

ORIGINAL

Recibido: 14/04/2021 | Aceptado: 10/06/2021

Problemas de validez y métodos de experto en investigaciones de la educación especial.

Validity problems and expert methods in special education research.

Mayelín Caridad Martínez Cepena [cepena@uho.edu.cu] 
Doctor en Ciencias Pedagógicas. Prof. Titular.
Universidad de Holguín. Holguín, Cuba.

Miguel Cruz Ramírez [macruz@uho.edu.cu] 
Doctor en Ciencias Pedagógicas. Prof. Titular.
Universidad de Holguín. Holguín, Cuba.

Juan Eduardo Nápoles Valdés [jnapoles@exa.unne.edu.ar] 
Doctor en Ciencias Matemáticas. Prof. Titular.
Universidad Nacional del Nordeste. Corrientes, Argentina.

Resumen

El concepto de validez ha sido explorado desde perspectivas diversas, donde el enfoque basado en argumentos provee una mediación entre lo cualitativo y lo cuantitativo. En el caso de las investigaciones de la educación especial, el problema de la validez se acentúa frente a muestras frecuentemente pequeñas y en contextos relativamente complejos. El presente trabajo aborda esta problemática, tomando como centro los métodos de experto, implementados frecuentemente en la parte conclusiva de investigaciones contemporáneas del ámbito educacional. Con un enfoque de la validez basado en argumentos, se exploran 45 tesis doctorales del campo de la educación especial. El análisis de contenido revela un conjunto de problemáticas relacionadas con la validez, como la necesidad de lograr mayor coherencia, consistencia interna, fiabilidad, entre otros aspectos. A partir de aquí se establecen recomendaciones para el logro de un mayor grado de validez en los procesos investigativos.



Abstract

The concept of validity has been explored from diverse perspectives, where the argument-based approach provides a mediation between the qualitative and the quantitative. In the case of special education research, the problem of validity is accentuated in the face of frequently small samples and in relatively complex contexts. The present work addresses this problem, taking as its center the expert methods, frequently implemented in the conclusive part of contemporary research in the educational field. Using an argument-based approach to validity, 45 doctoral theses from the field of special education are explored. The content analysis reveals a set of problems related to validity, such as the need to achieve greater coherence, internal consistency, and reliability, among other aspects. From here recommendations are established to achieve a higher degree of validity in the investigative processes.

Palabras claves: validez, métodos de experto; educación especial; investigación educacional.

Keywords: validity; expert methods; special education; educational research.

Introducción

Un rasgo característico del método científico consiste en su validez. La aspiración a la validez ha evolucionado con el tiempo, principalmente desde los inicios del siglo XX, donde predominó una visión eminentemente pragmática y signada por el positivismo. Muchas investigaciones pioneras en este campo, basaban la validez en dos cuestiones fundamentales: (1) el hecho de ser algo inherente al test, y (2) la certidumbre de que la validez es el grado en que el test mide lo que realmente supone medir. El primero de estos postulados ha sido ampliamente superado en la actualidad, ya que la validez se refiere más exactamente a las acciones o interacciones que se hacen respecto a los resultados del test u otro instrumento, e incluso de un



método científico (Sireci, 2009). Un instrumento, por sí solo (el *qué*), no da señales de validez. Es necesario verlo aplicado en un contexto específico, produciendo evidencias fiables, y en correspondencia con objetivos bien definidos (el *dónde* y el *para qué*).

El segundo postulado llevó a Guilford a expresar que, en general, “...un test es válido para algo con lo cual se correlaciona” (Guilford, 1946, p. 429). Bajo esta mirada, el mecanismo fundamental para examinar la validez era el análisis factorial, dirigido a determinar ciertas variables latentes y como una especie de simplificación conceptual. Además de la “validez factorial”, Guilford ya avizoraba lo que denominó “validez práctica”, aunque todavía ataba este constructo a la correlación estadística respecto a un criterio relevante. Pocos años después, en 1952, la *American Psychological Association* adoptó un cambio radical en la forma de enfocar la validez. En sus *Technical Recommendations for Psychological Test and Diagnostic Techniques: A Preliminary Proposal*, aparece la nueva figura de “categorías” de validez: predictiva, de status, de contenido, y congruente. Estas categorías transmutaron sucesivamente, como “atributos” y “tipos” de validez. Ya en 1954 se habían cambiado algunas denominaciones, como validez congruente por validez de constructo, y validez de status por validez concurrente. Desde 1999, con la publicación de los *Standards for Educational and Psychological Testing*, la concepción de validez ha cambiado de manera paulatina (Kane, 2016).

