



INFLUENCIAS HISTÓRICAS DE LA TEORÍA DE LA VALIDEZ EN LA INVESTIGACIÓN DE DIAGNÓSTICOS DE ENFERMERÍA

INFLUÊNCIAS HISTÓRICAS DA TEORIA DA VALIDADE SOBRE A PESQUISA EM DIAGNÓSTICOS DE ENFERMAGEM

HISTORICAL INFLUENCES OF VALIDITY THEORY ON NURSING DIAGNOSIS RESEARCH

Marcos Venícios de Oliveira Lopes¹
Viviane Martins da Silva¹

ORCID: 0000-0001-5867-8023
ORCID: 0000-0002-8033-8831

¹ Universidade Federal do Ceará, Departamento de Enfermagem.
Fortaleza, CE, Brasil

Cómo citar: Lopes MVO, Silva VM Historical influences of validity theory on nursing diagnosis research. Online Braz J Nurs. 2026;25(1):e20266851. <http://doi.org/10.17665/1676-4285.20266851>

RESUMO

Objetivo: Explorar a teoria da validade por meio de uma abordagem histórica e analisar sua influência na pesquisa e no desenvolvimento de diagnósticos de enfermagem. **Método:** Estudo teórico-reflexivo baseado em pesquisa bibliográfica, conduzido para compreender a evolução semântica, epistemológica e metodológica da teoria da validade. **Resultados:** O desenvolvimento histórico da teoria da validade organiza-se em cinco fases principais: primórdios na China antiga, gênese do conceito, fragmentação teórica, reunificação e desconstrução. A análise revelou como essa teoria tem sido fundamental na construção de estruturas diagnósticas e modelos de pesquisa, sendo, ao mesmo tempo, objeto de críticas e ampliações. A teoria da validade evoluiu sob influências de diversas correntes epistemológicas, incluindo as abordagens trinitária, unificada e baseada em argumentos. **Conclusão:** Este estudo oferece uma visão abrangente e crítica da teoria da validade, destacando suas implicações para o desenvolvimento e o uso de diagnósticos de enfermagem na teoria e na prática.

Descritores: Diagnóstico de Enfermagem; Validade dos Testes; Estudo de Validação; Avaliação em Enfermagem.

ABSTRACT

Objective: To explore validity theory through a historical approach and analyze its influence on research and the development of nursing diagnoses. **Method:** A theoretical-reflective study based on bibliographic research was conducted to understand the semantic, epistemological, and methodological evolution of validity theory. **Results:** The historical development of validity theory is organized into five main phases: beginnings in ancient China, the genesis of the concept, theoretical fragmentation, reunification, and deconstruction. The analysis revealed how this theory has been fundamental in constructing diagnostic structures and research models, while simultaneously being an object of criticism and expansion. Validity theory evolved under the influence of various epistemological currents, including the trinitarian, unified, and argument-based approaches. **Conclusion:** This study offers a comprehensive and critical view of validity theory, highlighting its implications for the development and use of nursing diagnoses in both theory and practice.

Descriptors: Nursing Diagnosis; Test Validity; Validation Study; Nursing Assessment.

RESUMEN

Objetivo: Explorar la teoría de la validez a través de un enfoque histórico y analizar su influencia en la investigación y el desarrollo de los diagnósticos de enfermería. **Método:** Se realizó un estudio teórico-reflexivo basado en una investigación bibliográfica para comprender la evolución semántica, epistemológica y metodológica de la teoría de la validez. **Resultados:** El desarrollo histórico de la teoría de la validez se organiza en cinco fases principales: inicios en la antigua China, la génesis del concepto, la fragmentación teórica, la reunificación y la desconstrucción. El análisis reveló cómo esta teoría ha sido fundamental en la construcción de estructuras diagnósticas y modelos de investigación, siendo simultáneamente objeto de crítica y expansión. La teoría de la validez evolucionó bajo la influencia de diversas corrientes epistemológicas, incluyendo los enfoques trinitario, unificado y basado en argumentos. **Conclusión:** Este estudio ofrece una visión integral y crítica de la teoría de la validez, destacando sus implicaciones para el desarrollo y uso de los diagnósticos de enfermería tanto en la teoría como en la práctica.

Descritores: Diagnóstico de Enfermería; Validez de las Pruebas; Estudio de Validación; Evaluación en Enfermería.

