

Hemos leído

Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal¹

Learning what works in sensory disabilities: establishing causal inference

J. B. Cooney, J. Young, J. L. Luckner, K. A. Ferrell²

Resumen

El propósito del presente artículo es ayudar a profesores e investigadores a diseñar estudios que examinen la eficacia de una determinada intervención o estrategia con alumnos con distintos tipos de discapacidad sensorial. Se analizan diez diferentes diseños de investigación que pueden determinar la inferencia causal (capacidad para atribuir cualquier efecto a la intervención) con y sin aleatorización.

Palabras clave

Investigación. Diseño de investigación. Estudio de intervención. Inferencia causal.

1 Artículo publicado con el título *Learning what works in sensory disabilities: establishing causal inference* en la revista *Journal of Visual Impairment and Blindness*, vol. 109, n.º 5 (noviembre-diciembre 2015), p. 469-486. Copyright © 2015 by American Foundation for the Blind (AFB). Todos los derechos reservados. Traducción de María Dolores Cebrián-deMiguel, publicada con permiso de la AFB.

2 **John B. Cooney**, Ph. D., vicepresidente asociado de Administración. Office of Academic Affairs, University of Colorado, Campus Box 35 UCA, 1800 Grant Street, Suite 800, Denver, CO 80203 (Estados Unidos). Correo electrónico: <john.cooney@cu.edu>. **John Young III**, Ph. D., vicepresidente de Ciencias del Desarrollo y Estadística. Applied Analytics Group, DST Systems, 333 West 11th Street, Kansas City, MO 64105 (Estados Unidos). Correo electrónico: <outcome32@yahoo.com>. **John L. Luckner**, Ed. D., catedrático de Educación Especial y coordinador del Programa de Educación para Sordos. College of Education and Behavioral Sciences, McKee 29, Campus Box 141, University of Northern Colorado, Greeley, CO 80639 (Estados Unidos). Correo electrónico: <john.luckner@unco.edu>. **Kay Alicyn Ferrell**, Ph. D., catedrática emérita de Educación Especial. School of Special Education, College of Education and Behavioral Sciences, University of Northern Colorado (Estados Unidos). Correo electrónico: <kay.ferrell@bears.unco.edu>. Se ruega dirigir toda la correspondencia al Dr. Cooney.

Abstract

This article is intended to assist teachers and researchers in designing studies that examine the efficacy of a particular intervention or strategy with students with sensory disabilities. Ten research designs that can establish causal inference (the ability to attribute any effects to the intervention) with and without randomization are discussed.

Key words

Research. Research design. Intervention study. Causal inference.

Desde que se publicó la ley *No Child Left Behind Act* (2001) y la subsiguiente reautorización de la ley de educación *Individuals with Disabilities Education Act* (2004), el sector de la educación especial ha puesto un mayor énfasis en las investigaciones con base científica y en las prácticas basadas en evidencias. La revista *Exceptional Children*, revista puntera en educación especial, ha dedicado un número monográfico a especificar los indicadores de calidad para distintos diseños de investigación (Brantlinger, [Jimenez, Klingner, Pugach y Richardson,] 2005; Gersten, [Fuchs, Compton, Coyne, Greenwood e Innocenti,] 2005; Horner, [Carr, Halle, McGee y Odom,] 2005; Odom, [Brantlinger, Gersten, Horner, Thompson y Harris,] 2005; Thompson, [Diamond, K. E., McWilliam, R., Snyder, P., y Snyder, S. W.,] 2005), a título de guía para evaluar los estudios de investigación y mejorar la base científica en educación especial.

En materia de discapacidad sensorial, es, a menudo, difícil alcanzar estos indicadores de calidad. Puesto que las distintas discapacidades sensoriales forman parte de las llamadas discapacidades de baja prevalencia (entre 2012 y 2013 los alumnos con discapacidad visual representaban el 0,1 % de la población escolar de entre 3 y 21 años, mientras que los niños con sordera o deficiencia auditiva representaban el 0,2 %, según Snyder y Dillow, 2015), investigar en este terreno es mucho más difícil, y no solo por el reducido número de afectados, sino por el hecho de que los sujetos están dispersos por una zona geográfica extremadamente amplia. Muchos distritos escolares de zonas rurales no matriculan a los alumnos con discapacidad sensorial, mientras que otros pueden integrar a un solo alumno, que es el único con ese tipo de deficiencia sensorial. Llegar hasta estos alumnos con intención de realizar una investigación representa un enorme gasto. Además, los niños con discapacidad sensorial están considerados como un grupo de población muy heterogéneo, lo que significa que dos niños de una misma edad, diagnóstico y nivel de estudios pueden tener dife-

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

rentes grados de deficiencia visual, o de pérdida auditiva, que afectan a su capacidad para responder a cualquier intervención concreta. La naturaleza de baja prevalencia de las distintas discapacidades sensoriales supone que los estudios de replicación no abundan, puesto que no se puede acudir a los mismos niños para replicar los estudios aunque lo hagan distintos investigadores, y que el transcurso del tiempo conduce a confundir las variables de exposición y madurez anteriores.

Una serie de artículos han examinado el nivel de evidencia de varios estudios de investigación sobre deficiencias sensoriales, a fin de identificar las prácticas más deseables basadas en la evidencia. Entre estos artículos se incluyen los de Abou-Gareeb, Lewallen, Bassett y Courtright, 2001; Botsford (2013); Cawthon y Leppo (2013); Ferrell, Buettel, Sebald y Pearson (2006); Ferrell, Dozier y Monson (2011); Ferrell, Mason, Young y Cooney (2006); Graeme et al. (2011); Kelly y Smith (2011); Luckner (2006); Luckner y Cooke (2010); Luckner y Handley (2008); Luckner, Sebald, Cooney, Young y Muir (2005, 2006); Luckner y Urbach (2012); Parker, Davidson y Banda (2007); Parker, Grimmett y Summers (2008); Parker y Ivy (2014); Parker y Pogrund (2009); Wang y Williams (2014), y Wright, Harris y Sticken (2010). Todos estos estudios han llegado a la conclusión de que la pedagogía para la discapacidad sensorial se caracteriza por la escasez de evidencia científica. Un reciente análisis efectuado para el proyecto *Collaboration for Effective Educator Development, Accountability, and Reform* (CEEDAR) de la Universidad de Florida, que atiende a estados e instituciones de educación superior para que reformen sus programas de capacitación del profesorado, reveló que la mayor parte de la investigación realizada sobre la discapacidad visual y la educación de la sordera y sordoceguera presentaba bajos niveles de evidencia (Ferrell, Bruce y Luckner, 2014). Se localizaron varias áreas en las que la evidencia apenas empezaba a aflorar, como es el caso de la serie de publicaciones que cuestiona la práctica de los centros escolares locales de contratar a educadores auxiliares (frecuentemente referidos como «paraprofesionales») para prestar servicios a alumnos con discapacidad visual (Conroy, 2007; Forster y Holbrook, 2005; Griffin-Shirley y Matlock, 2004; Harris, 2011; Koenig y Holbrook, 2000; Lewis y McKenzie, 2010; Marks, Schrader y Levine, 1999; McKenzie y Lewis, 2008; Russotti y Shaw, 2001), ya que tal evidencia estaba basada en limitados estudios de intervención y en opiniones. También había áreas consideradas como consolidadas, como la de los servicios de identificación e intervención tempranas de niños con sordera o deficiencia auditiva de edad inferior a los seis meses (Joint Committee on Infant Hearing, 2007; Meinzen-Derr, Wiley y Choo, 2011; Moeller, 2000; Vlastarakos, Proikas, Papacharalampous, Exadaktylou, Mochloulis y Nikolopoulos, 2010; Vohr et al., 2008; Yoshinaga-Itano, 2003a, 2003b;

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

Yoshinaga-Itano y Gravel, 2001; Yoshinaga-Itano, Sedey, Coulter y Mehl, 1998), sustentadas por un cierto número de estudios de intervención replicados por otros investigadores. Aunque la base de la investigación parece ir en aumento, Ferrel et al. (2014) concluyeron que la investigación en discapacidad sensorial se caracterizaba, por lo general, por estudios sin inferencias causales que incluían escasa evidencia de las estrategias que constituían la práctica.

Además de escasear las afirmaciones basadas en la evidencia sobre la eficacia de las intervenciones con personas con discapacidad sensorial, la investigación sobre las intervenciones derivadas de los modelos teóricos también es inexistente. Por otra parte, hay pocos estudios de replicación (si es que existe alguno) e investigaciones sistemáticas de cualquier tipo concreto sobre intervenciones diseñadas para mejorar la educación de los estudiantes con discapacidad sensorial. Esta falta de tales replications o investigaciones no significa que no exista un sólido conocimiento sobre este tipo de personas. Sin embargo, lo que sí es cierto es que hay escasa evidencia científica sobre la eficacia de las intervenciones que promueven el aprendizaje y el desarrollo de los niños con discapacidad sensorial.

