluación de la memoria. La confusión introducida por la respuesta verbal sirve como filtro a través del cual se ha de interpretar esa evaluación de la memoria, y la “impureza” de la tarea resta capacidad para interpretar cualquier problema que surja en la ejecución. Para abordar este tema de que la validez quede comprometida, se recomienda que se elijan mediciones con las mínimas exigencias verbales, que reduzcan la complejidad de respuestas motoras, y que mantenga la motivación en el estudio lo mejor posible. Se recomienda también que las exigencias de lenguaje receptivo sean igualmente reducidas, con instrucciones claras y ensayos previos para asegurar que el participante entiende la tarea antes de que le sea ya presentada para el análisis.

3.3 Fiabilidad y efectos debidos a la repetición de la práctica

Otras propiedades psicométricas importantes para establecer la utilidad de las evaluaciones en el SD son la fiabilidad test-retest y los efectos debidos a la práctica (efectos “práctica”). Se determina la información esencial al comprobar si las puntuaciones son constantes a lo largo de intervalos cortos entre los test (fiabilidad test-retest), y si las puntuaciones cambian cuando los test se repiten de forma múltiple

(efectos “práctica”). Son muchos los factores que influyen sobre la fiabilidad test-retest y los efectos “práctica”, incluidas la dificultad de la medición, la motivación, los efectos del examinador, la hora del día y el contexto en que se realiza la evaluación. La evidencia para establecer la fiabilidad de una prueba de medición utilizando datos normativos en individuos con desarrollo típico no garantiza la misma fiabilidad en las personas con SD. Tanto la fiabilidad test-retest como los efectos “práctica” influyen sobre cómo se interpretan las puntuaciones y, en último término, sobre los hallazgos que se describen tras intervenciones conductuales o farmacológicas. Hasta la fecha, se ha hecho alguna variación aceptable en la psicométrica de estos índices. Avanzando, será importante demostrar consistencia entre los estudios para que las evaluaciones de las mediciones sean válidas en lo referente a ambas propiedades. Para los estudios en que las visitas sean múltiples, es necesario determinar si se necesita que el examinador sea el mismo para mejorar la consistencia entre los intervalos. También es necesario evaluar otras mediciones en la batería, para determinar si estas mediciones pueden impactar efectos “práctica” o mejorar la ejecución. Por ejemplo, si “gato” es una palabra diana en una tarea de listado sobre la memoria, habrá de evitarse ese animal en las mediciones realizadas antes de la tarea de memoria, ya que su repetición podría aumentar la probabilidad de que esa palabra fuese recordada.

4. DICTAMEN DEL INFORMANTE

Dependiendo de la naturaleza del resultado que interese, las mediciones pueden ser solicitadas a los padres, maestros, hermanos u otros informantes. Cada uno de ellos rellena los impresos con puntuaciones basadas en su propia perspectiva y experiencia en relación con la persona con SD. Como resultado de haber observado a los individuos en contextos diferentes, hay abundante literatura que demuestra baja fiabilidad entre informantes en la población general, especialmente al puntuar conductas (Achenbach et al., 1987). Sin embargo, esta baja fiabilidad no niega el valor de recoger información a partir de informantes y contextos múltiples. Evaluaciones divergentes, como pueden ser las de un padre y un profesor, pueden arrojar luz sobre la presentación de diferentes conductas en contextos diferentes (Daunhauer et al., 2014), o pueden señalar a comparaciones con diferentes líneas basales para un determinado conjunto de evaluaciones frente a otro (Esbensen et al., 2021). Por tanto, el comprender la influencia de diferentes informantes o contextos sobre las mediciones de resultados tiene su propio valor.

5. ELEGIBILIDAD DE LOS PARTICIPANTES EN EL ESTUDIO

Los requisitos y los tipos de mediciones elegidas para realizar un estudio pueden tener consecuencias para quienes han sido elegidos e incluidos en el estudio. Excluir a participantes que son mínimamente verbales en un estudio en el que haya múltiples indicadores verbales podría estar justificado; y sin embargo deben considerarse los efectos de dicha exclusión. Si bien la utilización de tareas verbales está a veces justificada, es importante tener en cuenta hasta qué punto las mediciones verbales limitan el espectro de individuos elegibles para un estudio. En el ejemplo de los requisitos verbales, los investigadores deben tener la precaución de no crear sólo tratamientos para un subgrupo de individuos. En estudios anteriores, los criterios de inclusión habían tenido como requisito para participar un mínimo de puntuación del *Vineland Adaptive Behavior* de >55 (Kishnani et al., 2010). Aunque es razonable que un estudio exija un nivel mínimo de conducta adaptativa, es igualmente importante diseñar estudios con mediciones apropiadas de resultados que permitan criterios de inclusión más amplios.