Tradicionalmente, ha sido común encontrar análisis de validez en el marco de investigaciones donde predominan los métodos cuantitativos. En este ámbito, son frecuentes los estudios que separan la validez interna de la externa. Por un lado, la validez interna expresa el grado de confianza en que los resultados de un experimento se interpretan adecuadamente, es decir, responde a la causalidad y no a la casualidad de que una variable dependiente dependa de la manipulación de otras variables independientes (Hernández-Sampieri y Mendoza, 2018). Por



otro lado, la validez externa expresa el grado de generalización en condiciones no experimentales. El primer caso se preocupa por el control experimental, mientras que el segundo pondera la representatividad de la muestra. Asimismo, también subsisten análisis contemporáneos de la validez de constructo, de contenido, aparente, retrospectiva, concurrente, lógica, predictiva, convergente, discriminante o divergente, entre otras.

Por ejemplo, la validez de contenido representa el grado en que un instrumento refleja un dominio específico de contenido de la variable que se mide. Hernández-Sampieri y Mendoza (2018) ponen como ejemplo una prueba de operaciones aritméticas, donde solo se examinan problemas de resta y se excluyen las restantes operaciones de suma, división y multiplicación. La validez de contenido se pierde, pues no es capaz de aprehender todos o la mayoría de sus componentes o dimensiones. En el caso de la validez aparente, no se trata de una validez en su sentido estricto ya que describe la subjetividad percibida y no necesariamente tiene significación real. A modo de ejemplo, la validez aparente tiene lugar en la valoración individual de un experto, acerca de un aspecto objeto de estudio. La experticia favorece una opinión valiosa, pero no exenta de juicios apresurados, de sesgos, excesos de confianza, temores, sobrestimación, u otro aspecto que limite la calidad de la información proporcionada (Cruz, 2021; Landeta, 2006).

Bajo la mirada cualitativa, el problema de la validez ha sido explorado de un modo menos frecuente. Onwuegbuzie y Leech (2007) califican de oxímoron la validez en la investigación cualitativa, pero no como un binomio discordante sino como expresión de una nueva cualidad. Para potenciar la validez como cualidad, estos autores introducen el Modelo de Legitimación Cualitativa, dirigido a integrar varios conceptos relacionados como los niveles de generalización (poblacional, ecológica y temporal), así como la validez catalítica, comunicativa, consensual, entre otros conceptos. En esta investigación se discuten numerosos métodos para evaluar el valor



de verdad en los estudios cualitativos, como el análisis de casos extremos, el chequeo de la representatividad, el contraste y comparación, la observación perseverante, la triangulación, entre otros.

Bajo un enfoque cualitativo heurístico, Kleining y Witt (2000) defienden la idea de que la validez de los resultados está limitada por el carácter social y espacio-temporal de sus objetos de análisis. El rango de validez depende de cuán estrecha o ampliamente esté representado el objeto en los datos. Para estos autores, la regla de “variación estructural máxima de perspectivas” es esencial en la recopilación de los datos, junto a la búsqueda sistemática de similitudes. De este modo, se establece un diálogo entre el investigador y el objeto de investigación, donde la validez se va construyendo en la medida que las evidencias son cada vez más coherentes.

Whittemore et al. (2001) destacan la importancia de procurar la validez de los estudios cualitativos, bajo el reconocimiento de la complejidad que le es inherente a dicho proceso. Con base en la revisión de un amplio número de fuentes, estos autores identifican criterios primarios para la validación, como la autenticidad, integridad, credibilidad y criticismo en los resultados. Luego analizan otros criterios secundarios, como la perceptibilidad, la congruencia, la claridad, la creatividad, entre otros aspectos. Por su parte, Hayashi et al. (2019) conciben la validez de un estudio cualitativo como un proceso contentivo de cinco etapas que median entre el diseño y la recogida de información, hasta su correspondiente procesamiento y análisis. A partir de estas reflexiones, puede resumirse que el problema de la validez no obedece a un enfoque donde predomine un tipo de método, ya sea cualitativo o cuantitativo. Por tanto, resulta razonable considerar algún enfoque que no dependa de la naturaleza de la información, y que supere las características de los métodos e instrumentos de investigación.



En la actualidad, los *Standards* consideran que el concepto de validez responde "...al grado en que la evidencia y la teoría respaldan las interpretaciones de los puntajes de una prueba para usos propuestos de las pruebas. [...] El proceso de validación involucra acumular evidencia pertinente para proporcionar una base científica sólida para las interpretaciones de puntajes propuestas" (AERA et al., 2018). Por tanto, lo que se evalúa no es la prueba propiamente dicha, sino las interpretaciones de los resultados para los usos propuestos. Esta perspectiva se apoya en el enfoque de validación basado en argumentos (Kane, 1992, 2013, 2016), lo cual desestima el establecimiento de cualquier tipología de validez, adoptando en su lugar un proceso de aproximación a la validez. Se trata de punto de vista donde este constructo no puede establecerse absolutamente, y lo mejor que puede hacerse consiste en aportar todas las evidencias disponibles para verificar que los argumentos interpretativos son altamente verosímiles (Sireci, 2009).