Este artículo describe las estrategias para diseñar investigaciones que fortalezcan las afirmaciones causales sobre intervenciones previstas para personas con distintas deficiencias sensoriales. En primer lugar, se describen aquí las tácticas para mejorar lo que muchos consideran un potente diseño para diseñar inferencias causales sobre las intervenciones, es decir, un diseño que utilice pretest, postest y asignación aleatoria de participantes en el estudio para grupos experimentales y grupos de control. Debido a que este diseño concreto, pese a ser considerado el «estándar de oro» en la investigación sobre educación (Sullivan, 2011), no siempre es posible con las personas con discapacidad sensorial e, incluso, no especialmente relevante (Cronbach, 1982), a continuación se describen ciertas mejoras que pueden incorporarse a los diseños casi experimentales —en los que los participantes no son asignados aleatoriamente a los grupos experimentales y de control— y que mejoran las inferencias causales. El diseño de regresión discontinua [RDD, por sus siglas en inglés], por ejemplo, tiene mucho que ofrecer a los investigadores que trabajan con sujetos con discapacidad sensorial. No obstante, los análisis de Ferrell (2006) y Luckner (2006) no identificaron ni un solo estudio que hubiera empleado este diseño. Así pues, nuestro tercer objetivo es introducir el diseño de regresión discontinua como técnica útil para la construcción de aseveraciones causales sobre estrategias de intervención. En cuarto lugar, se exploró el uso de dos diseños en los cuales los participantes actúan como propio grupo de

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

control: el diseño unifactorial intrasujeto y el diseño de series temporales. Aunque los diseños de caso único no figuraban entre los diseños que Valentine y Cooper (2004) identificaron por cumplir los estándares de evidencia para efectuar demandas causales en la época en que Ferrell y Luckner hicieron sus revisiones, los actuales estándares de evidencia (Valentine y Cooper, 2008; What Works Clearinghouse, 2014) ofrecen directrices para construir afirmaciones causales procedentes de diseños de casos únicos. No obstante, existen otros diseños de investigación que pueden utilizarse para demostrar la eficacia de una intervención.

Las preguntas en la metodología de la investigación

Se ha puesto excesivamente de moda enmarcar los debates educativos en torno a la metodología de la investigación. Sin embargo, en nuestra opinión y en la de Bartlett[, MacNab, MacArthur, Mandwich, Magill-Evans, Young, Beal, Conti-Becker y Polatajko] (2006), es más productivo pensar en seleccionar una metodología de la investigación en relación con las preguntas de esta y su interrelación. Por ejemplo, Shavelson, Towne, y el Committee on Scientific Principles for Education Research (2002) han sugerido que la investigación en Educación, Ciencias Sociales y Ciencias Naturales adopta una de estas tres formas: (1) comprensión de lo que está ocurriendo, (2) detección de la relación causal, y (3) explicación de una relación causal. Visto de esta forma, tiene poco sentido discutir si se debe utilizar metodología cualitativa o cuantitativa hasta que uno haya estudiado en detalle las preguntas de las que se quiere obtener respuesta.

Comprender lo que está ocurriendo

Para comprender lo que sucede, los investigadores tienen que emplear herramientas que produzcan abundantes descripciones de las personas, sus entornos y sus interacciones, así como de las personas que aportan precisas estimaciones cuantitativas sobre importantes características. Por una parte, el deseo de los investigadores podría ser comprender las experiencias de las personas con discapacidades sensoriales en centros de enseñanza segregada y en qué se diferencian de los centros de inclusión. Las descripciones que aportan las entrevistas y los estudios etnográficos serían esenciales para acumular la información que diera respuesta a esta pregunta. Por otra parte, puede que los investigadores quisieran describir las características

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

de la población a la que se prestan servicios en los distintos centros, en materia de género, raza, edad, gravedad de la discapacidad, nivel de formación académica y otras cuestiones. Las descripciones precisas de esta naturaleza requieren métodos de muestreo adecuados y métodos estadísticos descriptivos e inferenciales. Los métodos que calculan la magnitud de las relaciones entre características medidas e inferidas pueden proporcionar también información útil para poder comprender lo que sucede.

DetECCIÓN DE LA RELACIÓN CAUSAL

Cuando el núcleo de las preguntas de la investigación se orienta a determinar los efectos causales, los experimentos que emplean la asignación aleatoria para conformar los grupos de tratamiento y control proporcionan la base referencial más justificable sin causa lógica contra la que estimar los efectos de una intervención (Cook, 2002). Aunque los diseños experimentales que utilizan asignación aleatoria proporcionan esta línea de base, hay que reconocer que existen ocasiones en las que los experimentos aleatorizados son simplemente imposibles o prematuros. Convendría saber, por ejemplo si la magnitud de los efectos del tratamiento varía con la gravedad de la deficiencia sensorial. Esta deficiencia es, sin embargo, una patología preexistente que no es apta para la asignación aleatoria. Hay que confiar en los métodos analíticos (por ejemplo, modelos de ecuaciones estructurales) y lógicos para hacer aseveraciones causales (véase Thompson, Diamond, McWilliam, Snyder y Snyder (2005) para más comentarios).

En último lugar, habría que tener en cuenta que los estudios descriptivos que aborda la cuestión de qué sucede juegan un papel básico en la determinación de la validez interna de la investigación experimental. La investigación descriptiva aborda el tema de lo que de hecho sucede al poner en práctica la intervención. En investigación experimental, las preguntas relacionadas con la fidelidad de la implementación de una intervención juegan un papel fundamental a la hora de responder preguntas que buscan explicar las relaciones causales.

Explicar una relación causal

Cuando en los experimentos aleatorizados se encuentran evidencias consistentes de una relación causal, el foco gira hacia las preguntas que intentan explicar el proceso cau-

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

sal. Si hay diferencias en el desarrollo de la alfabetización de los estudiantes que cursan estudios en centros de inclusión educativa relacionados con los contextos independientes, los investigadores desean saber a qué se debe. En este caso, la investigación descriptiva juega un papel esencial a la hora de generar potenciales explicaciones para hipotéticas relaciones causales. Las descripciones de las interacciones entre personas que se encuentran en centros independientes e inclusivos, o de estudios sobre la forma en que se desarrollan los procesos cognitivos en los dos tipos de centros, a través de estudios que permiten que los alumnos externalicen sus procesos cognitivos mientras hacen el esfuerzo de leer y escribir, son proclives a incluir puntos de vista sobre estrategias que son procesos clave para llegar a desarrollar la alfabetización. Una vez se han definido claramente los procesos cognitivos a través de estudios descriptivos, pueden diseñarse las intervenciones en torno a ellos y evaluar su eficacia en experimentos controlados.

Eliminar las barreras a la inferencia causal

Diseñar y poner en práctica los diseños de investigación más sencillos que sustenta la inferencia causal es tarea abrumadora. Por ello, el propósito del investigador no es encontrar el fallo en la investigación publicada, sino animar a que se continúe refinando la investigación en curso para encontrar vías que proporcionen evidencia de eficacia incuestionable. Hay dos diseños de investigación básicos que cumplen los estándares de evidencia para la validez causal: los experimentos aleatorizados y los diseños de regresión discontinua. Los cuasiexperimentos con grupos equivalentes cumplen estándares de evidencia con reservas [los estándares de diseño para diseños de regresión discontinua y diseños de caso único siguen estando en fase de revisión y se consideran estándares piloto (What Works Clearinghouse, 2014)].³ Los diseños cuasiexperimentales, que pueden establecer validez causal e incluyen grupos no equivalentes con múltiples pretest, los diseños de factor único intrasujeto y los diseños de caso único se describen todos a continuación.