RESUMEN DE DIRECTRICES FUTURAS PARA MEDIR RESULTADOS

Es cada vez más frecuente en diversos países que las personas con SD tengan actualmente acceso a programas de intervención temprana, prácticas de educación inclusiva, oportunidades de educación post-secundaria, y oportunidades laborales superiores a las de las generaciones anteriores (Roizen et al., 2013). Se están beneficiando de los avances en la medicina y la atención a la salud, que van asociados a incrementos sustanciales en su longevidad. En consecuencia, es importante medir de forma precisa los resultados conseguidos mediante estas oportunidades y terapias. Por otra parte, y puesto que nos movemos en un mundo en el que se están desarrollando nuevas intervenciones farmacológicas y clínicas, resulta imprescindible identificar mediciones muy sensibles de los resultados que sean capaces de detectar los cambios producidos por los tratamientos en dominios de la conducta y la vida de todas las personas con SD. La selección de los métodos que midan los resultados y cómo se aborden los actuales problemas de evaluación tienen implicaciones importantes a la hora de captar con fidelidad los efectos de la intervención. En último término, eso permitirá que un mayor número de personas con SD se beneficie de los programas clínicos y de los estudios de investigación.

BIBLIOGRAFÍA

- Abbeduto, L., McDuffie, A., Thurman, A., & Kover, S. (2016). En Vol. 50. Language development in individuals with intellectual and developmental disabilities: From phenotypes to treatments International Review of Research in Developmental Disabilities (pp. 71–118). Elsevier.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101(2), 213–232.
- Berry-Kravis, E., Hessel, D., Abbeduto, L., Reiss, A. L., Beckel-Mitchener, A., Urv, T. K., et al. (2013). Outcome measures for clinical trials in fragile X syndrome. *Journal of Developmental and Behavioral Pediatrics*, 34(7), 508.
- Bouck, E. C., Satsangi, R., Doughty, T. T., & Courtney, W. T. (2014). Virtual and concrete manipulatives: A comparison of approaches for solving mathematics problems for students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44(1), 180–193. <https://doi.org/10.1007/s10803-013-1863-2>.
- Breslin, J., Spano, G., Bootzin, R., Anand, P., Nadel, L., & Edgin, J. O. (2014). Obstructive sleep apnea syndrome and cognition in Down syndrome. *Developmental Medicine & Child Neurology*, 56(7), 657–664.
- Buchanan, L. H. (1990). Early onset of presbycusis in Down syndrome. *Scandinavian Audiology*, 19(2), 103–110.
- Buckley, S., & Johnson-Glenberg, M. (2008). Increasing literacy learning for individuals with Down syndrome and fragile X syndrome. En J. Roberts, R. Chapman, & S. Warren (Eds.), *Speech and language development and intervention in Down syndrome and fragile X syndrome*. Paul H. Brookes.
- Bull, M. J., & Genetics, C. O. (2011). Health supervision for children with Down syndrome. *Pediatrics*, 128(2), 393–406. <https://doi.org/10.1542/peds.2011-1605>.
- Capone, G., Goyal, P., Ares, W., & Lannigan, E. (2006). Neurobehavioral disorders in children, adolescents, and young adults with Down syndrome. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, 142(3), 158–172.
- Channell, M. M. (2020). The social responsiveness scale (SRS-2) in school-age children with Down syndrome at low risk for autism spectrum disorder. *Autism & Developmental Language Impairments*, 5, 2396941520962406.
- Channell, M. M., Loveall, S. J., Conners, F. A., Harvey, D. J., & Abbeduto, L. (2018). Narrative language sampling in typical development: Implications for clinical trials. *American Journal of Speech-Language Pathology*, 27(1), 123–135.
- Channell, M. M., Phillips, B. A., Loveall, S. J., Conners, F. A., Bussanich, P. M., & Klinger, L. G. (2015). Patterns of autism spectrum symptomatology in individuals with Down syndrome without comorbid autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, 7(1), 1–9.
- Chapman, R. (2006). Language learning in Down syndrome: The speech and language profile compared to adolescents with cognitive impairment of unknown origin. *Down Syndrome Research and Practice*, 10(2), 61–66.
- Chapman, R. S., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, 6(2), 84–95.
- Conners, F. A., Tungate, A. S., Abbeduto, L., Merrill, E. C., & Faught, G. G. (2018). Growth and decline in language and phonological memory over two years among adolescents with Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 123(2), 103–118.
- Coppus, A., Evenhuis, H., Verberne, G. J., Visser, F., Van Gool, P., Eikelenboom, P., et