Editores:

Rosimere Ferreira Santana (ORCID: 0000-0002-4593-3715)
Geilsa Soraia Cavalcanti Valente (ORCID: 0000-0003-4488-4912)
Nuno Felix (ORCID: 0000-0002-0102-3023)

Editora:

Escola de Enfermagem Aurora de Afonso Costa – UFF
Rua Dr. Celestino, 74 – Centro, CEP: 24020-091 – Niterói, RJ, Brasil
Correo electrónico de la revista: objn.cme@id.uff.br

Autora correspondiente:

Marcos Venícios de Oliveira Lopes
Correo electrónico: marcos@ufc.br

Lo que ya se sabe:

- La teoría de la validez tiene un desarrollo histórico complejo, marcado por diferentes corrientes epistemológicas y debates sobre su significado, métodos y aplicaciones.
- Los procesos de validación de los diagnósticos de enfermería se han basado históricamente en métodos clásicos, a menudo alejados de las discusiones contemporáneas sobre la validez.
- Aún existen vacíos importantes, especialmente en lo que respecta a la comprensión de las implicaciones teóricas y prácticas de la validez diagnóstica.

Lo que este artículo aporta:

- Ofrece un análisis histórico exhaustivo de la teoría de la validez, evidenciando cómo las diferentes fases influyeron en el desarrollo de los diagnósticos de enfermería.
- Demuestra que, aunque con cierto retraso en comparación con la psicometría, la enfermería ha incorporado progresivamente métodos de validación avanzados, como los modelos de causalidad, la Teoría de Respuesta al Ítem (TRI) y las clases latentes.
- Destaca la necesidad de profundizar en la discusión sobre el significado y las consecuencias de la interpretación de los resultados diagnósticos, señalando una brecha crítica que aún no ha sido abordada en el campo.

INTRODUCCIÓN

El concepto de validez constituye el fundamento central para la investigación y la práctica en diversas disciplinas, funcionando como una característica esencial de la adecuación de las inferencias extraídas de los datos. En áreas como la psicología, la educación, las ciencias de la salud y las ciencias sociales, la integridad de los resultados y sus aplicaciones dependen de la solidez de las evaluaciones⁽¹⁾.

En enfermería, el interés por los procesos de validación ha aumentado en las últimas décadas, especialmente en el desarrollo de instrumentos destinados a evaluar el estado de salud del paciente. Estos procesos se han basado en gran medida en la metodología psicométrica y en los estándares establecidos por entidades como la *American Psychological Association* (APA), la *American Educational Research Association* (AERA) y el *National Council on Measurement in Education* (NCME). Aunque influyentes, estas recomendaciones han sido objeto de críticas, alimentando un debate de larga duración sobre el concepto de validez⁽²⁻⁴⁾.

Este debate puede resumirse en cuatro problemas esenciales: ontológico (el significado de la validez), epistemológico (las características de una prueba válida), metodológico (cómo investigar la validez) y ético (cómo aplicar la prueba). Entre ellos, la cuestión ontológica —la más fundamental— es la que ha recibido menor atención en la literatura especializada⁽⁵⁾. No obstante, los problemas restantes persisten sin consenso, reflejando diferentes posiciones teórico-metodológicas que han influido en la evolución del concepto.

El término “validez” comenzó a integrarse sistemáticamente en la investigación de enfermería a medida que la evaluación de la salud ganó reconocimiento como un componente esencial del cuidado. Aunque el uso de los diagnósticos de enfermería aún genera divergencia entre los profesionales, se ha vuelto imperativo desarrollar instrumentos rigurosos e identificar indicadores que guíen las intervenciones basadas en resultados. En este contexto, los procesos de validación basados en la psicometría encontraron terreno fértil en la enfermería, impulsados por el desarrollo de lenguajes estandarizados que exigen métodos capaces de legitimar componentes de clasificaciones y taxonomías.

Mientras tanto, en los campos de la psicología y la educación, las discusiones sobre la teoría de la validez avanzaron y dieron lugar a diferentes corrientes de pensamiento^(2,6-7). Sin embargo, una gran parte de la investigación en enfermería permanece distanciada de estas discusiones, empleando definiciones, métodos y análisis que reflejan enfoques obsoletos. En casos extremos, este desajuste puede conducir a interpretaciones erróneas de los resultados y com-

prometer las respuestas a los cuatro problemas descritos anteriormente⁽⁵⁾.

En el campo epistemológico, la validez diagnóstica se ha vinculado tradicionalmente al grado en que las características definitorias, cuando se observan en un número suficiente de casos, reflejan la realidad de la interacción cliente-entorno⁽⁸⁻⁹⁾. Esta perspectiva sigue siendo dominante y concentra los esfuerzos en identificar las características definitorias como evidencia del contenido diagnóstico. Sin embargo, pocos estudios abordan las relaciones clínicas o la causalidad, elementos centrales en las corrientes contemporáneas que definen la validez como la relación causal entre el atributo y la puntuación de la prueba⁽¹⁰⁾.