Gasmalla y Tahir (2020) sostienen que no existen tipos de validez, sino fuentes de validez. Estas fuentes tienen como soporte los argumentos, pero no vistos de forma aislada sino en forma necesariamente integrada y coherente (Addey et al., 2020). La evidencia de validez es más efectiva cuando se direcciona hacia los puntos débiles de los argumentos interpretativos (Kane, 2013). O sea, es preferible un argumento más en favor de desechar una interpretación incorrecta, a un argumento que adicione evidencias donde ya existen suficientes. Desde el punto de vista práctico, Kane (2013) establece dos momentos: (1) la especificación de la interpretación y uso del instrumento como una red de inferencias y suposiciones, lo cual va desde la aplicación del instrumento hasta el establecimiento de las conclusiones, y (2) la evaluación crítica de los argumentos de interpretación/uso (IUA = *interpretation/use argument*).

Kane (2013) también identifica cuatro tipos de inferencias fundamentales: de puntuación (respecto a la fiabilidad de los datos, a la consistencia de los calificadores, a la claridad de la



regla que asigna puntuaciones por niveles de desempeño), de generalización (hacia otros contextos, en otras ocasiones, con otras tareas, con otros experimentadores), de extrapolación (en qué medida un examen representa el contenido que se desea evaluar, y si logra captar toda la gama posible de niveles de desempeño), basadas en la teoría (cómo justificar la correspondencia de los resultados con determinados constructos teóricos, en qué medida los hallazgos empíricos son coherentes con los presupuestos teóricos, cuán bien los aspectos investigados reflejan un concepto con cabalidad). En este sentido, los *Standards* proveen seis formas específicas de evidencia de validación, las cuales proveen de orden y sistematicidad al proceso analítico. Por ejemplo, en lugar de validez de contenido, se demandan evidencias orientadas al contenido. Asimismo, se sugiere el análisis de evidencias relacionadas con procesos cognitivos, con la estructura interna, con las interrelaciones conceptuales de los constructos, con la pertinencia y calidad técnicas de las variables de criterios, y también con evidencias basadas en consecuencias de las pruebas (AERA et al., 2018).

Las investigaciones que hacen uso del método de criterio de expertos también requieren del análisis de validez. Al respecto, Bolger y Wright (1992) señalan que todo juicio experto requiere de validez cuando su finalidad es científica. Un aspecto primario de la validez consiste en la fiabilidad, pues todo juicio válido es necesariamente fiable, pero un juicio fiable no tiene por qué ser válido. En principio, para estos autores, la fiabilidad del juicio experto debe satisfacer al menos tres requisitos: la consistencia, el consenso y la coherencia. Estos autores destacan una cualidad de la validez que denominan “ecológica”, relacionada con el control sistemático de los factores que moderan el desempeño del experto. Por ejemplo, es necesario equiparar el nivel de complejidad de los cuestionarios con las motivaciones y experiencias de los expertos, los



procedimientos de evaluación con sus términos de fidelidad, las tareas con sus grados de dificultad, entre otros aspectos.

En el presente trabajo se centra en el problema de la validez en investigaciones cubanas de maestría y doctorado, en el campo de la educación especial. A fin de presentar una primera aproximación al estado actual de esta temática, se analiza el camino seguido en aquellas investigaciones que utilizan, fundamentalmente, el método de criterio de expertos para analizar la validez de sus aportaciones teórico-prácticas. El campo de la educación especial es favorable, en el sentido de que permite el desarrollo de investigaciones cualitativas vinculadas a discapacidades muy específicas, junto a otras investigaciones cuantitativas con muestras más numerosas (McNeill, 2019; Pua et al., 2021; Rumrill et al., 2020), e incluso con enfoque mixto (Corr et al., 2020). Por otra parte, el método de criterio de expertos ha sido empleado de forma significativa en investigaciones educacionales cubanas, principalmente durante las dos últimas décadas (Díaz et al., 2020), de modo que resulta viable la presente investigación.

Población y muestra

El presente estudio es de tipo documental y tiene el propósito de conformar una muestra lo suficientemente amplia y representativa, que comprenda la mayoría de las tesis doctorales defendidas en el país en el campo de la Educación Especial, y enmarcadas dentro de las dos últimas décadas. Por consiguiente, la población constituye este tipo de informe de investigación científica, y la muestra responde al propio proceso de conformación, tratando de aprehender el mayor volumen posible respecto a la población.