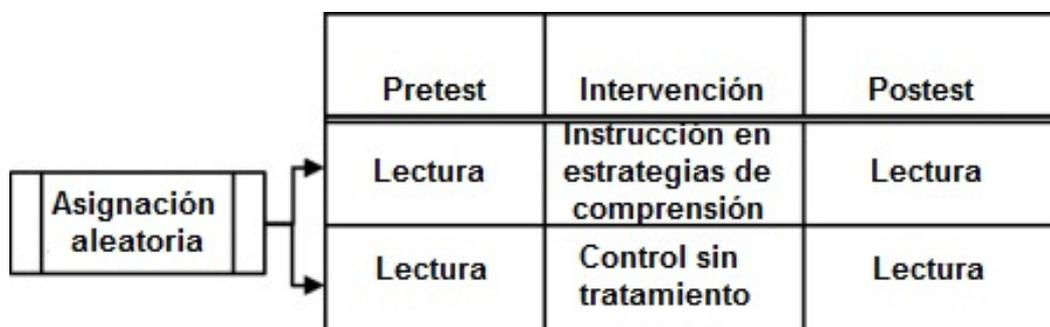
Experimentos aleatorios

La Figura 1 muestra los componentes de los experimentos aleatorios que apoyan las afirmaciones sobre validez causal. El componente más importante de este diseño

³ Entre corchetes en el original. [Nota del editor].

de investigación es la asignación aleatoria de los participantes en el estudio a las condiciones de control de tratamiento y no tratamiento. La asignación de los participantes al tratamiento y a los grupos de control es el mejor método para equilibrar cualquier diferencia entre dos grupos de participantes al inicio del estudio. Así pues, el grupo que no recibe la intervención actúa como la línea de base sin causa más lógicamente defendible para evaluar los efectos de la intervención. Aunque no sea necesario incluir un pretest, el hecho de hacerlo puede ser utilizado para aumentar el potencial del test estadístico a la hora de detectar la presencia de los efectos del tratamiento, cuando estos se tratan como covariable para reducir la variancia del error.

Figura 1. Experimento aleatorizado con pretest y posttest



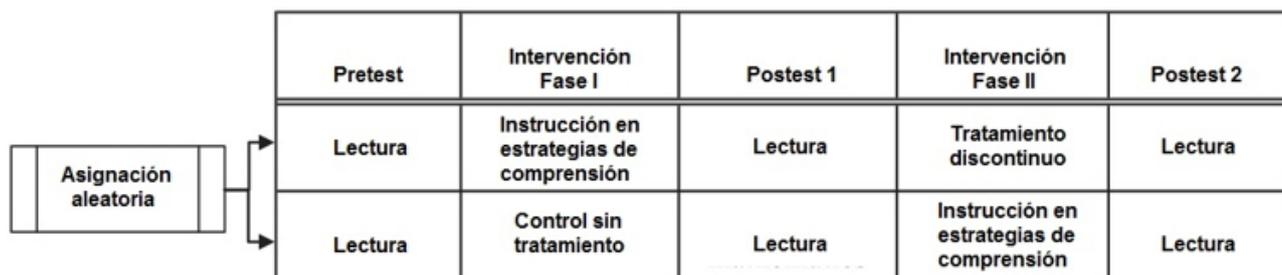
Los estándares de diseño *The What Works Clearinghouse* (2014) requieren de un mínimo de 350 sujetos para efectuar una prueba controlada aleatoria (PCA), cifra que es, de lejos, imposible de alcanzar en el campo de la discapacidad sensorial. No obstante, una obvia objeción a los experimentos aleatorizados es que una intervención potencialmente beneficiosa haya de ser ocultada a algunos participantes. Aun cuando se desconozca la eficacia de la intervención o si esta puede ser perjudicial, e incluso aunque un sorteo sea la forma más justa de hacer la asignación, la utilización de tales diseños resulta cuestionable desde una perspectiva ética (Cook y Campbell, 1979). De hecho, son muchas las objeciones filosóficas y prácticas que se pueden utilizar en los experimentos aleatorizados (Cook, 2002) y, dado que los cuasiexperimentos implican una retirada de la intervención y que no proporcionan evidencias claras a la validez causal, se hace difícil justificar la elección de los cuasiexperimentos frente a la opción de los experimentos aleatorios.

Una estrategia para superar las objeciones, respecto a la retirada de una intervención que se percibe como válida, es la de utilizar un diseño con grupo en lista de espera que simplemente retrase la introducción de la intervención para uno de los

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

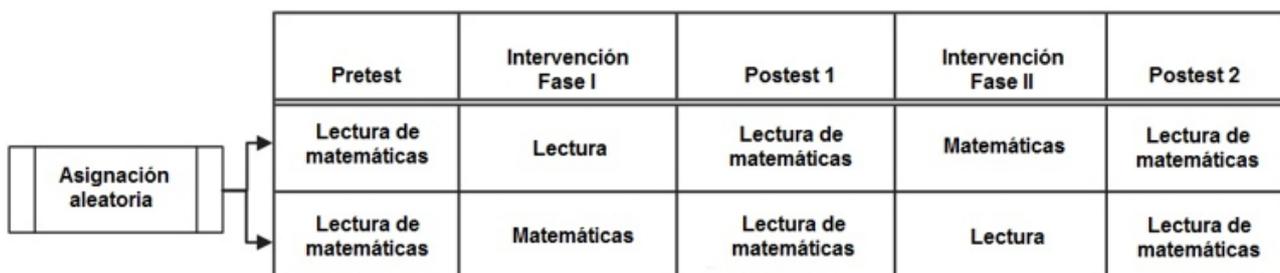
grupos de participantes. Un ejemplo de este diseño es el que se muestra en la Figura 2. Utilizando múltiples posttest, el diseño emplea la asignación aleatoria para situar a los participantes en el estudio en las condiciones de intervención inmediata (Fase I) o posterior (Fase II).

Figura 2. Experimento aleatorio de grupo de control en lista de espera con pretest y múltiple posttest



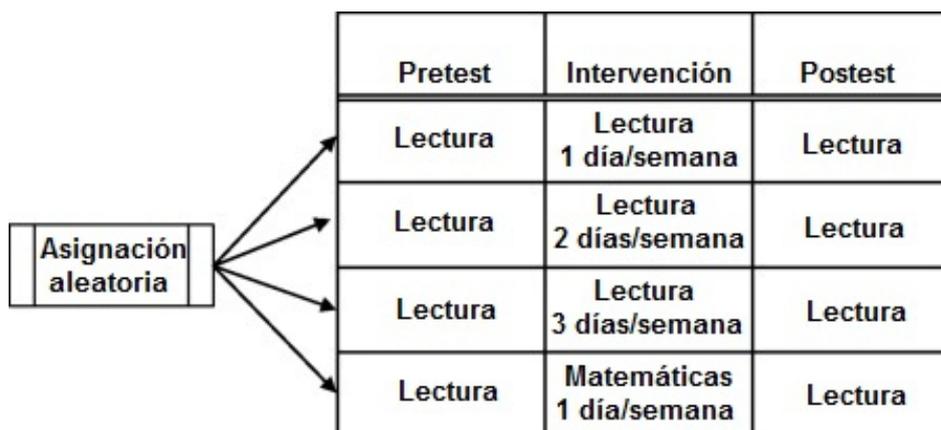
Otra estrategia para superar las objeciones a la retirada del tratamiento es la de realizar, de hecho, dos estudios en un mismo estudio. Atkinson (1968) asignó aleatoriamente a los estudiantes para que recibieran una instrucción asistida por ordenador, tanto para la lectura como para las matemáticas. Ambos grupos completaron ambos temas en los pretest y los posttest para, posteriormente, intercambiar las intervenciones. El grupo que estaba recibiendo la formación en lectura asistida por ordenador pasó a recibirla en matemáticas, mientras que los que inicialmente recibieron la formación en matemáticas pasaron a recibirla en lectura asistida por ordenador. Tras el segundo tratamiento de la Fase II, ambos grupos completaron un segundo posttest en ambas materias. En el diseño de tratamiento múltiple (véase Figura 3), cada grupo actúa como grupo de control (la línea de base no existe) del otro grupo, a fin de calcular el impacto de las intervenciones. La clave para utilizar este diseño es que las dos intervenciones sean independientes en términos de impacto en el funcionamiento de los campos que se evalúan (aunque los logros de los alumnos en materia de lectura y matemáticas estén probablemente correlacionados, no lo es que la enseñanza de las matemáticas afecte a la lectura, o que la formación lectora ejerza impacto alguno sobre el razonamiento matemático). En la medida en que ambas intervenciones son valoradas positivamente, el diseño consigue superar las objeciones sobre la retirada del tratamiento, incluso cuando se ha completado la primera fase. La precisión en la estimación de la magnitud del efecto del tratamiento puede incrementarse en el análisis si se utilizan las puntuaciones en lectura y matemáticas del pretest de los alumnos como covariantes en el análisis del modelo de estructura de covarianza.