En el campo metodológico, ha habido una gran evolución desde que se propusieron inicialmente los modelos clásicos. Los enfoques recientes buscan ampliar la comprensión del proceso diagnóstico, utilizando pruebas diagnósticas, teorías de mediano alcance y métodos que incorporan análisis de causalidad⁽¹¹⁻¹⁴⁾. Los autores también han relacionado la validez diagnóstica con los conceptos de validez interna y externa, ampliando su operacionalización metodológica⁽¹⁵⁾.

En el campo ético, aunque la idea de que los diagnósticos válidos deben resistir el escrutinio de los enfermeros en la práctica clínica es reiterada por diversos autores^(13,16-20), la discusión se ha restringido a la necesidad de alineación entre la evidencia científica y la práctica profesional. Sin embargo, la implementación práctica aún depende excesivamente de la opinión de expertos y carece de debates más amplios que comparen enfoques diseñados para cerrar la brecha entre la teoría y la práctica, como la investigación traslacional y las teorías de mediano alcance.

Definir la validez aplicada a los diagnósticos de enfermería requiere, por tanto, la comprensión de su evolución teórica y el impacto de esta trayectoria en el razonamiento clínico. Sin embargo, esta discusión parece haber sido abreviada en el campo, posiblemente debido a la producción acelerada de conocimiento mediada por tecnologías digitales, la búsqueda de respuestas simplificadas a problemas complejos y la competencia entre diferentes enfoques. Como consecuencia, el concepto de validez se ha empleado frecuentemente de forma inadecuada, confundiendo con eficacia, adecuación o evaluación⁽²¹⁾.

Basándose en estas consideraciones, este artículo pretende explorar la evolución de la teoría de la validez, discutir su influencia en el desarrollo de los diagnósticos de enfermería y ofrecer perspectivas sobre cómo su contexto histórico continúa moldeando los métodos de investigación sobre el tema.

DESARROLLO HISTÓRICO DE LA TEORÍA DE LA VALIDEZ

Inicios de la teoría de la validez

El interés por pruebas supuestamente válidas para identificar características esenciales de los sujetos tiene raíces en la antigua China, al menos 2.200 años antes de Cristo, cuando se aplicaban exámenes a los funcionarios del imperio cada tres años para verificar su aptitud funcional. Estos exámenes eran necesarios porque los cargos gubernamentales no eran hereditarios, lo que requería una selección y evaluación constantes. Este sistema se perfeccionó con el tiempo; en el año 115 a.C., por ejemplo, los candidatos eran evaluados en seis “artes”: música, tiro con arco, equitación, escritura, aritmética y los ritos y ceremonias de la vida pública y privada⁽²²⁾.

Los exámenes escritos se introdujeron en la dinastía Han (entre 200 a.C. y 200 d.C.) y alcanzaron su forma final alrededor de la década de 1370. En ese momento, los exámenes constaban de tres etapas, con tasas de aprobación muy bajas y condiciones de aislamiento rigurosas. Solo cerca del 3% de los seleccionados finales se convertían en mandarines elegibles para cargos públicos. Tales exámenes, aunque rigurosos, carecían de validez procesal e incluían prácticas excesivamente severas, lo que llevó a su abolición en 1906 por decreto real⁽²²⁻²³⁾.

A pesar de la falta de una relación directa con la idea de validez del diagnóstico de enfermería durante este período, el sistema chino influyó en las prácticas europeas. A partir del siglo XVI, liberales de la Revolución Francesa, como Voltaire y Quesnay, abogaron por prácticas similares. A principios del siglo XIX, diplomáticos británicos también sugirieron exámenes escritos para el servicio civil, con la primera implementación ocurriendo en 1833 para seleccionar internos para el servicio civil indio⁽²²⁾.

Los historiadores consideran que el desarrollo moderno de la teoría de la validez se originó a principios del siglo XIX en Europa y Estados Unidos. Una organización útil de esta trayectoria divide la evolución en cuatro períodos: 1) la génesis de la teoría de la validez (mediados de 1800 a 1951), 2) la fragmentación de la validez (1952 a 1974), 3) la reunificación de la validez (1975 a 1999) y 4) la deconstrucción de la validez (2000 a 2012)⁽²⁴⁾.