Se trata de un propósito complejo, pues no siempre los documentos aparecen registrados en centros de documentación. Frecuentemente, esta valiosa información aparece distribuida de modo irregular en repositorios y colecciones institucionales que comprenden una región



específica. Por otro lado, en algunas ocasiones los documentos poseen dispositivos de seguridad que impide la minería de datos, o bien se encuentran en archivos múltiples e incluso incompletos. Todo ello pone en riesgo la consecución de una muestra suficientemente representativa.

La base material del presente estudio se apoya en un repositorio conformado en la universidad de Holguín, contentivo de 1112 tesis doctorales, defendidas desde el año 2001. Este repositorio contiene aproximadamente la tercera parte de todas las tesis doctorales cubanas defendidas en el campo de las Ciencias Pedagógicas, procedentes de instituciones autorizadas por la Comisión Nacional de Grados Científicos, correspondientes a diez provincias del país. El volumen representa aproximadamente el 30% de las tesis de este tipo defendidas por autores cubanos, durante las dos últimas décadas.

Metodología

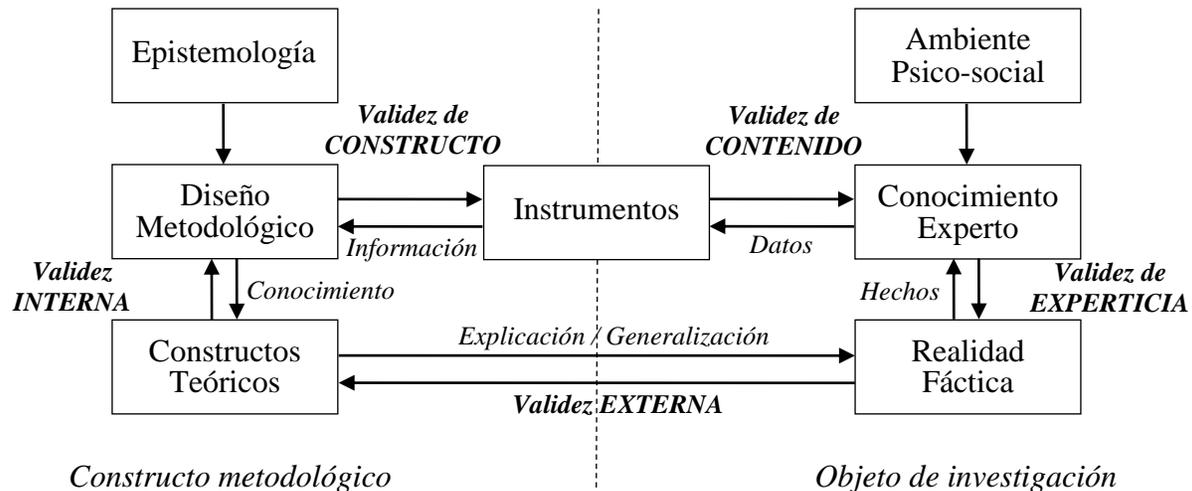
La base metodológica parte de una modelación estructural y funcional del análisis de validez, con enfoque basado en argumentos. Todo el proceso de gestión y procesamiento científico de la información experta se concibe bajo la pirámide datos-información-conocimiento-sabiduría, donde es esencial la postura epistémica subyacente. Al igual que Frické (2019), el presente trabajo toma distancia de posturas mecanicistas que consideran los niveles básicos como componentes triviales de los superiores; en su lugar, se percibe un proceso iterativo de constante actualización y enriquecimiento, con base socio-cultural. El valor de los datos, por ejemplo, está condicionado por su veracidad y nivel de certidumbre dentro de un contexto. Los datos sostienen los niveles superiores bajo un proceso analítico-sintético, pero no solo se reducen a hechos y evidencias a procesar, sino que pueden ser suficientemente relevantes



y convincentes para establecer un nuevo conocimiento. El esquema general se ilustra en la Figura 1 (adaptado de Cruz, 2021).

Figura 1

Un modelo general para la búsqueda de evidencias de validez durante la implementación del método de criterio de expertos (adaptado de Cruz, 2021).