Figura 3. Experimento aleatorio con tratamientos múltiples, pretest y múltiples postest



Experimentos aleatorios con diseños de variabilidad planificada

Figura 4. Experimento aleatorio con variación de intervención planificada



Una variante del diseño de tratamiento múltiple es el diseño de variabilidad planificada (Rivlin y Timpane, 1975), aunque también este precisa un mínimo de 350 participantes estándares para el diseño (What Works Clearinghouse, 2014). Siguiendo con el ejemplo arriba mencionado, un investigador puede modificar el nivel o la dimensión del tipo de instrucción a 20 minutos en uno, tres, o cinco días por semana (un enfoque reminiscente del diseño dosis-respuesta, frecuentemente empleado en la investigación médica). Sin embargo, en ausencia de un grupo de control, este diseño solo proporciona una estimación del impacto relativo de una intervención. Añadir un grupo de control que recibe una valiosa intervención, sin impacto sobre el resultado de interés principal, proporciona la necesaria línea de base *sin actividad* para evaluar el efecto total (Cook y Campbell, 1979). El diseño de variabilidad programada se recoge

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

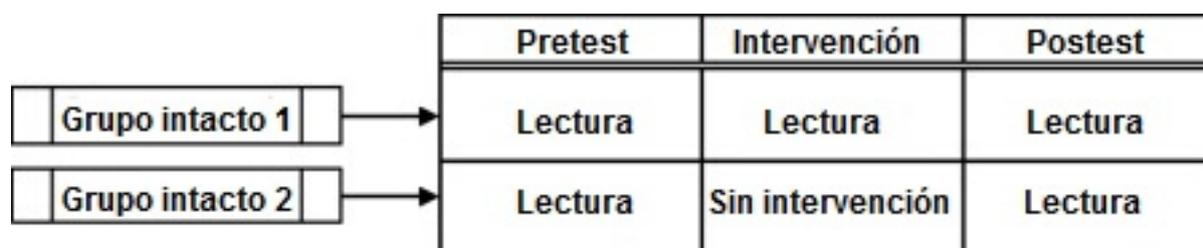
en la Figura 4, incluyendo la mención a la creación de un grupo de control sobre el que se interviene en matemáticas. En la medida en que la intervención en matemáticas no se espera que tenga impacto alguno en el resultado de la medición de la lectura, puede considerarse que no existen causas de referencia para valorar los efectos de los distintos niveles de tratamiento.

Cuasiexperimentos

Diseños de grupo de control no equivalente

Los diseños de investigaciones basados en grupos intactos de participantes, para que formen grupos experimentales y de control, han sido los que se han utilizado con mucha más frecuencia para investigar la alfabetización de personas con discapacidad sensorial (la Figura 5 recoge el diseño de grupo no equivalente que se encuentra más a menudo). El diseño incluye frecuentemente un grupo de control no tratado, y medidas pretest y posttest. Dado que se desconoce el proceso de formación de los grupos y que los participantes no pueden equipararse en todas y cada una de las características que pueden afectar a los resultados, hay muchas diferencias potenciales entre los dos grupos que podrían aparecer como efectos de la intervención, aun cuando los investigadores tomen medidas extraordinarias para encajar los dos grupos y asignar aleatoriamente los grupos a las condiciones experimentales y de control. Dependiendo del resultado del experimento, tales resultados podrían ser o no ser interpretables (Cook y Campbell, 1979; Shadish, Cook y Campbell, 2002). Las cuatro principales amenazas a las inferencias causales que hay que eliminar en este diseño son: selección-madurez, instrumentación, regresión estadística diferencial e historia local.

Figura 5. Los cuasiexperimentos de grupos no equivalentes con pretest y posttest son los más frecuentes

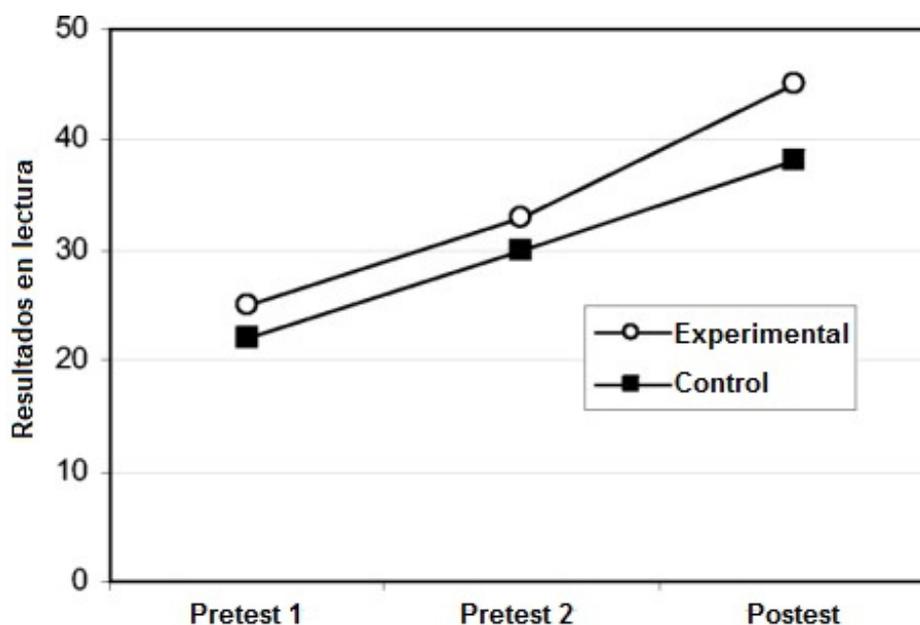


COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

Grupos no equivalentes con múltiples pretest

La interpretación de los diseños no equivalentes con un grupo de control puede mejorarse considerablemente si se añade al diseño uno o más pretest (véase Figura 6). Para eliminar las amenazas a las inferencias causales basadas en los factores de selección y madurez y en la regresión estadística, puede resultar útil incluir un pretest adicional. Si los dos grupos presentan niveles de crecimiento similares entre los dos pretest, ello eliminaría la hipótesis rival sobre madurez diferencial entre los dos grupos (analizar los resultados hipotéticos que aparecen en la Figura 7, comparándolos con el diseño de la Figura 6). Si el segundo pretest fuese el único, cabría argumentar que el grupo experimental ya era el aventajado y que su superior rendimiento en el postest reflejaba un aprovechamiento anterior que nada tenía que ver con la intervención. Sin embargo, el pretest adicional conduce a una interpretación diferente. Aunque el grupo experimental manifieste en el primer pretest un escaso nivel de logro superior en competencia lectora, los dos grupos (experimental y de control) llevan un desarrollo aproximadamente similar entre los dos pretest. Por tanto, como explicación de las diferencias en el postest entre los dos grupos de participantes puede deducirse la existencia de una madurez diferencial, aunque sea necesario tomar en cuenta las diferencias entre las medidas de los dos grupos en el pretest, así como su correlación.

Figura 6. Resultados hipotéticos de un experimento con pretest múltiples



COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

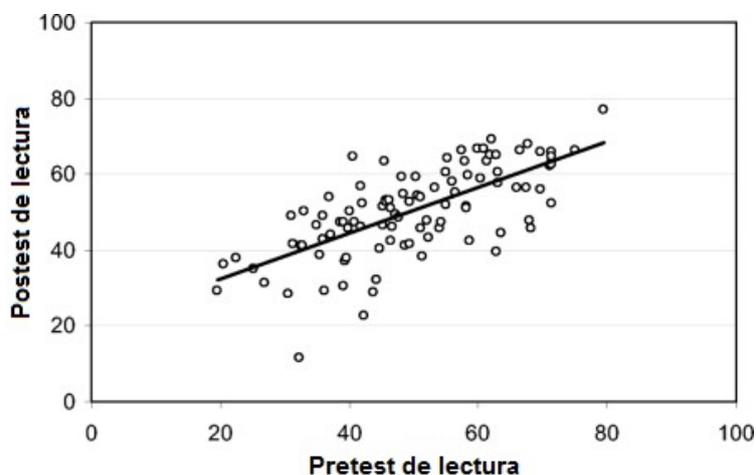
Figura 7. Grupos no equivalentes, cuasiexperimento, con pretest y postest múltiples



Diseños de regresión discontinua

Debido a la sencillez de este diseño y a su capacidad para controlar las amenazas a su validez interna, resulta sorprendente que no se utilice con más frecuencia —los estudios de Ferrell (2006) y Luckner (2006) no identificaron ningún estudio que utilizara un diseño de regresión discontinua—. Supongamos que una persona ha realizado una prueba de lectura al inicio y término del trimestre escolar. Para una habilidad como la lectora, sería razonable arrojar una correlación positiva, entre moderada y elevada, como la presentada entre los dos grupos de puntuaciones de la Figura 8.