Aunque el texto carece de un análisis crítico de esta historia, se describen detalladamente los principales hechos que marcaron el desarrollo de la teoría de la validez⁽²⁵⁾. En la Figura 1 se puede ver un resumen de esta evolución, adaptado al contexto de los diagnósticos enfermeros.

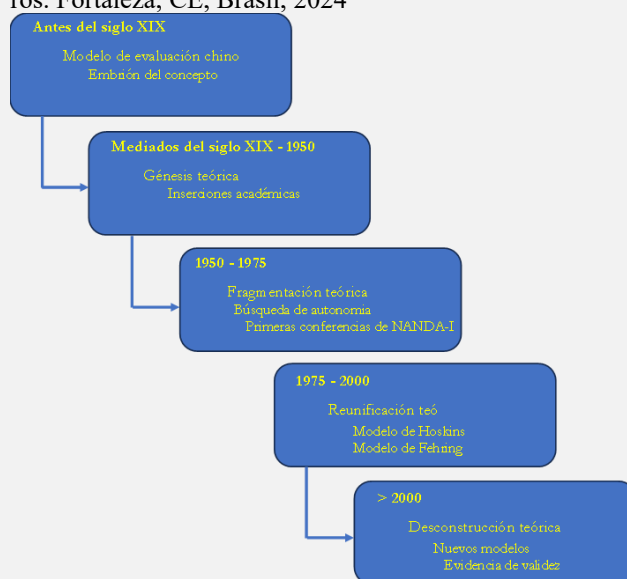
La génesis de la teoría de la validez

El primer período en la historia de la validez puede dividirse en dos momentos⁽²⁴⁾: un período gestacional, que se extiende desde mediados del siglo XIX hasta 1920, y un período de cristalización, entre 1921 y 1951. El período gestacional estuvo marcado por el creciente interés en Europa y Estados Unidos por desarrollar exámenes estructurados para la admisión escolar, basándose en la creencia de que las pruebas estructuradas aumentarían la precisión de la evaluación y producirían decisiones más justas.

En el Reino Unido, la Universidad de Londres fue fundada en 1836 para regular los exámenes de sus unidades, y en 1871 comenzó a aceptar resultados de otras institucio-

nes. Oxford y Cambridge adoptaron exámenes de ingreso en 1857 y 1858, lo que fomentó mejoras en la educación secundaria. En los Estados Unidos, los exámenes escritos se introdujeron en Massachusetts en 1845 para la certificación de la escuela secundaria, expandiéndose posteriormente a la selección profesional. En este contexto, se desarrollaron técnicas estadísticas aplicadas al estudio de la herencia⁽²⁶⁾, y se propuso un enfoque empírico de la validación, sugiriendo la comparación de medidas con estimaciones independientes de las capacidades humanas para identificar las más informativas⁽²⁷⁾. Los investigadores británicos ampliaron estas ideas desarrollando métodos de correlación aplicables a los atributos psicológicos⁽²⁸⁾. El coeficiente de correlación se convirtió entonces en la principal herramienta estadística en los estudios de validación.

Figura 1 - Resumen de la evolución histórica de la teoría de la validez aplicada al contexto de los diagnósticos enfermeros. Fortaleza, CE, Brasil, 2024



Fuente: elaboración propia, 2024.

Paralelamente a esto, los esfuerzos por comprender la mente humana llevaron a los investigadores franceses, entre 1905 y 1911, a desarrollar pruebas con niveles progresivos de dificultad para identificar el retraso mental en los niños⁽²²⁾. Estos instrumentos se difundieron rápidamente. Las pruebas de inteligencia también se aplicaron a más de 1,5 millones de reclutas estadounidenses durante la Primera Guerra Mundial. El entusiasmo por las pruebas psicológicas fue tan significativo que se afirmó que la mayoría de los países “civilizados” ya poseían instituciones de orientación vocacional con psicólogos capacitados⁽²⁹⁾.

Sin embargo, a principios del siglo XX, los académicos estadounidenses empezaron a cuestionar si los exámenes estructurados eran realmente tan precisos como se creía, destacando la subjetividad inherente a los juicios. En respuesta, se intensificó el desarrollo de instrumentos con menor subjetividad, recurriendo a ítems de verdadero/falso, opción múltiple, recuerdo y completación⁽³⁰⁻³¹⁾. Este cambio fomentó la búsqueda de métodos objetivos y explica por qué la teoría de la validez surgió inicialmente en los Estados Unidos⁽²⁴⁾. El período también vio la proliferación de pruebas estandarizadas en la práctica profesional y en la investigación científica sobre la inteligencia y la personalidad. Durante esta fase, el coeficiente de correlación se consolidó como elemento cen-

tral para evaluar la calidad de las pruebas.