El modelo refleja una dinámica del procesamiento de la información, donde los aspectos de validez no expresan una tipología, sino demandas de argumentos de validez. Las denominaciones no obedecen a una teoría en concreto, sino que constituyen ajustes y adecuaciones al proceso de implementación del método de criterio de expertos en las investigaciones educacionales. En las evidencias de validez perceptiva, la función fundamental consiste en aportar evidencias que vindiquen la legitimidad del conocimiento experto. Por ejemplo, aquí se requieren evidencias relacionadas con el procedimiento de selección de los expertos, los niveles de experticia, la cantidad y diversidad, el hecho de si existió algún tipo de interacción, entre otros aspectos. Los argumentos de contenido reflejan la capacidad de las entrevistas, encuestas, u otro instrumento, para captar los datos con calidad y fiabilidad. Por



ejemplo, la suficiencia del tiempo disponible, la factibilidad del contexto donde tiene lugar la experimentación, la claridad de las preguntas, etcétera.

Las evidencias de validez de constructo inquietan en la correspondencia entre los aspectos objeto de estudio y la capacidad del instrumento para representarlos. Aquí reside la aspiración clásica de que el instrumento logre aprehender la variable investigada. Entre los aspectos a chequear pueden enumerarse no solo los relacionados con la estructura del instrumento y sus ítems, sino también otros vinculados al procesamiento de los datos, y a la complementación de la información con aquella procedente de otros métodos utilizados. Las evidencias de validez interna justifican la necesaria correspondencia entre la información producida, durante el procesamiento de la información, y el constructo teórico investigado. Es útil indagar si la triangulación de fuentes resulta efectiva, si existe consistencia en el nuevo conocimiento, así como coherencia respecto a la epistemología subyacente. Finalmente, se requieren evidencias de validez externa, pues no basta con la robustez y relevancia científica de los hallazgos, sino también su aplicabilidad y potencialidades de generalización en el objeto investigado.

Con ayuda de la tipología de argumentos de validez, contenida en el modelo antes descrito, se desarrolla un análisis de contenido en el conjunto de tesis que conforman la base de datos.

Análisis de los resultados

Por medio de una búsqueda automatizada basada en conceptos típicos de la educación especial, tales como “discapacidad” y “necesidades educativas especiales” (NEE), se obtienen 152 tesis del repositorio. Estas se examinan manualmente y se conforma una base de datos compuesta por 65 tesis doctorales de autores cubanos que comprenden el periodo 2001-2021,



procedentes de universidades de 10 provincias del país. Las tesis abordan una amplia diversidad de objetos de investigación relacionados con discapacidad sensorial (visual y/o auditivas), discapacidad intelectual, limitaciones físico-motoras, trastornos del espectro autista, dificultades del aprendizaje, inclusión de alumnos con NEE asociadas o no a una discapacidad, alteraciones emocionales y del comportamiento, formación inicial y permanente del profesional de la educación especial, entre otros. De las 65 tesis mencionadas, un total de 47 (72.3%) consignan de manera explícita el empleo del método de criterio de expertos, pero dos fueron desestimadas por no carecer de los anexos correspondientes. Definitivamente, resultan 45 tesis doctorales para el presente análisis.

El empleo del método de criterio de expertos se describe en los terceros capítulos, donde 17 casos refieren un Delphi (37.8%). En seis casos se utiliza la terminología “consulta a especialista” y en las restantes “criterio de experto”, sin que se observen detalles acerca de estos conceptos. Un total de 36 tesis emplea el cálculo de un índice de competencia experta (el coeficiente k), consistente en el promedio de dos autoevaluaciones: el coeficiente de conocimiento y el de fuentes de argumentación. Se observa cierta inercia en el cálculo de este coeficiente, obviando aspectos que podrían lacerar su validez, como el origen en campos no pedagógicos (Cruz y Martínez, 2020) y problemas de fiabilidad en los mecanismos diseñados para el procesamiento de la autoevaluación (Cruz y Martínez, 2012). En las restantes tesis se emplean criterios de selección elaborados por los autores, sin evidencias de argumentación, mientras que en 7 casos se combina más de un procedimiento de selección. Todos los informes especifican el número de expertos, que oscila entre 10 y 78 (media = 28.7, desv. típ. = 10.9). Los instrumentos empleados para obtener información experta generalmente aparecen en solitario, y



son elaborados por los propios autores. Se utilizan 22 encuestas, 18 cuestionarios, 3 entrevistas, y dos casos en que no se refiere instrumento alguno.