Figura 8. Hipotéticos efectos de la intervención sobre la relación entre puntuaciones obtenidas en los pretest y postest de la evaluación lectora de los alumnos

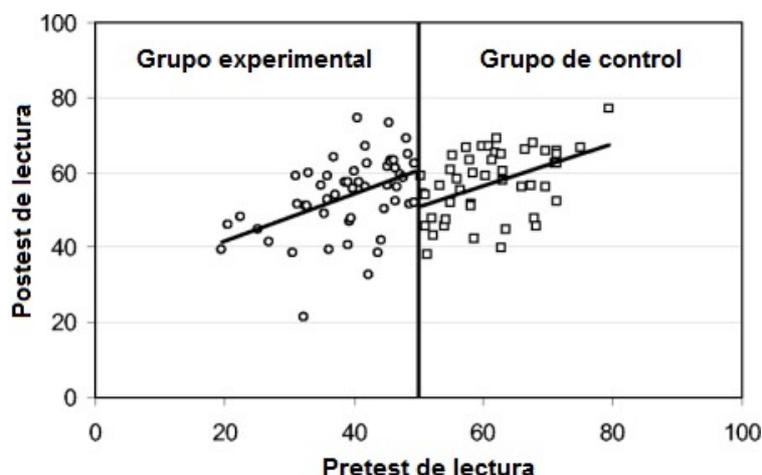


Además, supongamos que un investigador decide asignar a todos los estudiantes unas puntuaciones para la lectura inferiores a una determinada puntuación de corte (por ejemplo, la nota media), a efectos de ser beneficiario de una intervención para la alfabetización. Si dicha intervención fuera efectiva, ¿cómo afectaría a la esperada relación entre puntuaciones del pretest y postest? La Figura 9 presenta la visualiza-

ción de un simple efecto principal de la intervención. Si el tratamiento fuese efectivo, se produciría una discontinuidad en la línea de regresión que describe la relación entre puntuaciones del pretest y postest. Lord y Novick (1968) demostraron que la discontinuidad en la línea de regresión es una estimación objetiva del efecto de la intervención, si los requisitos básicos del diseño pueden satisfacerse:

1. Si, y solo en ese caso, la asignación de participantes está basada en la nota de corte;
2. si la nota de corte es la media de la distribución; y
3. si se conoce la forma funcional de la relación entre variable de asignación y variable dependiente.

Figura 9. Efectos hipotéticos de la intervención sobre la relación entre puntuaciones de la evaluación de la lectura de los estudiantes en el pretest y el postest



En el informe *Technical Methods Report* (2008), elaborado y publicado por el Institute of Education Sciences en su página web, se puede obtener más información sobre los diseños de regresión discontinua.

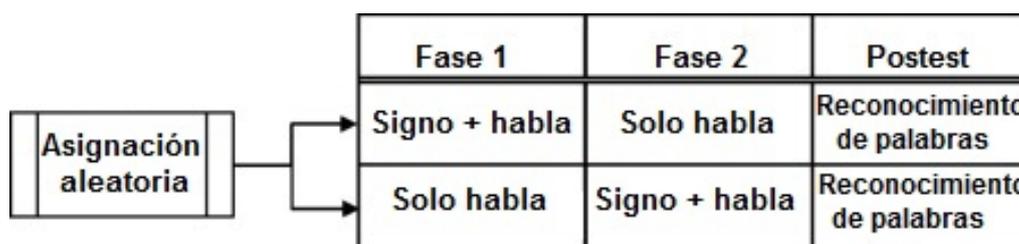
Diseños unifactoriales intrasujeto

En ciertos casos, se puede demostrar la evidencia para la validez causal utilizando un diseño en el que los participantes actúan como su propio grupo de control, participando de todas las situaciones del experimento (Keppel y Zedeck, 1989). En la Figura

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

10, los participantes se asignan aleatoriamente para recibir ambas intervenciones, pero con arreglo a dos órdenes diferentes. Por ejemplo, Wauters [, Knoors, Vervloed y Aarnoutse] (2001) investigaron si los estudiantes con sordera (pérdida auditiva media de 104 dB) reconocían las palabras escritas que habían aprendido mediante signos (lengua de signos holandesa) y hablaban en lengua oral más eficazmente que cuando las palabras escritas habían sido aprendidas únicamente mediante el habla. Se entrenó a los participantes en ambas situaciones por orden diferente y con distintas listas de palabras; se equilibró tanto el orden de las intervenciones como las listas de palabras a efectos de la práctica y de la lista de palabras. A falta de compensación, cualquier diferencia entre las dos situaciones de aprendizaje podría ser atribuida a su orden o a algunas diferencias desconocidas en las listas de palabras. Wauters [et al.] (2001) avanzaron un paso más al incluir una tercera lista de palabras que no había formado parte del entrenamiento; de esta forma, las dos condiciones del aprendizaje podrían evaluarse tanto en términos absolutos como en términos relativos.

Figura 10. Experimento aleatorizado intrasujeto



No obstante, hay una clara situación en la que el diseño unifactorial intrasujeto no sería considerado apropiado. Cuando una intervención tiene un efecto continuado o prorrogado, es posible que pueda interaccionar con una de las otras condiciones para producir un efecto prorrogado diferencial. En este caso, el equilibrio no controla la amenaza a la validez causal. Por ejemplo, supongamos que un investigador está investigando los efectos de una estrategia para la comprensión lectora (A) relativa a las habituales estrategias lectoras de los estudiantes (B). Es poco probable que los alumnos que reciben la estrategia de comprensión lectora (A) retornen en un primer momento y por completo a sus hábitos de lectura habituales (B), incluso cuando se les dice que lo hagan. Por ello, los resultados de la lectura en las dos situaciones no constituyen una justa comparación, debido a que los estudiantes que reciben una orden AB es probable que tengan un mejor funcionamiento en condiciones B que los estudiantes que reciben la orden BA. Es decir, la intervención A lleva a la intervención B, pero la B no conduce a la intervención A.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

Diseños de caso único

Otra categoría de diseños de investigación, en la que los participantes actúan como su propio control, es la de los diseños de caso único (Kazdin, 1982). Aunque existen muchas variaciones de los diseños de caso único, todas comparten la propiedad de repetidas mediciones de la variable de resultados en distintas condiciones, como antes y después de la introducción de una intervención o de intervenciones alternativas (una intervención que implica un cambio de criterio para el funcionamiento). Los diseños de caso único brindan una gran flexibilidad, debido a que las unidades de análisis pueden consistir en un único individuo, varios individuos, diadas, pequeños grupos, clases escolares o incluso instituciones (Levin, O'Donnell y Kratochwill, 2003) y que, cuando están correctamente aplicados, los diseños de caso único pueden proporcionar sólidas pruebas para la validez causal (Kratochwill et al., 2010; Kratochwill y Levin, 2010), aunque la generalización a una población de mayor tamaño siga siendo limitada debido a que el tamaño de la muestra es demasiado pequeño.

En este artículo nos hemos centrado en el diseño de línea base múltiple. Es un diseño flexible de caso único que, según argumentan algunas personas, ofrece el máximo sustento a la generación de inferencias causales sobre eficacia de las intervenciones (Kratochwill y Levin, 2010). La Figura 11 muestra una representación esquemática de un posible diseño de línea base múltiple. Este tipo de diseño se caracteriza por varios elementos significativos que son esenciales para reforzar la validez interna. Su sencillez radica en que conlleva solo dos fases por caso: una fase de línea de base (A) y una fase de intervención (B). Un sencillo diseño AB de caso único no controlaría la amenaza a la validez interna o la generalización del apoyo, y no cumpliría los estándares de la evidencia para la formulación de reclamaciones causales sobre una determinada intervención. Sin embargo, cabe señalar que las fases AB se replican en tres unidades o casos (por ejemplo, los estudiantes). La repetición de un efecto de la intervención en los casos es un requisito esencial para demostrar el control del experimento, pero no es suficiente. Si la introducción de la intervención se produce simultáneamente en cada caso, es que no se han controlado muchas de las amenazas a la validez interna, tales como los sucesos concurrentes. Sin embargo, escalonar con cada sujeto la introducción de la intervención en diferentes momentos y puntos permite controlar muchas amenazas a la validez interna. Si la medición del resultado cambia en el momento en que se introduce la intervención, y únicamente en ese momento, el investigador tiene constancia de la relación funcional o causal entre intervención y medición del resultado.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

Figura 11. Representación esquemática del diseño de línea base múltiple

Caso	Observación														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B
1	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B
3	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B
2	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B

Nota: A = fase de línea base; B = fase de intervención.

El control adicional sobre las amenazas a la validez interna puede introducirse en el diseño de la línea base múltiple mediante dos esquemas de aleatorización (Kratochwill y Levin, 2010). Una vía de introducción de la aleatorización en este diseño es la asignación al azar de los participantes a las condiciones de intervención escalonadas. Adicionalmente, el investigador podría asignar aleatoriamente el tiempo de intervención a cada participante. No obstante, la asignación aleatoria deja abierta la posibilidad de tener una única línea de base u observación tras la intervención, e idénticos puntos de inicio en los casos. Un enfoque alternativo es el procedimiento de aleatorización regulado (Koehler y Levin, 1998).