En la transición al periodo de cristalización, la Asociación Nacional de Directores de Investigación Educativa de Norteamérica buscó establecer un consenso sobre términos y procedimientos. Es en este momento cuando surge la definición inaugural de validez: el grado en que una prueba mide lo que se supone que debe medir⁽³²⁾. Pocos años después, esta redacción se modificó ligeramente, consolidando la llamada definición clásica: el grado en que una prueba mide lo que pretende medir⁽³³⁾. Esta definición guio el pensamiento psicométrico durante décadas e influyó en las formulaciones posteriores adoptadas en el campo de los diagnósticos de enfermería.

A partir de esta definición, surgió la necesidad de determinar cómo verificar empíricamente la validez. Se propusieron dos enfoques iniciales: el análisis lógico del contenido y la evidencia empírica por correlación, siendo esta última considerada la más robusta en aquel momento⁽³⁴⁾. El reto central consistía en definir criterios adecuados de comparación⁽³⁵⁻³⁷⁾. Las recomendaciones incluían el juicio de expertos, la comparación entre versiones reducidas y completas de una prueba y el análisis longitudinal. Sin embargo, la validez de los propios criterios seguía siendo difícil de demostrar⁽³⁸⁾.

Simultáneamente, los estudiosos de la educación cuestionaron la centralidad exclusiva de la estadística. Sostenían que una prueba podía considerarse válida si los expertos, mediante un análisis lógico, juzgaban que su contenido era coherente con el plan de estudios⁽³⁸⁾. Se propuso entonces que el juicio de los expertos complementara los enfoques empíricos⁽³⁶⁾. La división entre la validez basada en criterios y la validez basada en el contenido se intensificó. Aunque la literatura destaca el predominio del enfoque empírico, el periodo también valoró el contenido, incluyendo el propósito del instrumento y la llamada “validez de apariencia”⁽³⁹⁾.

A pesar de la atención prestada al contenido, el periodo anterior a 1950 se describe como una fase orientada a la “predicción”⁽⁴⁰⁻⁴¹⁾. El énfasis en los métodos de criterio empírico condujo a la degradación gradual del concepto de validez, que empezó a confundirse con los “procedimientos de validación”. Mientras que estos se refieren a los métodos de obtención de datos, la validez corresponde a la sostenibilidad de las interpretaciones producidas. En consecuencia, las obras de instrucción redujeron la validez a la idea de un coeficiente de correlación empírico⁽²⁴⁾.

En enfermería, la producción científica a finales del siglo XIX y principios del XX fue incipiente; no se encontraron artículos sobre validez o pruebas en las bases de datos PubMed y Scopus durante este periodo. Las primeras referencias aparecen en la década de 1940 en estudios realizados por psicólogos sobre pruebas de personalidad aplicadas a enfermeras⁽⁴²⁻⁴³⁾. El primer estudio publicado por enfermeras directamente relacionado con la teoría de la validez data de finales de la década de 1950, describiendo un enfoque empírico para identificar las características del éxito profesional entre los estudiantes⁽⁴⁴⁾. Así, sólo en periodos posteriores la teoría de la validez se integró en los temas de la profesión. En esta etapa, la idea de los diagnósticos de enfermería era embrionaria y las discusiones sobre su validación eran inexistentes, aunque el interés por las estrategias de evaluación clínica empezaba a surgir en el ámbito académico.

La fragmentación de la validez

A principios de la década de 1950, la APA, el NCME

y la AERA publicaron nuevos Estándares para las pruebas psicológicas. Bajo la influencia de Lee Cronbach, el documento sistematizó la tendencia a clasificar la validez en tipos distintos, basándose en la distinción entre validez lógica y empírica⁽⁴⁵⁾. En 1952, la APA propuso cuatro tipos: de contenido, predictiva, de estado y congruente⁽⁴⁶⁾. La versión final de 1954 mantuvo los cuatro tipos, pero los renombró: el estado se convirtió en validez concurrente, y la congruente en validez de constructo⁽⁴⁷⁾. Esta última se aplicaba cuando los análisis lógicos o empíricos directos eran insuficientes.

En un artículo fundamental, se destacó que la validez de constructo es esencialmente científica y requiere la definición del atributo teórico responsable del desempeño⁽⁴⁸⁾. La contribución más significativa fue la defensa de que la validez se relaciona con las interpretaciones de las puntuaciones, las cuales deben ser coherentes con una red nomológica (un sistema de leyes que vincula términos teóricos con indicadores observables)⁽⁵⁾.