El empleo de expertos o especialistas adoptan objetivos que pueden estructurarse en la forma finalidad-cualidad-resultado. Por ejemplo, una de estas tesis declara que "...se utilizó el método Delphi [...], el que permitió realizar la validación teórica del modelo y el nivel de pertinencia y relevancia del modelo y la estrategia". Como *finalidad* aquí se trata de validar, la *cualidad* se enmarca en satisfacer niveles de pertinencia y relevancia, mientras que el *resultado* referido consiste un modelo y una estrategia. En general, las investigaciones emplean el método de criterio de experto con las finalidades de valorar (19), evaluar (10), constatar (3), obtener juicios o criterios (3), y ocasionalmente validar, comprobar, corroborar, demostrar, analizar, precisar, y conformar. Las cualidades más frecuentes son la pertinencia (23), la factibilidad (12), y viabilidad (7), aunque también aparecen términos tales como eficacia, relevancia, calidad, coherencia, y valor científico. Por su parte, los resultados consisten en metodologías, estrategias, sistemas, concepciones, modelos, programas, y sistematizaciones, principalmente. En la mayoría de los casos, el método de experto se centra en un solo aporte, pero también se observan casos como el siguiente: "valorar la pertinencia del marco teórico-conceptual, la coherencia en la estructura de la concepción didáctica, y la factibilidad de las orientaciones metodológicas propuestas", con una mirada más abarcadora.

Respecto a la exploración de la validez, un total de 10 tesis refiere este aspecto de modo explícito, pero no necesariamente relacionado con el empleo del método de experto. De forma ocasional, se observan casos donde los autores someten los instrumentos empíricos a estudios pilotos, con la finalidad de perfeccionarlos. Más implícita que explícitamente, en nueve de los estudios se evalúan indicadores u otros aspectos por medio del modelo de Torgerson, con datos



procedentes de escalas tipo Likert. En estos casos, el número de experto alcanza valores generalmente superiores a 30, pero sin discutirse aspectos relacionados con requisitos de normalidad poblacional. La validez de los resultados, en ocasiones, responde al nivel de certeza en que una población de expertos podría aportar un juicio similar ante condiciones similares. Este tipo de cadena de inferencias no indica, por sí misma, que el juicio es necesariamente el más adecuado. Restarían argumentos de validez que legitimen, más específicamente, las posibilidades de la población de expertos y del panel para aportar un juicio suficientemente verosímil acerca del aspecto investigado.

Algunos estudios Delphi constan de una sola ronda, o bien denominan ronda al diligenciado de un cuestionario de autoevaluación que antecede a un cuestionario de evaluación. Ciertamente existen numerosas modificaciones del método Delphi, como un caso especial del método de criterio de experto. Sin embargo, en todos los casos se conservan las rondas de consenso o convergencia de criterios, como un rasgo distintivo en el orden metodológico. En todo estudio Delphi existen oportunidades para inquirir en aspectos de validez. Por ejemplo, Varanda y Freitas (2018) consideran que la implementación del método Delphi adquiere mayor validez, en el sentido de que la opinión colectiva es más objetiva que la individual. También sugieren la necesidad de conformar un panel heterogéneo para “asegurar la validez de los resultados”, lo cual es razonable en procesos que requieren de una pluralidad de miradas. Ya este aspecto ha sido abordado por Landeta (2006), quien señala que no pocas veces se desatienden algunas dificultades que restan validez al método Delphi. Además del problema de la selección, es necesario discutir acerca de la permanencia y calidad de la información que aporta el panel de expertos. Este autor sugiere “ponerse en el lugar del experto”, de manera que podrían sacrificarse



preguntas y rondas, a fin de garantizar la participación del panel y su continuidad en la secuencia de rondas.

Por otra parte, en 15 de las tesis analizadas (33.3%) los autores se esfuerzan por lograr mayor robustez en sus conclusiones, a partir del uso de otros métodos de investigación. Los más frecuentes son pre-experimentos (6) y talleres de socialización y opinión crítica (5), aunque también se observan experimentos, cuasi-experimentos, y experimentaciones parciales sobre el terreno. Sin embargo, no es frecuente que se expliciten los motivos que conducen a combinar estos métodos, en el sentido de reforzar la colección de evidencias sobre determinado aspecto, de complementar información, o bien de integrarla. Por ejemplo, en una de las tesis se señala que: “...la validez de la propuesta es respaldada por la triangulación de los resultados de la consulta a los especialistas y la sistematización de las experiencias vividas en diferentes momentos del proceso de construcción y socialización de la investigación”. Este propósito es loable, pero no se aportan argumentos suficientes que expresen el *para qué* en esa búsqueda de validez.

A modo de resumen, puede señalarse que las tesis analizadas siguen un enfoque metodológico predominantemente mixto, donde los métodos de experto persiguen la valoración o evaluación de la pertinencia y/o factibilidad de los resultados teórico-prácticos. El análisis de la validez no ocupa un lugar central, sino que aparece diluido en el discurso científico, mayormente en un plano implícito. Los argumentos de validez de la experticia (ver Figura 1) utilizan primordialmente el cálculo de un índice de competencia experta basado en la autoevaluación, o bien criterios predefinidos, pero no siempre justificados. Tampoco suelen discutirse elementos que respalden una adecuada percepción del conocimiento experto, acerca de la realidad fáctica concreta, ni se esclarecen todas las condiciones en que tiene lugar el diligenciado de las encuestas, cuestionarios y otros instrumentos empíricos.