- al. (2006). Dementia and mortality in persons with Down's syndrome. *Journal of Intellectual Disability Research*, 50(10), 768–777.
- Couzens, D., Cuskelly, M., & Haynes, M. (2011). Cognitive development and Down syndrome: Age-related change on the Stanford-Binet test. *American Journal on Intellectual and Developmental Disabilities*, 116(3), 181–204.
- d'Ardhuy, X. L., Edgin, J. O., Bouis, C., de Sola, S., Goeldner, C., Kishnani, P., et al. (2015). Assessment of cognitive scales to examine memory, executive function and language in individuals with Down syndrome: Implications of a 6-month observational study. *Frontiers in Behavioral Neuroscience*, 9, 300.
- Daunhauer, L. A., Fidler, D. J., Hahn, L., Will, E., Lee, N. R., & Hepburn, S. (2014). Profiles of everyday executive functioning in young children with Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 119(4), 303–318. <https://doi.org/10.1352/1944-7558-119.4.303>.
- Dekker, M. C., & Koot, H. M. (2003). DSM-IV disorders in children with borderline to moderate intellectual disability. I: Prevalence and impact. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(8), 915–922.
- del Hoyo Soriano, L., Thurman, A. J., & Abbeduto, L. (2018). Specificity: A phenotypic comparison of communication-relevant domains between youth with Down syndrome and fragile X syndrome. *Frontiers in Genetics*, 9, 424.
- del Hoyo Soriano, L., Thurman, A. J., Harvey, D., Kover, S. T., & Abbeduto, L. (2020). Expressive language development in adolescents with Down syndrome and fragile X syndrome: Change over time and the role of family-related factors. *Journal of Neurodevelopmental Disorders*, 12(1), 1–18.
- DiGuiseppi, C., Hepburn, S., Davis, J. M., Fidler, D. J., Hartway, S., Lee, N. R., et al. (2010). Screening for autism spectrum disorders in children with Down syndrome: Population prevalence and screening test characteristics. *Journal of Developmental and Behavioral Pediatrics*, 31(3), 181–191. <https://doi.org/10.1097/DBP.0b013e-3181d5aa6d>.
- Duff, K., Clark, H. J. D., O'Bryant, S. E., Mold, J. W., Schiffer, R. B., & Sutker, P. B. (2008). Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's disease: Sensitivity, specificity, and positive and negative predictive powers. *Archives of Clinical Neuropsychology*, 23(5), 603–612.
- Dykens, E. M., Hodapp, R. M., & Finucane, B. M. (2000). Genetics and mental retardation syndromes: A new look at behavior and interventions. Paul H Brookes Publishing.
- Edgin, J. O., Anand, P., Rosser, T., Pierpont, E. I., Figueroa, C., Hamilton, D., et al. (2017). The Arizona cognitive test battery for Down syndrome: Test-retest reliability and practice effects. *American Journal on Intellectual and Developmental Disabilities*, 122(3), 215–234.
- Edgin, J. O., Mason, G. M., Allman, M. J., Capone, G. T., DeLeon, I., Maslen, C., et al. (2010). Development and validation of the Arizona cognitive test battery for Down syndrome. *Journal of Neurodevelopmental Disorders*, 2, 149–164.
- Ekstein, S., Glick, B., Weill, M., Kay, B., & Berger, I. (2011). Down syndrome and attention-deficit/hyperactivity disorder. *Journal of Child Neurology*, 26, 1290–1295.
- Esbensen, A. J. (2010). Health conditions associated with aging and end of life of adults with Down syndrome. *International Review of Research in Mental Retardation*, 39, 107–126.
- Esbensen, A. J., Epstein, J. N., Vincent, L. B., Kamimura-Nishimura, K., Wiley, S., Angkustsiri, K., et al. (2021). Comparison of attention-deficit hyperactivity disorder in typically developing children and children with Down syndrome. *Journal of Developmental & Behavioral Pediatrics*.
- Esbensen, A. J., Hoffman, E. K., Shaffer, R., Chen, E., Patel, L., & Jacola, L. (2018). Reliability of parent report measures of behaviour in children with Down syndrome.

- Journal of Intellectual Disability Research, 62(9), 785–797. <https://doi.org/10.1111/jir.12533>.
- Esbensen, A. J., Hoffman, E. K., Shaffer, R., Chen, E., Patel, L., & Jacola, L. (2019). Reliability of informant-report measures of executive functioning in children with Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 124(3), 220–233. <https://doi.org/10.1352/1944-7558-124.3.220>.
- Esbensen, A. J., Hooper, S. R., Fidler, D., Hartley, S. L., Edgin, J. O., d’Ardhuy, X. L., et al. (2017). Outcome measures for clinical trials in Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 122(3), 247–281.
- Esbensen, A. J., & MacLean, W. (2017). Down syndrome. In M. L. Wehmeyer, I. Brown, M. Percy, K. A. Shogren, & W. L. A. Fung (Eds.), *A comprehensive guide to intellectual and developmental disabilities* (2nd Edition). Brookes: Paul H.
- Faroqi-Shah, Y., Treanor, A., Ratner, N. B., Ficek, B., Webster, K., & Tsapkini, K. (2020). Using narratives in differential diagnosis of neurodegenerative syndromes. *Journal of Communication Disorders*, 85, 105994.
- Feeley, K., & Jones, E. (2008). Strategies to address challenging behaviour in young children with Down syndrome. *Down Syndrome Research and Practice*, 12(2), 153–163.
- Fidler, D., Most, D., & Philofsky, A. (2008). The Down syndrome behavioural phenotype: Taking a developmental approach. *Down Syndrome Research and Practice*, 12(3), 37–44.
- Fidler, D., & Nadel, L. (2007). Education and children with Down syndrome: Neuroscience, development, and intervention. *Mental Retardation and Developmental Disabilities Research Reviews*, 13(3), 262–271.
- Fidler, D. J., Philofsky, A., Hepburn, S. L., & Rogers, S. J. (2005). Nonverbal requesting and problem-solving by toddlers with Down syndrome. *American Journal on Mental Retardation*, 110(4), 312–322.
- Finestack, L. H., & Abbeduto, L. (2010). Expressive language profiles of verbally expressive adolescents and young adults with Down syndrome or fragile X syndrome. *Journal of Speech, Language, and Hearing Research*, 53(5), 1334–1348.
- Finestack, L. H., Sterling, A. M., & Abbeduto, L. (2013). Discriminating Down syndrome and fragile X syndrome based on language ability. *Journal of Child Language*, 40(1), 244–265.
- Finlay, W. M., & Lyons, E. (2001). Methodological issues in interviewing and using self-report questionnaires with people with mental retardation. *Psychological Assessment*, 13(3), 319
- Fleming, V., Hom, C. L., Clare, I. C. H., Hurd-Thomas, S. L., Krinsky-McHale, S., Handen, B. L., et al. (2022). Cognitive outcome measures for tracking Alzheimer’s disease in down syndrome. *International Review of Research in Developmental Disabilities*.
- Frank, K., & Esbensen, A. (2015). Fine motor and self-care milestones for individuals with Down syndrome using a retrospective chart review. *Journal of Intellectual Disability Research*, 59, 719–729.
- Froehlich, T. E., Lanphear, B. P., Epstein, J. N., Barbaresi, W. J., Katusic, S. K., & Kahn, R. S. (2007). Prevalence, recognition, and treatment of attention-deficit/hyperactivity disorder in a national sample of US children. *Archives of Pediatrics & Adolescent Medicine*, 161(9), 857–864.
- Gardiner, K. J., Haurault, Y., Lott, I. T., Antonarakis, S. E., Reeves, R. H., & Dierssen, M. (2010). Down syndrome: From understanding the neurobiology to therapy. *The Journal of Neuroscience*, 30(45), 14943–14945.
- Gathercole, S. E., & Alloway, T. P. (2006). Practitioner review: Short-term and working memory impairments in neurodevelopmental disorders: Diagnosis and remedial support. *Journal of Child Psychology and Psychiatry*, 47(1), 4–15.