La revisión de los Estándares de 1966 unificó la validez concurrente y la predictiva bajo el término validez de criterio, dejando tres tipos reconocidos, estructura que se mantuvo en 1974⁽⁴⁹⁾. Entonces se consolidó la percepción (hoy considerada errónea) de que estos tipos eran concepciones alternativas de la validez, en lugar de categorías de evidencia. La validez de constructo adquirió una posición central, aunque los estudios de criterio continuaron proliferando.

A partir de 1960, los estudios de validación realizados por enfermeras se multiplicaron, centrándose en la atención pediátrica⁽⁵⁰⁻⁵²⁾, el desempeño profesional⁽⁵³⁻⁵⁸⁾, el conocimiento técnico⁽⁵⁹⁾ y la satisfacción laboral⁽⁶⁰⁾. El interés se centró en las aptitudes y competencias, en consonancia con la tradición de la validez de criterio. Dado que el primer grupo de trabajo para la clasificación de los diagnósticos de enfermería no surgiría hasta la década de 1970, aún no existían procesos de validación dirigidos a este fin. El concepto fragmentado de validez no sería incorporado por la enfermería sino hasta finales de la década de 1970 y principios de la de 1980.

La reunificación de la validez

A mediados de la década de 1970, el interés de los estudiosos de la teoría de la validez se desplazó hacia el proceso de validación de constructo. Este movimiento defendió la superioridad de la validez de constructo sobre otros tipos⁽⁴⁸⁾ y enfatizó que no se valida una prueba, sino más bien las interpretaciones derivadas de los datos producidos por procedimientos específicos, trasladando el enfoque de la validez del instrumento a las inferencias asociadas con el proceso de medición⁽⁶¹⁾. Estas concepciones prepararon el terreno para el desarrollo de un enfoque unificado de la teoría de la validez.

El inicio formal de este cambio ocurrió en 1974, cuando la APA, la AERA y el NCME presentaron una nueva revisión de los Standards. A partir de este hito, dos corrientes comenzaron a predominar: la concepción trinitaria, según la cual los tres tipos de validez eran igualmente importantes, y la concepción unitaria, centrada exclusivamente en la validez de constructo⁽³⁸⁾. Samuel Messick desempeñó un papel fundamental en este debate al expandir las ideas de otros académicos como Gulliksen y Loevinger. Para él, toda la validez debía entenderse como validez de constructo, y su contribución más significativa fue reposicionar la medición como un elemento central en todos los contextos de validez, desa-

fiando la visión anterior que separaba la validez para la medición y la validez para la predicción⁽⁶²⁾.

Los defensores del enfoque unitario argumentaron que la validez debería referirse únicamente a las inferencias derivadas de la prueba, rechazando la validez de contenido como un tipo independiente, ya que se refería solo al universo de los ítems⁽³⁸⁾. También defendieron que la evidencia de la validez de criterio era insuficiente, ya que dependía de criterios externos que igualmente necesitarían ser validados. Por lo tanto, la validez de criterio no representaría un tipo de validez, sino solo una de las muchas fuentes de evidencia para el proceso de validación de constructo⁽⁶²⁾. Además, se argumentó que las diferentes formas de evidencia no significaban diferentes tipos de validez⁽⁶³⁻⁶⁴⁾.

La validez fue entonces redefinida como un juicio evaluativo integrado sobre el grado en que la evidencia empírica y los fundamentos teóricos respaldan la adecuación y relevancia de las inferencias y acciones derivadas de los resultados de las pruebas⁽⁶⁵⁾. También se destaca que un elemento central de esta definición fue la introducción de las consecuencias sociales de las interpretaciones⁽⁶⁶⁾. Así, el enfoque unitario incorporó la relevancia, la utilidad, los valores y los impactos sociales —elementos ya mencionados anteriormente pero nunca integrados en el concepto de validez. Paradójicamente, el intento de simplificación produjo un modelo más complejo y un blanco de críticas.

Las críticas al modelo unitario se concentraron en la dificultad de operacionalizarlo en el campo aplicado. Muchos autores argumentaron que expandir el concepto para incluir las consecuencias sociales conducía a la desorganización conceptual y generaba más confusión que claridad, lo que llevó a la recomendación de restringir el concepto de validez y dio inicio a una deconstrucción del enfoque unitario⁽⁶⁷⁻⁶⁸⁾.