En la aleatorización regulada, el investigador especifica el número mínimo de observaciones que se requieren para las fases de línea base y de intervención del estudio, al analizar el número total de observaciones existentes. En la Figura 11 se tomó la decisión de contar con un mínimo de tres observaciones de la línea base, con un mínimo de cinco observaciones tras la intervención. Dado que los intervalos de medición que existían eran 15, la intervención podría introducirse antes de la cuarta observación pero nunca después de la décima. En consecuencia, los puntos de intervención posibles que podrían seleccionarse al azar eran siete. Para mantener con claridad la separación temporal de la introducción de la intervención en todos los casos, la aleatorización regulada permite que el investigador especifique un intervalo de puntos de inicio para la intervención, de modo que no se produzcan solapamientos de los puntos de inicio entre casos (Koehler y Levin, 1998). Al disponer de 15 periodos de observación, el investigador podría optar aleatoriamente por iniciar la intervención antes de la cuarta o quinta observación para un caso; con el siguiente, el investigador podría utilizar al azar la selección para iniciar la intervención antes de la sexta o séptima observación, y así sucesivamente en los dos casos restantes. La Figura 11 recoge el resultado de la asignación aleatoria de puntos de inicio para la intervención y asignación aleatoria de los estudiantes a los puntos de inicio. Si hay

más de cuatro casos, los estudiantes se asignan aleatoriamente a uno de los cuatro puntos de inicio.

En suma, la replicación de ocasiones para observar los efectos de la intervención, la separación temporal de puntos de inicio de la intervención en las replicaciones, la utilización de la asignación aleatoria para estructurar los puntos de inicio y la asignación al azar de casos a los puntos de inicio refuerzan la validez interna del diseño de línea base múltiple y cualquier reclamación del resultado sobre la relación causal entre la intervención y la medición de interés resultante.

Para alcanzar los estándares de evidencia, el diseño de línea base múltiple (y otros diseños de caso único) ha de cumplir cuatro criterios (Kratochwill et al., 2010). Primero: el investigador tiene que documentar el control sobre el cómo y el cuándo se implementa la intervención; segundo: es necesario crear un elevado estándar de acuerdo interobservador para la determinación de la medida del resultado (superior al 80 %); este acuerdo interobservador debe calcularse para cada uno de los casos en un 20 % de las observaciones de cada fase (línea base e intervención); tercero: un diseño de línea base múltiple tiene que incluir al menos tres replicaciones de la intervención en tres momentos diferentes; cuarto: cada fase debe incluir cinco datos, como mínimo. Para alcanzar los estándares de prueba con reservas, cada fase tiene que incluir, al menos, tres datos. El diseño de la Figura 11 cumple todos los criterios excepto el último, porque un caso tiene menos de cinco observaciones en una fase de la línea de base. Por lo tanto, este diseño de línea base múltiple cumple los estándares de prueba en las reservas. Si se aumentara el número de los periodos de observación hasta los 18, ello permitiría al observador alcanzar el criterio de un mínimo de cinco observaciones por fase.

Para satisfacer estos criterios, cada estudio debe efectuar un mínimo de tres demostraciones de control experimental sobre la variable del resultado (un efecto) para poder proclamar la «prueba fehaciente» de existencia de una relación causal. En caso contrario, el estudio no proporcionaría «prueba alguna» que sustentara la reivindicación de existencia de una relación causal. Si un estudio incluye tres demostraciones de un efecto y, al menos, un ejemplo de no efecto, entonces es que existe una «evidencia moderada» para la relación causal (Kratochwill et al., 2010). Adicionalmente, se debería contemplar la interpretación de los resultados de un diseño de línea base múltiple, en el contexto de si las intervenciones se definen en términos de constructos teóricos significativos (por ejemplo, soporte, tutoría entre iguales, proceso de supervisión de la comprensión, etc.) y de si la intervención introdujo mejoras conforme a lo previsto.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

Si bien existe un amplio consenso sobre los procedimientos necesarios para controlar las amenazas a la validez interna del diseño de línea base múltiple, hay un gran debate sobre los métodos de detección de la presencia de los efectos de la intervención en diseño de caso único. Análisis visual (Hersen y Barlow, 1976; Kennedy, 2005; Kratochwill et al., 2010), porcentaje de datos que no se solapan (Scruggs y Mastropieri, 1998; 2013), análisis de regresión con inclusión de modelos de crecimiento jerarquizados (Raudenbush y Bryk, 2002, Singer y Willet, 2003) y por pruebas de aleatorización (Edgington, 1975, 1992; Koehler y Levin, 1998; Levin y Wampold, 1999) son tres de los principales métodos propuestos para analizar las series cronológicas que aparecen en diseños de caso único y para detectar los efectos de la intervención. Aunque un planteamiento detallado de estos tópicos exceda el ámbito de este artículo, es importante que se tome en cuenta que, al margen de propio enfoque analítico, no existen claras directrices sobre el modo de integrar los resultados obtenidos de los diseños de caso único en los estudios que utilizan diseños más tradicionales para la comparación de grupos.

Pese a la controversia en torno al cálculo de las estimaciones del tamaño del efecto procedentes de diseños de caso único, las ventajas de un diseño de caso único bien controlado sobrepasan las de diseños intergrupo, en los que no quedan controladas las amenazas a la validez interna. Por ello, se invita a que los investigadores publiquen los datos en bruto, puesto que ello hará posible los reanálisis de datos por parte de los revisores basándose en los nuevos desarrollos. Además, las estimaciones del tamaño del efecto, procedentes de diseños de casos únicos, pueden analizarse separadamente de los tamaños del efecto obtenidos en tradicionales diseños intergrupo, y así identificar lo que de verdad mejor funciona. Wendel, Cawthon, Ge y Beretvas (2015) aportan un útil ejemplo de revisión de estudios sobre sujetos con sordera o deficiencia visual. Como sugieren Kratochwill et al. (2010), el orden de importancia de las intervenciones en estudios intrasujetos y en estudios entre sujetos es probablemente similar, pese a que no se puedan, de hecho, comparar sus estimaciones del tamaño del efecto.

Conclusión

En resumen, los diseños de investigación que hemos destacado en este artículo no son, en absoluto, novedosos. Sin embargo, cada diseño ofrece estrategias útiles que respaldan las afirmaciones sobre la eficacia de las intervenciones con personas con discapacidad sensorial. Indudablemente, describir estrategias para diseñar investiga-

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

ciones que sustenten aseveraciones de causalidad, sobre la eficacia de las intervenciones con sujetos con discapacidades sensoriales, es mucho más fácil que realizar los estudios en los miles de establecimientos en los que investigadores y profesores trabajan para ayudarles. El progreso en este terreno, descubriendo lo que realmente es eficaz, depende de la capacidad de los integrantes del sector para establecer los argumentos que puedan defenderse desde la lógica, en torno a las reivindicaciones causales sobre la eficacia de las intervenciones y la base para la valoración de la magnitud de sus efectos.

Hemos escrito el presente artículo porque confiamos en servir de inspiración a una nueva generación de investigadores, para que piensen en términos más creativos sobre la utilización de los diseños de investigación que pueden fortalecer la evidencia, y para extraer inferencias causales sobre las intervenciones que promuevan el aprendizaje y el desarrollo de personas con discapacidad sensorial. Nuestra revisión ha puesto de relieve varias de las dificultades para realizar investigaciones con poblaciones de baja prevalencia, pero confiamos en haber sugerido también métodos que puedan fortalecer, en un futuro, la investigación. A continuación, ofrecemos una serie de recomendaciones para nuevos y experimentados investigadores, así como para profesores que dirijan a sus alumnos en investigaciones propiamente dichas:

1. Las cuestiones planteadas en investigaciones deberían determinar la selección de la metodología.
2. Los investigadores deberían pensar en términos creativos sobre el modo de diseñar investigaciones que sustenten la validez causal y que, por lo tanto, eliminen las hipótesis rivales.
3. Sin embargo, las cuestiones sobre las relaciones causales requieren descripciones cualitativas sobre la naturaleza y la fidelidad de implementación de las intervenciones.
4. Se deberán comunicar las estadísticas descriptivas que permitan a otros investigadores replicar los análisis y sintetizar los resultados de un estudio, especialmente en relación a otros estudios que examinen los mismos constructos. Aunque las revistas profesionales pudieran requerir cortes en el manuscrito, creemos que tales estadísticas conforman la base de cualquier estudio, y que son esenciales para su interpretación y aplicación por parte de terceros.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

5. Aumentar al máximo su potencialidad, incluyendo covariables que reduzcan la varianza de error en la comprobación de las hipótesis (véase también Wright (2010)).
6. Calcular las estimaciones del tamaño del efecto y los intervalos de confianza para los principales resultados del estudio (Fitz, Morris y Richler, 2012; Kim, 2015; Peng, Long y Abaci, 2012).

Muchas de estas recomendaciones reflejan principios básicos del diseño de investigación. Sin embargo, en un esfuerzo por investigar campos de baja prevalencia que se han visto comprometidos por grupos de población reducidos, extremadamente heterogéneos y enormemente dispersos por la geografía, se hace a menudo difícil reconocer qué es lo que mejor funciona y qué es lo que no funciona. Esperamos que estas recomendaciones sirvan para arrojar luz sobre este ámbito y sobre sus potenciales soluciones.