- Grieco, J., Pulsifer, M., Seligsohn, K., Skotko, B., & Schwartz, A. (2015). Down syndrome: Cognitive and behavioral functioning across the lifespan. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*.
- Hahn, L. J., Loveall, S. J., Savoy, M. T., Neumann, A. M., & Ikuta, T. (2018). Joint attention in Down syndrome: A meta-analysis. *Research in Developmental Disabilities*, 78, 89–102.
- Hawkins, B. A., Eklund, S. J., James, D. R., & Foose, A. K. (2003). Adaptive behavior and cognitive function of adults with Down syndrome: Modeling change with age. *Intellectual and Developmental Disabilities*, 41(1).
- Heller, J. H., Spiridigliozzi, G. A., Crissman, B. G., Sullivan, J. A., Eells, R. L., Li, J. S., et al. (2006). Safety and efficacy of rivastigmine in adolescents with Down syndrome: A preliminary 20-week, open-label study. *Journal of Child and Adolescent Psychopharmacology*, 16(6), 755–765.
- Heller, J. H., Spiridigliozzi, G. A., Sullivan, J. A., Doraiswamy, P. M., Krishnan, R. R., & Kishnani, P. S. (2003). Donepezil for the treatment of language deficits in adults with Down syndrome: A preliminary 24-week open trial. *American Journal of Medical Genetics Part A*, 116(2), 111–116.
- Hessl, D., Nguyen, D. V., Green, C., Chavez, A., Tassone, F., Hagerman, R. J., et al. (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: The case of fragile X syndrome. *Journal of Neurodevelopmental Disorders*, 1(1), 33.
- Hessl, D., Sansone, S. M., Berry-Kravis, E., Riley, K., Widaman, K. F., Abbeduto, L., et al. (2016). The NIH toolbox cognitive battery for intellectual disabilities: Three preliminary studies and future directions. *Journal of Neurodevelopmental Disorders*, 8(1), 35.
- Hodapp, R., & Dykens, E. (2004). Genetic and behavioural aspects: Application to maladaptive behavior and cognition. In J. Rondal, R. Hodapp, S. Soresi, E. Dykens, & L. Nota (Eds.), *Intellectual disabilities: Genetics, behaviour and inclusion* (pp. 13–48). Whurr.
- Hodapp, R., Leckman, J., Dykens, E., Sparrow, S., Zelinsky, D., & Ort, S. (1992). K-ABC profiles in children with fragile X syndrome, Down syndrome, and nonspecific mental retardation. *American Journal on Mental Retardation*, 97(1), 39–46.
- Holland, A., Hon, J., Huppert, F., & Stevens, F. (2000). Incidence and course of dementia in people with Down's syndrome: Findings from a population-based study. *Journal of Intellectual Disability Research*, 44(2), 138–146.
- Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., et al. (2013). Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, 12(2), 207–216. [https://doi.org/10.1016/S1474-4422\(12\)70291-0](https://doi.org/10.1016/S1474-4422(12)70291-0)
- Kanno, K., & Ikeda, Y. (2002). Word-length effect in verbal short-term memory in individuals with Down's syndrome. *Journal of Intellectual Disability Research*, 46(8), 613–618.
- Karmiloff-Smith, A., Al-Janabi, T., D'Souza, H., Groet, J., Massand, E., Mok, K., et al. (2016). The importance of understanding individual differences in Down syndrome. *F1000Research*, 5.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test—second edition (KBIT-2)*. Circle Pines, MN: American Guidance Service.
- Kishnani, P. S., Heller, J. H., Spiridigliozzi, G. A., Lott, I., Escobar, L., Richardson, S., et al. (2010). Donepezil for treatment of cognitive dysfunction in children with Down syndrome aged 10–17. *American Journal of Medical Genetics Part A*, 152(12), 3028–3035.
- Klein, B. P., & Mervis, C. B. (1999). Contrasting patterns of cognitive abilities of 9- and 10-year-olds with Williams syndrome or Down syndrome. *Developmental Neuropsychology*.