En enfermería, el periodo entre 1970 y 1990 mostró un crecimiento gradual del interés por la validez. Aunque se observó un aumento moderado de las publicaciones entre mediados de los años 70 y finales de los 80, la expansión más significativa ocurrió en la década de 1990. Los primeros seis artículos sobre la validación de diagnósticos de enfermería se publicaron en 1985, casi todos en el volumen 20⁽⁴⁾ de *Nursing Clinics of North America*⁽⁶⁹⁻⁷⁴⁾. La excepción fue un artículo publicado en *Orthopaedic Nursing* sobre la validación del diagnóstico “Imagen corporal alterada”⁽⁷⁵⁾.

En la década de 1970, la principal preocupación de la comunidad dedicada a los diagnósticos de enfermería era definir el concepto mismo y establecer una estructura taxonómica inicial. La primera Conferencia Nacional sobre Clasificación de Diagnósticos de Enfermería tuvo lugar en 1973 en la Universidad de Saint Louis y se basó en enfoques inductivos y deductivos para desarrollar diagnósticos⁽⁷⁶⁾. Aunque el enfoque inductivo incluía la validación clínica, el significado de esta etapa no fue detallado. Parte del enfoque estuvo influenciado por la teoría de conjuntos, los diagramas de Venn y el álgebra de Boole para clasificar entidades clínicas⁽⁷⁷⁾. En esta conferencia, se listaron 30 diagnósticos primarios y 100 categorías diagnósticas, diferenciadas por etiología y duración; sin embargo, las características definitorias se describieron solo para los diagnósticos principales y 14 categorías, lo que refleja la ausencia de intención de crear una taxonomía estructurada, sino únicamente de identificar posibles problemas clínicos⁽⁷⁶⁾.

Fue solo en 1979 cuando surgió el primer modelo metodológico de validación para los diagnósticos de enferme-

ría⁽⁷⁸⁾, centrado en la identificación de características definitorias y evidencia que respaldara la existencia del diagnóstico. Sin embargo, el modelo reflejaba una concepción de la validez alineada con las fases iniciales de la teoría psicométrica —análisis lógico y evidencia empírica— e ignoraba los avances de la teoría unificada de la validez. También existía una confusión conceptual, como la asociación de la validez con el grado de acuerdo entre enfermeras.

En 1982, se fundó la *North American Nursing Diagnosis Association* (NANDA), que comenzó a colaborar con la *American Nursing Association* (ANA). Un intento de incluir los diagnósticos de la NANDA en la Clasificación Internacional de Enfermedades (CIE) en 1986 fue rechazado por falta de representatividad internacional. La Taxonomía I se publicó en 1987 y, al año siguiente, comenzaron los esfuerzos para ampliar la participación internacional⁽⁷⁹⁾.

Durante la década de 1980, los estudios sobre la validación de los diagnósticos de enfermería se intensificaron. La validez diagnóstica se definió como dependiente de la ocurrencia simultánea de claves clínicas y de la congruencia entre claves y diagnósticos en muestras grandes, y se criticó el empirismo de la NANDA argumentando que las validaciones basadas únicamente en la opinión de expertos eran insuficientes⁽¹⁹⁾. Esta crítica se refería al periodo de cristalización de la teoría de la validez, marcado por la tensión entre los análisis lógicos y empíricos. Se propusieron dos fuentes de evidencia para la validación: de expertos y clínica⁽¹³⁾. Esto era esencialmente un análisis lógico basado en la opinión de expertos y un análisis empírico similar a la validez concurrente, utilizando un promedio ponderado de las evaluaciones diagnósticas como criterio.

En 1989, se presentó un modelo de tres fases: análisis de concepto, validación por expertos y validación clínica, reflejando las discusiones de la teoría psicométrica al final de la fragmentación⁽⁸⁰⁾. A mediados de la década de 1990, las críticas a los modelos anteriores sostuvieron que una de las principales debilidades era la falta de una conceptualización teórica previa adecuada⁽⁸⁰⁻⁸¹⁾. Este movimiento acercó a la enfermería al debate sobre la reunificación de la teoría de la validez, aun cuando en la psicometría, la validez de constructo estaba ganando centralidad y el concepto se estaba expandiendo para incluir dimensiones éticas y sociales. Así, este periodo marca los primeros movimientos efectivos en la búsqueda de métodos para guiar los procesos de construcción y validación de los diagnósticos de enfermería. Sin embargo, estos movimientos parecieron incorporar no solo las ideas actuales de ese periodo, sino también estructuras y fases típicas de periodos anteriores, produciendo una mezcla de distintas interpretaciones de lo que sería la validez de un diagnóstico de enfermería.