Referencias bibliográficas

ABOU-GAREEB, I., LEWALLEN, S., BASSETT, K., y COURTRIGHT, P. (2001). Gender and blindness: a meta-analysis of population-based prevalence surveys. *Ophthalmic Epidemiology*, 8, 39-56.

ATKINSON, R. C. (1968). Computerized instruction and the learning process. *American Psychologist*, 23, 225-239.

BARTLETT, D. J., MACNAB, J., MACARTHUR, C., MANDWICH, A., MAGILL-EVANS, J., YOUNG, N. L., BEAL, D., CONTI-BECKER, A., y POLATAJKO, H. J. (2006). [Advancing rehabilitation research: an interactionist perspective to guide question and design \[página web\]](#). *Disability and Rehabilitation*, 28, 1169-1176.

Botsford, K. D. (2013). [Social skills for youths with visual impairments: a meta-analysis \[formato PDF\]](#). *Journal of Visual Impairment & Blindness*, 107, 497-508.

BRANTLINGER, E. V.[, JIMENEZ, R., KLINGNER, J. PUGACH, M., y RICHARDSON, V.] (2005). [Qualitative studies in special education \[formato PDF\]](#). *Exceptional Children*, 71, 195-207.

CAWTHON, S., y LEPPA, R. (2013). Assessment accommodations on tests of academic achievement for students who are deaf or hard of hearing: a qualitative meta-analysis of the research literature. *American Annals of the Deaf*, 158, 363-376.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

- CONROY, P. W. (2007). Paraprofessionals and students with visual impairments: Potential pitfalls and solutions. *RE:view*, 39, 43-55.
- COOK, T. D. (2002). Randomized experiments in educational policy research: a critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis*, 24(3), 175-199.
- COOK, T. D., y CAMPBELL, D. T. (1979). *Quasi-experimentation: design and analysis issues for field settings*. Chicago: Rand McNally.
- CRONBACH, L. J. (1982). *Designing evaluations of educational and social problems*. San Francisco: Jossey-Bass.
- EDGINGTON, E. S. (1975). Randomization tests for one-subject operant experiments. *Journal of Psychology*, 90, 57-68.
- EDGINGTON, E. S. (1992). Nonparametric tests for single-case experiments. En: T. R. KRATOCHWILL y J. R. LEVIN (eds.), *Single-case research design and analysis* (pp. 133-157). Hillsdale, NJ: Erlbaum.
- FERRELL, K. A. (2006). Evidence-based practices for students with visual disabilities. *Communication Disorders Quarterly*, 28(1), 42-48.
- FERRELL, K. A., BRUCE, S., y LUCKNER, J. L. (2014). *Evidence-based practices for students with sensory impairments* [formato PDF] (Documento núm. IC-4). Gainesville, FL: Collaboration for Effective Educator, Development, Accountability, and Reform Center, University of Florida.
- FERRELL, K. A., BUETTEL, M., SEBALD, A. M., y PEARSON, R. (2006). *American Printing House for the Blind mathematics research analysis* [formato PDF] [Informe técnico]. Greeley, CO: University of Northern Colorado, National Center on Low-Incidence Disabilities.
- FERRELL, K. A., DOZIER, C., y MONSON, M. (2011). *Meta-analysis of the educational research in low vision* (sic) [formato PDF].⁴ Greeley, CO: University of Northern Colorado.

⁴ El título real de este informe es *A meta-analysis of educational applications of low vision research* [Nota del editor].

FERRELL, K. A., MASON, L., YOUNG, J., y COONEY, J. (2006). *Forty years of literacy research in blindness and visual impairment* [formato DOC] [Informe técnico]. Greeley, CO: University of Northern Colorado, National Center on Low-Incidence Disabilities.

FITZ, C. O., MORRIS, P. E., y RICHLER, J. J. (2012). *Effect size estimates, current use, calculations, and interpretation* [formato PDF]. *Journal of Experimental Psychology*, 141, 2-18.

FORSTER, E. M., y HOLBROOK, M. C. (2005). Implications of paraprofessional supports for students with visual impairments. *RE:view*, 36, 155-163.

GERSTEN, R. S.[, FUCHS, L. S., COMPTON, D., COYNE, M., GREENWOOD, C., e INNOCENTI, M. S.] (2005). *Quality indicators for group experimental and quasi-experimental research in special education* [formato PDF]. *Exceptional Children*, 71[(2)], 149-164.

GRAEME, D., MCLINDEN, M., MCCALL, S., PAVEY, S., WARE, J., y FARRELL, A. M. (2011). Access to print literacy for children and young people with visual impairment: findings from a review of literature. *European Journal of Special Needs Education*, 26, 25-38.

GRIFFIN-SHIRLEY, N., y MATLOCK, D. (2004). Paraprofessionals speak out: a survey. *RE:view*, 36, 127-136.

HARRIS, B. (2011). Effects of the proximity of paraeducators on the interactions of braille readers in inclusive settings. *Journal of Visual Impairment & Blindness*, 105, 467-478.

HERSEN, M., y BARLOW, D. H. (1976). *Single-case experimental designs: strategies for studying behavior change*. Nueva York: Pergamon Press.

HORNER, R. H.[, CARR, E. G., HALLE, J., MCGEE, G., y ODOM, S.] (2005). *The use of single-subject research to identify evidence-based practice in special education* [formato PDF]. *Exceptional Children*, 71[(2)], 165-179.

Individuals with Disabilities Education Act [formato PDF], 20 U. S. C. § 1400 (2004).

INSTITUTE OF EDUCATION SCIENCES, NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE (2008). *Technical methods report: statistical power for regression*

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

discontinuity designs for education evaluation [formato PDF]. Washington, DC: U. S. Department of Education.

JOINT COMMITTEE ON INFANT HEARING (2007). *Year 2007 position statement: principles and guidelines for early hearing detection and intervention programs* [formato PDF]. *Pediatrics*, 120, 898-921.

KAZDIN, A. E. (1982). *Single-case research designs: methods for clinical and applied settings*. Nueva York: Oxford University Press.

KELLY, S. M., y SMITH, D. W. (2011). The impact of assistive technology on the educational performance of students with visual impairments: a synthesis of the research. *Journal of Visual Impairment & Blindness*, 105, 73-83.

KENNEDY, C. H. (2005). *Single-case designs for educational research*. Boston: Allyn and Bacon.

KEPPEL, G., y ZEDECK, S. (1989). *Data analysis for research designs*. Nueva York: W. H. Freeman.

KIM, D. S. (2015). Power, effect size, and practical significance: how the reporting in *Journal of Visual Impairment & Blindness* articles has changed in the past 20 years. *Journal of Visual Impairment & Blindness*, 109, 214-218.

KOENIG, A. J., y HOLBROOK, M. C. (2000). Professional practice. En: M. C. HOLBROOK y A. J. KOENIG (eds.), *Foundations of education, Vol. 1: History and theory of teaching children and youths with visual impairments* (2.ª ed., pp. 260-276). Nueva York: AFB Press.

KOEHLER, M. J., y LEVIN, J. R. (1998). *Regulated randomization: a potentially sharper analytical tool for the multiple-baseline design* [formato PDF]. *Psychological Methods*, 3(2), 206-217.

KRATOCHWILL, T. R., HITCHCOCK, J., HORNER, R. H., LEVIN, J. R., ODOM, S. L., RINDSKOPF, D. M., y SHADISH, W. R. (2010). *Single-case designs technical documentation* [formato PDF]. [Washington, DC: U. S. Department of Education].