- chology, 16(2), 177–196.
- Krinsky-McHale, S. J., Silverman, W., Gordon, J., Devenny, D. A., Oley, N., & Abramov, I. (2014). Vision deficits in adults with Down syndrome. *Journal of Applied Research in Intellectual Disabilities*, 27(3), 247–263. <https://doi.org/10.1111/jar.12062>.
- Laws, G., & Hall, A. (2014). Early hearing loss and language abilities in children with Down syndrome. *International Journal of Language & Communication Disorders*, 49(3), 333–342. <https://doi.org/10.1111/1460-6984.12077>.
- Magyar, C. I., Pandolfi, V., & Dill, C. A. (2012). An initial evaluation of the social communication questionnaire for the assessment of autism spectrum disorders in children with Down syndrome. *Journal of Developmental and Behavioral Pediatrics*, 33(2), 134–145. <https://doi.org/10.1097/DBP.0b013e318240d3d9>.
- Mantry, D., Cooper, S. A., Smiley, E., Morrison, J., Allan, L., Williamson, A., et al. (2008). The prevalence and incidence of mental ill-health in adults with Down syndrome. *Journal of Intellectual Disability Research*, 52(2), 141–155.
- Martin, G. E., Klusek, J., Estigarribia, B., & Roberts, J. E. (2009). Language characteristics of individuals with Down syndrome. *Topics in Language Disorders*, 29(2), 112–132.
- McCabe, L. L., Hickey, F., & McCabe, E. R. (2011). Down syndrome: Addressing the gaps. *Journal of Pediatrics*, 159(4), 525–526.
- McCarron, M., Gill, M., McCallion, P., & Begley, C. (2005). Health co-morbidities in ageing persons with Down syndrome and Alzheimer's dementia. *Journal of Intellectual Disability Research*, 49(7), 560–566.
- McDuffie, A., & Abbeduto, L. (2009). Language disorders in children with mental retardation of genetic origin: Down syndrome, fragile X syndrome, and Williams syndrome. In *Handbook of child language disorders* (pp. 44–66).
- McGuire, D., & Chicoine, B. (2006). *Mental wellness in adults with Down syndrome*. Woodbine House.
- Mircher, C., Sacco, S., Bouis, C., Gallard, J., Pichot, A., Le Galloudec, E., et al. (2020). Thyroid hormone and folic acid in young children with Down syndrome: The phase 3 ACTHYF trial. *Genetics in Medicine*, 22(1), 44–52.
- Moyer, A. (2021). Too good to be true: Reflections on a Down syndrome clinical trial. Johns Hopkins Medicine. Retrieved 8/5 from <https://biomedicalodyssey.blogs.hopkinsmedicine.org/2021/03/too-good-to-be-true-reflections-on-a-down-syndrome-clinical-trial/>.
- Patel, L., Wolter-Warmerdam, K., Leifer, N., & Hickey, F. (2018). Behavioral characteristics of individuals with Down syndrome. *Journal of Mental Health Research in Intellectual Disabilities*, 11(3), 221–246.
- Pezzuti, L., Nacinovich, R., Oggiano, S., Bomba, M., Ferri, R., La Stella, A., et al. (2018). Beyond the floor effect on the WISC-IV in individuals with Down syndrome: Are there cognitive strengths and weaknesses? *Journal of Intellectual Disability Research*, 62(7), 593–603.
- Phillips, B. A., Loveall, S. J., Channell, M. M., & Conners, F. A. (2014). Matching variables for research involving youth with Down syndrome: Leiter-R versus PPVT-4. *Research in Developmental Disabilities*, 35(2), 429–438.
- Pulina, F., Vianello, R., & Lanfranchi, S. (2019). Cognitive profiles in individuals with Down syndrome. In *International Review of Research in Developmental Disabilities*: Vol. 56 (pp. 67–92). Elsevier.
- Roid, G. H. (2003). *Stanford-Binet intelligence Scales (SB5)*. Riverside.
- Roizen, N. (2013). Down syndrome. In M. Batshaw, N. Roizen, & G. R. Lotrecchiano (Eds.), *Children with disabilities* (7th ed., pp. 307–318). Paul H Brookes.
- Sansone, S. M., Schneider, A., Bickel, E., Berry-Kravis, E., Prescott, C., & Hessel, D. (2014). Improving IQ measurement in intellectual disabilities using true deviation from population norms. *Journal of Neurodevelopmental Disorders*, 6(1), 1–15.