Los argumentos de validez externa se concentran en la relación muestra/población de la opinión experta que, si bien constituye un tipo de generalización, todavía dista de las inferencias hacia la realidad fáctica donde se origina el problema de la investigación. Existen amplias potencialidades para la búsqueda de argumentos de validez de contenido, de constructo, y de consistencia interna. En los diseños metodológicos podría especificarse más la finalidad con la cual se combinan o complementan los métodos, a fin de comprender con más claridad el recorrido que adopta la pirámide del conocimiento, desde los datos aportados por los expertos hasta las conclusiones emergentes de los procesos de triangulación. Algunas veces, estos aspectos se esclarecen en las conclusiones generales de la tesis, pero sin la amplitud y especificidad que podría lograrse en las conclusiones particulares del capítulo correspondiente.

Conclusiones

1. Los estudios de contenido constituyen un aspecto importante para el desarrollo de la actividad científica, pues contribuyen a identificar tendencias, a reconocer logros y también a corregir ciertas dificultades que pudieran surgir en la marcha del quehacer investigativo. Las tesis exploradas contienen resultados teóricos y prácticos de indudable calidad científica. Todas son valiosas y muestran un gran esfuerzo por contribuir al noble propósito de educar los niños y las niñas con NEE, desde bases científicas sólidas. El presente estudio ha explorado apenas un aspecto muy puntual, visto de manera transversal y en una muestra relativamente pequeña de tesis doctorales defendidas durante los últimos 20 años a escala nacional. Se trata de una mirada hacia el problema neurálgico de la validez científica, lo cual ha sido reconocido como un problema latente en numerosos ámbitos de las ciencias sociales, donde las ciencias pedagógicas y de la educación ocupan un importante lugar.



2. Las dificultades se identifican desde la óptica de la validez basada en argumentos, un enfoque de mucha actualidad que ha mostrado efectividad para explorar este constructo con amplitud y flexibilidad. Uno de los aspectos más sobresalientes consiste en la posibilidad de explorar la validez no solo desde una mirada predominantemente cuantitativa, sino también mixta e incluso cualitativa. El presente estudio toma como centro el método de criterio de expertos, el cual contiene aspectos cuantitativos relacionados con el procesamiento de los datos, muchas veces combinados con métodos también cuantitativos diseñados para el cálculo de índices de competencia experta. Por este motivo, en las investigaciones seleccionadas predomina el enfoque mixto, fruto de la creatividad de los investigadores por complementar sus hallazgos con otros métodos y técnicas más cualitativas. Restaría, para trabajos ulteriores, un análisis más orientado hacia la exploración de la validez en estudios predominantemente cualitativos, bajo el enfoque basado en argumentos.
3. Las problemáticas identificadas no solo son inherentes al campo de la Educación Especial. Se trata de aspectos metodológicos que exigen de mayor desarrollo conceptual. La selección de los métodos científicos debe tomar en cuenta numerosos aspectos, como con la coherencia en las bases epistémicas para evitar eclecticismo, la necesidad y suficiencia para lograr aprehender todo el objeto de investigación sin redundar, y la fiabilidad y validez para poder mostrar niveles de argumentación de un fenómeno, grados de extrapolación contextual o generalización, coherencia y consistencia interna, entre otros aspectos. Por tanto, a tenor de la importancia de este último aspecto, el presente trabajo desea llamar la atención acerca de la necesidad de profundizar más en aquellos problemas relacionados con la validez de un resultado



teórico-práctico, desde una mirada flexible que supere la impronta positivista que ha acompañado este concepto desde sus orígenes.

Referencias bibliográficas

- Addey, C., Maddox, B. & Zumbo, B. D. (2020). Assembled validity: rethinking Kane's argument-based approach in the context of International Large-Scale Assessments (ILSAs). *Assessment in Education: Principles, Policy & Practice*, 27(6), 1–19.
<https://doi.org/10.1080/0969594X.2020.1843136>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2018). *Estándares para Pruebas Educativas y Psicológicas*. American Educational Research Association.
- Bolger, F., & Wright, G. (1992). Reliability and validity in expert judgment. In: G. Wright, & F. Bolger (Eds.), *Expertise and Decision Support* (pp. 47–76). Plenum Press. https://doi.org/10.1007/978-0-585-34290-0_4
- Corr, C., Snodgrass, M. R., Love, H., Scott, I. M., Kim, J., & Andrews, L. (2020, in press). Exploring the landscape of published mixed methods research in special education: a systematic review. *Remedial and Special Education*.
<https://doi.org/10.1177/0741932520924030>
- Cruz, M. (2021). *Métodos de Experto en el Campo de las Investigaciones Educativas*. Informe del proyecto de investigación PS223LH001-026. Universidad de Holguín.
- Cruz, M., & Martínez, M. C. (2012). Perfeccionamiento de un instrumento para la selección de expertos en las investigaciones educativa. *Revista Electrónica de*