KRATOCHWILL, T. R., y LEVIN, J. R. (2010). *Enhancing the scientific credibility of single-case intervention research: randomization to the rescue* [formato PDF]. *Psychological Methods*, 15(2), 124-144.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

- LEVIN, J. R., y WAMPOLD, B. E. (1999). Generalized single-case randomization tests: flexible analyses for a variety of situations. *School Psychology Quarterly*, 14, 59-93.
- LEVIN, J. R., O'DONNELL, A. M., y KRATOCHWILL, T. R. (2003). Educational/psychological intervention research. En: I. B. WEINER (ed. de la serie), W. M. REYNOLDS y G. E. MILLER (eds. del volumen), *Handbook of psychology, Vol. 7: Educational psychology* (pp. 557-581). Nueva York: Wiley.
- LEWIS, S., y MCKENZIE, A. R. (2010). The competencies, roles, supervision, and training needs of paraeducators working with students with visual impairments in local and residential schools. *Journal of Visual Impairment & Blindness*, 104, 464-477.
- LORD, F. M., y NOVICK, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- LUCKNER, J. L. (2006). Evidence-based practices with students who are deaf. *Communication Disorders Quarterly*, 28, 49-52.
- LUCKNER, J. L., y COOKE, C. (2010). [A summary of the vocabulary research with students who are deaf or hard of hearing \[formato PDF\]](#). *American Annals of the Deaf*, 155, 38-67.
- LUCKNER, J. L., y HANDLEY, C. M. (2008). [A summary of the reading comprehension research undertaken with students who are deaf or hard of hearing \[formato PDF\]](#). *American Annals of the Deaf*, 153, 6-36.
- LUCKNER, J. L., SEBALD, A. M., COONEY, J., YOUNG, J., y MUIR, S. G. (2005, 2006). [An examination of the evidence-based literacy research in deaf education \[formato PDF\]](#). *American Annals of the Deaf*, 150, 443-456.
- LUCKNER, J. L., y URBACH, J. (2012). Reading fluency and students who are deaf or hard of hearing: synthesis of the research. *Communication Disorders Quarterly*, 33, 230-241.
- MARKS, S. U., SCHRADER, C., y LEVINE, M. (1999). [Paraeducator experiences in inclusive settings. Helping, hovering, or holding their own? \[página web\]](#). *Exceptional Children*, 65, 315-328.
- MCKENZIE, A. R., y LEWIS, S. (2008). The role and training of paraprofessionals who work with students who are visually impaired. *Journal of Visual Impairment & Blindness*, 102, 459-471.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

MEINZEN-DERR, J., WILEY, S., y CHOO, D. I. (2011). Impact of early intervention on expressive and receptive language development among young children with permanent hearing loss. *American Annals of the Deaf*, 155(5), 580-591.

MOELLER, M. P. (2000). [Early intervention and language development in children who are deaf and hard of hearing \[formato PDF\]](#). *Pediatrics*, 106(3), E43.

[No Child Left Behind \(NCLB\) Act of 2001 \[formato PDF\]](#), Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

ODOM, S. R., BRANTLINGER, E., GERSTEN, R., HORNER, R. H., THOMPSON, B., y HARRIS, K. R.] (2005). [Research in special education: scientific methods and evidence-based practices \[formato PDF\]](#). *Exceptional Children*, 71[(2)], 137-148.

PARKER, A. T., DAVIDSON, R., y BANDA, D. R. (2007). Emerging evidence from single-subject research in the field of deaf-blindness. *Journal of Visual Impairment & Blindness*, 101, 690-700.

PARKER, A. T., GRIMMETT, E. S., y SUMMERS, S. (2008). Evidence-based communication practices for children with visual impairments and additional disabilities: an examination of single-subject design studies. *Journal of Visual Impairment & Blindness*, 102, 540-552.

PARKER, A. T., y IVY, S. (2014). Communication development of children with visual impairment and deafblindness: synthesis of intervention research. En: D. D. HATTON (ed.), *International Review of Research in Developmental Disabilities*, 46, 101-144.

PARKER, A. T., y POGRUND, R. L. (2009). A review of research on the literacy of students with visual impairments and additional disabilities. *Journal of Visual Impairment & Blindness*, 103, 635-648.

PENG, C.-Y. J., LONG, H., y ABACI, S. (2012). Power analysis software for educational researchers. *The Journal of Experimental Education*, 80, 113-136.

RAUDENBUSH, S. W., y BRYK, A. S. (2002). *Hierarchical linear models* (2.ª ed). Thousand Oaks, CA: Sage Publications.

RIVLIN, A. M., y TIMPANE, P. M. [(eds.)] (1975). *Planned variation in education: should we give up or try harder?* Washington, DC: Brookings Institution.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

RUSSOTTI, J., y SHAW, R. (2001). Inservice training for teaching assistants and others who work with students with visual impairments. *Journal of Visual Impairment & Blindness*, 95, 483-487.

SCRUGGS, T. E., y MASTROPIERI, M. A. (1998). Summarizing single-subject research: issues and applications. *Behavior Modification*, 22, 221-242.

SCRUGGS, T. E., y MASTROPIERI, M. A. (2013). PND at 25: past, present, and future trends in summarizing single-subject research. *Remedial and Special Education*, 34(1), 9-19.

SHADISH, W. R., COOK, T. D., y CAMPBELL, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

SHAVELSON, R. J., TOWNE, L., y COMMITTEE ON SCIENTIFIC PRINCIPLES FOR EDUCATION RESEARCH (eds.) (2002). *Scientific research in education*. Washington, DC: National Academy Press.

SINGER, J. D., y WILLET, J. B. (2003). *Applied longitudinal data analysis*. Nueva York: Oxford University Press.

SNYDER, T. D., y DILLOW, S. A. (2015). *Digest of Education Statistics, 2013* [formato PDF] [Table 204.30]. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

SULLIVAN, G. M. (2011). [Getting off the «gold standard»: randomized controlled trials and education research \[formato PDF\]](#). *Journal of Graduate Medical Education*, 3, 285-289. doi: 10.4300/JGME-D-11-00147.1

THOMPSON, B., DIAMOND, K. E., McWILLIAM, R., SNYDER, P., y SNYDER, S. W. (2005). [Evaluating the quality of evidence from correlational research for evidence-based practice \[página web\]](#). *Exceptional Children*, 71[(2)], 181-194.

VALENTINE, J. C., y COOPER, H. (2004). *What Works Clearinghouse study design and implementation assessment device* (Version 1.1). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

VALENTINE, J. C., y COOPER, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: the

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

Study Design and Implementation Assessment Device (Study DIAD). *Psychological Methods*, 13, 130-149.

VLASTARAKOS, P. V., PROIKAS, K., PAPACHARALAMPOUS, G., EXADAKTYLOU, I., MOCHLOULIS, G., y NIKOLOPOULOS, T. P. (2010). Cochlear implantation under the first year of age — the outcomes: a critical systematic review and meta-analysis. *International Journal of Pediatric Otorhinolaryngology*, 74, 119-126.

VOHR, B., JODOIN-KRAUZYK, J., TUCKER, R., JOHNSON, M. J., TOPOL, D., y AHLGREN, M. (2008). Results of newborn screening for hearing loss: effects on the family in the first 2 years of life. *Archives of Pediatrics & Adolescent Medicine*, 162(3), 205-211.

WANG, Y., y WILLIAMS, C. (2014). [Are we hammering square pegs into round holes? An investigation of the meta-analyses of reading research with students who are deaf or hard of hearing and students who are hearing \[página web\]](#). *American Annals of the Deaf*, 159, 323-345.

WAUTERS, L. N.[, KNOORS, H. E. T., VERVLOED, M. P. J., y AARNOUTSE, C. A. J.] (2001). Sign facilitation in word recognition. *Journal of Special Education*, 35(1), 31-40.

WENDEL, E., CAWTHON, S. W., GE, J. J., y BERETVAS, S. N. (2015). [Alignment of single-case design \(SCD\) research with individuals who are deaf or hard of hearing with the What Works Clearinghouse Standards for SCD Research \[página web\]](#). *Journal of Deaf Studies and Deaf Education Advance Access*, February, 1-12. doi:10.1093/deafed/enu049.

WHAT WORKS CLEARINGHOUSE (2014). [Procedures and standards handbook version 3.0 \[formato PDF\]](#). Washington, DC: Institute of Education Sciences.

WRIGHT, T. (2010). Looking for power: The difficulties and possibilities of finding participants for braille research. *Journal of Visual Impairment & Blindness*, 104, 775-780.

WRIGHT, T., HARRIS, B., y STICKEN, E. (2010). A best-evidence synthesis of research on orientation and mobility involving tactile maps and models. *Journal of Visual Impairment & Blindness*, 104, 95-106.

YOSHINAGA-ITANO, C. (2003a). Early intervention after universal neonatal hearing screening: impact on outcomes. *Mental Retardation and Developmental Disabilities Research Reviews*, 9, 252-266.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.

YOSHINAGA-ITANO, C. (2003b). From screening to early identification and intervention: discovering predictors to successful outcomes for children with significant hearing loss [página web]. *Journal of Deaf Studies and Deaf Education*, 8, 11-30.

YOSHINAGA-ITANO, C., y GRAVEL, J. S. (2001). The evidence for universal newborn hearing screening. *American Journal of Audiology*, 10, 62-64.

YOSHINAGA-ITANO, C., SEDEY, A. L., COULTER, D. K., y MEHL, A. L. (1998). The language of early- and later-identified children with hearing loss. *Pediatrics*, 102, 1161-1171.

COONEY, J. B., YOUNG, J., LUCKNER, J. L., y FERRELL, K. A. (2016). Aprender lo que mejor funciona con la discapacidad sensorial y establecer la inferencia causal. *Integración: Revista digital sobre discapacidad visual*, 68, 148-176.