- Schworer, E. K., Esbensen, A. J., Fidler, D. J., Beebe, D. W., Carle, A., & Wiley, S. (2021). Evaluating working memory outcome measures for children with Down syndrome. *Journal of Intellectual Disability Research*. <https://doi.org/10.1111/jir.12833>.
- Schworer, E. K., Fidler, D. J., & Daunhauer, L. A. (2021). Early regulatory skills and social communication development in infants with Down syndrome. *Brain Sciences*, 11(2), 208.
- Schworer, E. K., Fidler, D., Kaur, M., Needham, A., Prince, M., & Daunhauer, L. (2020). Goal-directed action planning in infants with Down syndrome. *Journal of Intellectual Disability Research*, 64(9), 713–724.
- Schworer, E. K., Fidler, D. J., Kaur, M., Needham, A., Prince, M. A., & Daunhauer, L. A. (2022). Infant precursors of executive function in Down syndrome. *Journal of Intellectual Disability Research*, 66, 108–120. <https://doi.org/10.1111/jir.12824>.
- Schworer, E. K., Hoffman, E. K., & Esbensen, A. J. (2021). Psychometric evaluation of social cognition and behavior measures in children and adolescents with Down syndrome. *Brain Sciences*, 11(7). <https://doi.org/10.3390/brainsci11070836>.
- Schworer, E. K., Voth, K., Hoffman, E., & Esbensen, A. (2022). Short-term memory outcome measures: Psychometric evaluation and performance in youth with Down syndrome. *Research in Developmental Disabilities*, 120, 104147.
- Shields, R. H., Kaat, A. J., McKenzie, F. J., Drayton, A., Sansone, S. M., Coleman, J., et al. (2020). Validation of the NIH toolbox cognitive battery in intellectual disability. *Neurology*, 94(12), e1229–e1240.
- Shott, S. (2006). Down syndrome: Common otolaryngologic manifestations. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, 142C, 131–140.
- Smeyne, C., Esbensen, A. J., Schworer, E., Belzaira, S., Hoffman, E., Beebe, D., et al. (2022). Evaluating verbal fluency outcome measures for children with Down syndrome. *American Journal on Intellectual and Developmental Disabilities*.
- Stein, D. (2016). Supporting positive behavior in children and teens with Down syndrome. Woodbine House.
- Strydom, A., Shoostari, S., Lee, L., Raykar, V., Torr, J., Tsiouris, J., et al. (2010). Dementia in older adults with intellectual disabilities—Epidemiology, presentation, and diagnosis. *Journal of Policy and Practice in Intellectual Disabilities*, 7(2), 96–110.
- Thurman, A., & del Hoyo Soriano, L. (2021). Down syndrome. In L. Cummings (Ed.), *Pragmatic language disorders: Complex and underserved populations* (pp. 99–128). Springer Nature.
- Thurman, A. J., Edgin, J. O., Sherman, S. L., Sterling, A., McDuffie, A., Berry-Kravis, E., et al. (2021). Spoken language outcome measures for treatment studies in Down syndrome: Feasibility, practice effects, test-retest reliability, and construct validity of variables generated from expressive language sampling. *Journal of Neurodevelopmental Disorders*, 13(1), 1–17.
- Van Buggenhout, G., Trommelen, J., Schoenmaker, A., De Bal, C., Verbeek, J., Smeets, D., et al. (1999). Down syndrome in a population of elderly mentally retarded patients. Genetic-diagnostic survey and implications for medical care. *American Journal of Medical Genetics*, 85(4), 376–384.
- Visootsak, J., Mahle, W. T., Kirshbom, P. M., Huddleston, L., Caron-Besch, M., Ransom, A., et al. (2011). Neurodevelopmental outcomes in children with Down syndrome and congenital heart defects. *American Journal of Medical Genetics Part A*, 155(11), 2688–2691.
- Wiig, E., Secord, W. A., & Semel, E. (2004). Clinical evaluation of language fundamentals—Preschool, (CELF Preschool-2). San Antonio, TX: Harcourt Assessment.
- Winders, P. (2013). Gross motor skills for children with Down syndrome. In Bethesda, MD: Woodbine House.
- Wolpert, G. (2001). What general educators have to say about successfully including students with Down syndrome in their classes. *Journal of Research in Childhood*

- Education, 16(1), 28–38.
- Zackheim, C. T., & Conture, E. G. (2003). Childhood stuttering and speech disfluencies in relation to children's mean length of utterance: A preliminary study. *Journal of Fluency Disorders*, 28(2), 115–142.
- Zigman, W. B., Schupf, N., Sersen, E., & Silverman, W. (1996). Prevalence of dementia in adults with and without Down syndrome. *American Journal on Mental Retardation*, 100(4), 403-12.
- Zoia, S., Pelamatti, G., & Rumiati, R. I. (2004). Praxic skills in down and mentally retarded adults: Evidence for multiple action routes. *Brain and Cognition*, 54(1), 7–17

COMENTARIO FINAL (Down 21)