Investigación Educativa, 14(2), 167–179,

<http://redie.uabc.mx/vol14no2/contenido-cruzmtnz2012.html>

Cruz, M., & Martínez, M. C. (2020). Origen y desarrollo de un índice de competencia experta: el coeficiente k. *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*, 19(10), 40–56,

<http://www.relmis.com.ar/ojs/index.php/relmis/article/view/248>

Díaz, Y., Cruz, M., Pérez, M., C., & Ortiz, T. (2020). El método criterio de expertos en las investigaciones educacionales: visión desde una muestra de tesis doctorales. *Revista Cubana de Educación Superior*, 39(1).

<http://scielo.sld.cu/pdf/rces/v39n1/0257-4314-rces-39-01-e18.pdf>

Frické, M. (2019). The knowledge pyramid: the DIKW hierarchy. *Knowledge Organization*, 46(1), 33–46. <https://doi.org/10.5771/0943-7444-2019-1-33>

Gasmalla, H. E. E., & Tahir, M. E. (2020). The validity argument: addressing the misconceptions. *Medical Teacher*.

<https://doi.org/10.1080/0142159X.2020.1856802>

Guilford, J. P. (1946). New standards for test evaluation. *Educational and Psychological Measurement*, 6(4), 427–438. <https://doi.org/10.1177/001316444600600401>

Hayashi, P., Abid, G., & Hoppen, N. (2019). Validity in qualitative research: a processual approach. *The Qualitative Report*, 24(1), 98–112. <https://doi.org/10.46743/2160-3715/2019.3443>

Hernández-Sampieri, R., & Mendoza, C. P. (2018). *Metodología de la Investigación. Las Rutas Cuantitativa, Cualitativa y Mixta*. McGraw Hill.



- Kane, M. T. (2013). The argument-based approach to validation. *School Psychology Review*, 42(4), 448–457. <https://doi.org/10.1080/02796015.2013.12087465>
- Kane, M. T. (2016). Explicating validity. *Assessment in Education. Principles, Policy & Practice*, 23(2), 198–211. <https://doi.org/10.1080/0969594X.2015.1060192>
- Kleining, G., & Witt, H. (2000). The qualitative heuristic approach: a methodology for discovery in psychology and the social sciences. Rediscovering the method of introspection as an example. *Forum Qualitative Sozialforschung*, 1(1), Art. 13. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0001136>
- Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting & Social Change*, 73(5), 467–482. <https://doi.org/10.1016/j.techfore.2005.09.002>
- McNeill, J. (2019). Social validity and teachers' use of evidence-based practices for autism. *Journal of Autism and Developmental Disorders*, 49, 4585–4594. <https://doi.org/10.1007/s10803-019-04190-y>
- Onwuegbuzie, A. J., & Leech, N. L. (2007). Validity and qualitative research: an oxymoron? *Quality & Quantity. International Journal of Methodology*, 41(2), 233–249. <https://doi.org/10.1007/s11135-006-9000-3>
- Pua, D. J., Peyton, D. J., Brownell, M. T., Contesse, V. A., & Jones, N. D. (2021). Preservice observation in special education: a validation study. *Journal of Learning Disabilities*, 54(1), 6-19. <https://doi.org/10.1177/0022219420920382>
- Rumrill, P. D., Cook, B. G., & Stevenson, N. A. (2020). *Research in Special Education. Designs, Methods, and Applications* (3rd ed.). Charles C. Thomas Pub, Ltd.



Sireci, S. G. (2009). Packing and unpacking sources of validity evidence: history repeats itself again. In: R. W. Lissitz (Ed.), *The Concept of Validity: Revisions, New Directions, and Applications*. Information Age Publishing, Inc.

Varanda, J. B., & Freitas, D. (2018). Método DELPHI: caracterização e potencialidades na pesquisa em Educação. *Pro-Posições*, 29(2), 389–415.

<https://dx.doi.org/10.1590/1980-6248-2015-0140>

Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research*, 11(4), 522–537.

<https://dx.doi.org/10.1177/104973201129119299>