Quizá el hallazgo más relevante en los recientes análisis de las características que configuran al individuo con síndrome de Down sea la confirmación de su individual diversidad. Los datos estaban ahí, pero había primado la necesidad de configurar y definir la singularidad de un síndrome dentro del amplio mundo de la discapacidad intelectual. Razones no faltaban. Era ya evidente la precisión de su etiología: la presencia de un tercer cromosoma 21, total o parcial, en los núcleos de las células; pero había que delimitar las múltiples características en las áreas de las comorbilidades de la salud que podían determinar su vida, de las diversas habilidades cognitivas, de la capacidad adaptativa, de la conducta, de las posibilidades de desarrollo a lo largo de la vida y su respuesta a las múltiples influencias que cada persona había de recibir. Simultáneamente, se fue configurando el estudio de las raíces biológicas que debían explicar todo ese mundo tan complejo, mediante la utilización de la moderna tecnología en los campos de la genética, la biología molecular y celular, la neurociencia en sus múltiples dimensiones incluidas la neuropsicología y la salud mental.

Dónde nos encontramos en nuestros actuales conocimientos y hacia dónde debemos continuar para progresar ha sido el objetivo del estudio mostrado en el primer número de Cuadernos de Actualidad. El estudio ofrecido en el presente Cuaderno plantea el cómo hemos de hacerlo, al descubrir y desvelar el alto grado de diversidad individual que ofrecen las personas con síndrome de Down. Y lo hace analizando tres grandes dimensiones: la cognición, el lenguaje y la conducta/ psicopatología; añadiendo, además, el impacto que tienen la salud física y la psicométrica de la evaluación sobre las mediciones de los resultados.

Destacan dos aspectos en este estudio: la necesidad de seleccionar bien la muestra de los individuos a analizar, y la necesidad de elegir bien qué resultados elegimos para alcanzar la debida información y qué método hemos de utilizar para aplicar la medición correcta y su análisis.

La selección de la muestra va íntimamente ligada a la realidad que hemos aprendido: la variabilidad individual. Es preciso distinguir entre la investigación realizada en los primeros meses de la vida, cuando el cerebro es —en buena parte— una hoja en blanco, y la investigación realizada posteriormente cuando esa hoja se va llenando de múltiples trazos que irán perfilando el poderío y las características de la mente en todas sus dimensiones. La situación del cerebro en

los primeros meses depende muy prioritariamente de la dotación genética del individuo, en la que ciertamente influye la herencia parental; y por ello, la información de los estudios realizados en ese periodo informan sobre el estado “virgen” de la realidad cerebral propia de los bebés con síndrome de Down. Es decir, detectamos en su estado más puro el perfil de sus habilidades y limitaciones. Son muy notables la investigación y sus resultados conseguidos en esta fase, que definieron en buena parte nuestro conocimiento sobre características esenciales propias del síndrome de Down y propiciaron el desarrollo de los programas de atención temprana. Ya en esa precoz etapa se detectaba la variabilidad individual, algo que ha tardado mucho tiempo en reconocerse y sacar las debidas conclusiones. Pero, además, esos necesarios estudios no eran capaces de revelar las capacidades o posibilidades de desarrollo, es decir, de respuesta a la infinita diversidad de influencias que habrían de penetrar en cada cerebro. La intensidad y calidad de la respuesta del individuo dependen: 1) de su capacidad genética para movilizar un cambio neuroplástico en un cerebro que va madurando en el tiempo –factor intrínsecamente individual–, y 2) de la calidad y características de las influencias externas, es decir, de las intervenciones que provienen de múltiples fuentes o entornos: la capacidad de la familia (conocimientos, motivaciones, habilidades, fortaleza, constancia, implicación global de sus miembros); la capacidad de los profesionales en sus múltiples variantes; la capacidad del medio social; las comorbilidades que condicionan el aprovechamiento de esas intervenciones.

Así, pues, debemos preguntarnos en qué medida han sido tenidos en cuenta todos estos factores a la hora de seleccionar las muestras de los individuos con síndrome de Down que van a participar en un determinado estudio en el que, por ejemplo, se trata de analizar la respuesta a una concreta intervención, médica, conductual o farmacológica. Muchos de estos estudios no los han tenido en cuenta o han utilizado parámetros muy limitados. Por ejemplo, a la hora de definir el factor familia, sólo se suele contemplar el nivel social y el nivel educativo, como si eso ilustrara fehacientemente sobre su motivación o su constancia, factores determinantes en la educación. Son precisamente las debilidades propias de la trisomía las que hacen más imprescindible conocer la calidad de la atención anterior y permanentemente prestada: la de quien la aplica y la de los métodos que utiliza.

Quizá esto explique la perplejidad con que a veces uno contempla los resultados de estudios publicados, que no encajan con la realidad de los que uno observa y comprueba en su vida diaria en las personas con síndrome de Down que le rodean y a las que atiende. Cuanto más heterogénea sea la muestra –por querer abarcar a más personas, y ya

no digamos si es un estudio internacional— más diferencias existirán en los antecedentes de los participantes, en su historia, y más difícil será conocer el estado real de su estado cognitivo y de su estado afectivo-conductual que necesariamente condicionan la respuesta a la intervención. Puede que eso explique, al menos en parte, el fracaso de tantos ensayos clínicos para medir la eficacia de un posible producto, a la hora de trasladar los éxitos obtenidos en los estudios en modelos animales a las pruebas realizadas en las personas con síndrome de Down.

Junto a lo dicho, no menos importante es conocer por qué determinados individuos con síndrome de Down son excluidos de los estudios, porque ello evitará articular generalizaciones y obligará a matizar conclusiones excesivamente generosas.

El segundo principio destacado en esta revisión es la necesidad de elegir bien qué resultados elegimos para alcanzar la debida información y qué método hemos de utilizar para aplicar la medición correcta y su análisis. Las autoras se extienden en explicar las cualidades de las pruebas que ellas eligen a la hora de evaluar debidamente la cognición, el lenguaje y la conducta/psicopatología. Analizan además la influencia de la salud física en los resultados, hecho demasiadas veces olvidado en la mayoría de los estudios, y la psicométrica en la evaluación de los resultados. Cabe aprender de su experiencia a la hora de diseñar los estudios.

Publicaciones de la Fundación Iberoamericana Down21

Colección DOCUMENTOS

- 100 Editoriales en Canal Down21
- Programación educativa para escolares con síndrome de Down. *E. Ruíz*
- Los bebés con síndrome de Down tienen derecho a vivir. *J.-A. Rondal*
- Respirar por la boca: Consecuencias, prevención y tratamiento. *S. Galletti*
- Síndrome de Down. Comunicar la noticia: Primer acto terapéutico. 3ª ed. *J. Flórez*
- Progresando juntos. Nuevos editoriales de www.down21.org (2010-2017). 2ª ed.
- ¿Cómo favorecer la autonomía personal de mi hijo con síndrome de Down? *B. Garvía*
- Regresión de origen desconocido en el síndrome de Down. Una nueva entidad. *J. Flórez, M.C. Ortega*
- Terapia farmacológica: Neurodesarrollo y neurodegeneración en el síndrome de Down. *Fundación Iberoamericana Down21*

Colección RECURSOS

- Síndrome de Down: lectura y escritura (sólo en versión digital). *M.V. Troncoso, M. del Cerro*
- Bienestar mental en los adultos con síndrome de Down: Una guía para comprender y evaluar sus cualidades y problemas emocionales y conductuales. *D. McGuire, B. Chicoine*
- Síndrome de Down: habilidades tempranas de comunicación. Una guía para padres y profesionales. *L. Kumin*
- Síndrome de Down: Neurobiología, neuropsicología, salud mental. Bases para la intervención en el aprendizaje, la conducta y el bienestar mental. *J. Flórez, B. Garvía, R. Fernández-Olaria*
- Educación postsecundaria en entorno universitario para alumnos con discapacidad intelectual. Experiencias y resultados. *D. Cabezas, J. Flórez (Coord.)*
- Síndrome de Down y Autismo: Cuando dos mundos se encuentran. Una guía para padres y profesionales. *M. Froehlke, R. Zaborek (Coord.)*

- Psicoterapia y Farmacoterapia en la Discapacidad Intelectual. *B. Garvía, J. Flórez*
- Enfermedad de Alzheimer & Síndrome de Down: Guía práctica para cuidadores. *J. Moran y National Down Syndrome Society*
- Síndrome de Down. Mi visión y presencia. *M.V. Troncoso.*
- La esperanza pedagógica vivida por padres de hijos con síndrome de Down. *K. Armijos Yambay*
- La vida adulta en el síndrome de Down. *J. Flórez (Coord.)*
- Terapia farmacológica: Neurodesarrollo y neurodegeneración en el síndrome de Down. *Fundación Iberoamericana Down21*

Cuadernos de Actualidad

- La investigación en el síndrome de Down: propuestas actuales y recomendaciones. *J.A. Hendrix et al.*
- La medición de los resultados en los estudios de investigación sobre las personas con síndrome de Down. *Anna J Esbensen et al.*

REVISTAS

- Revista Virtual Canal Down21 , mensual
- Síndrome de Down: Vida Adulta, cuatrimestral

PÁGINAS WEB

- www.down21.org
- www.downciclopedia.org
- www.sindromedownvidaadulta.org

El estudio aquí presentado es una revisión sistemática que se centra en los problemas que plantea la medición de los resultados en el síndrome de Down, con un enfoque específico sobre las consideraciones que han de hacerse para evaluar los resultados relativos a dimensiones sustanciales de la persona, como son la cognición, el lenguaje y la conducta/psicopatología. Esta revisión también proporciona un resumen de las recomendaciones sobre la valoración de los resultados en estos dominios, así como recomendaciones para la futura investigación.

Al mismo tiempo, analiza el impacto que tienen la salud física y la psicométrica de la evaluación sobre las mediciones de los resultados.

Las autoras son investigadoras muy acreditadas que han aportado un gran beneficio en el conocimiento sobre aspectos esenciales de la personalidad y desarrollo de los individuos con síndrome de Down, en diversas etapas de su vida.

El presente estudio complementa el presentado en nuestro Cuaderno de Actualidad nº 1. A la vista de la necesidad de seguir profundizando en las investigaciones, se hace preciso clarificar con objetividad la sistemática a seguir en cada proyecto, con el fin de que los resultados sean bien recogidos y las conclusiones resulten fiables.