La deconstrucción de la validez

Entre la década de 1990 y principios de la de 2000, se desarrolló una nueva metodología para proporcionar una base a la práctica de la validación basada en la argumentación. A diferencia de la corriente que concebía la validez como un juicio evaluativo sustentado por la evidencia, esta nueva propuesta describía cómo dicha evidencia debería organizarse en un argumento de validez coherente^(40,82-83). El enfoque proponía que el proceso de validación debería comenzar con la definición de la interpretación y el uso previsto de las puntuaciones de las pruebas; continuar con la explicación de las afirmaciones que sustentarían esta interpreta-

ción en forma de un argumento estructurado y mediante la comprobación de sus supuestos; y finalizar con el juicio del argumento en términos de coherencia, integridad y plausibilidad⁽⁶⁶⁾.

A lo largo de la década de 2000, este enfoque se profundizó y se apoyó en la noción de red nomológica de Cronbach y Meehl, pero se distanció de la concepción unificada de Messick. Su principal crítica a las propuestas de Cronbach y Messick fue la llamada “búsqueda interminable de validez”, ya que esos autores vinculaban la validez a una investigación científica continua sobre el significado de las puntuaciones, dependiente de teorías que relacionan un constructo con una red extensa de otros constructos⁽²⁴⁾. En oposición, el nuevo enfoque argumentaba que la validación dependería exclusivamente de la interpretación y el uso de los resultados que el usuario de la prueba tiene en mente⁽⁴²⁾. Así, para las pruebas dirigidas a atributos relativamente simples, solo sería suficiente un conjunto limitado de evidencias y análisis.

Para sostener esta perspectiva, se estableció una distinción entre atributos observables y constructos teóricos⁽⁸³⁾. Las interpretaciones basadas en constructos requerirían una investigación científica amplia, típica de la validación de constructo tradicional, mientras que las interpretaciones asociadas a atributos observables, como la competencia matemática o el conocimiento de vocabulario, serían más manejables desde un punto de vista metodológico. De esta manera, la nueva metodología descompuso y simplificó el concepto de validez y redujo el lugar central de los constructos teóricos.

Otros autores también criticaron el modelo unificado. Por ejemplo, se rechazó la inclusión de aspectos éticos en el concepto de validación, considerando que los análisis científicos y los análisis éticos deberían permanecer en esferas distintas, pues de lo contrario se acentuaría la disociación entre la teoría y la práctica de la validación⁽⁸⁴⁾. Otra crítica aún más radical a la concepción de validez adoptada durante décadas por la APA, AERA y NCME argumentaba que la validez debería concebirse únicamente como una propiedad de las pruebas, volviendo a la definición clásica⁽³⁾.

Desde principios de la década de 2000 hasta el momento en que se escribió este texto, se observó un crecimiento significativo en el número de publicaciones sobre la teoría de la validez y los procesos de validación. Entre los psicometras persisten debates intensos, con defensores de las corrientes trinitaria, unitaria, argumentativa o propuestas alternativas. En 2012, la revista *Measurement* dedicó un número especial a discutir una propuesta para aclarar la definición consensuada de validez presentada en 1999 por AERA, APA y NCME⁽⁸⁵⁾. La participación de diversos especialistas reveló una amplia gama de posiciones favorables y contrarias, destacando la dificultad de alcanzar un consenso sobre el concepto de validez.

En enfermería, este periodo reciente sigue un movimiento similar, con un aumento progresivo del interés por el tema. Entre 2000 y 2023, una búsqueda en la base de datos Scopus con la estrategia (“Validity” OR “Validation”) AND “Nursing” identificó 7.574 publicaciones en el área, con una tendencia claramente ascendente, alcanzando los 574 artículos en 2023. De estos, 378 incluyeron el término diagnóstico de enfermería, también con una tendencia ascendente, llegando a 31 artículos sobre la validación de diagnósticos de enfermería en 2021. El aumento más pronunciado de estos estudios se produjo a partir de 2015, cuando se empezaron a

publicar más de 20 artículos anuales.

En este contexto, se propusieron nuevos enfoques para la validación de los diagnósticos de enfermería. En 1989, la NANDA promovió una reunión con especialistas en métodos de investigación y diagnósticos de enfermería para discutir el uso de metodologías cuantitativas y cualitativas en el proceso de validación⁽⁸⁶⁾. Entre los trabajos presentados, solo seis artículos definieron explícitamente el término validez; tres utilizaron la definición clásica de la década de 1920, a pesar de ser citada a partir de referencias contemporáneas⁽⁸⁷⁻⁹⁰⁾. Otros dos artículos se apoyaron en modelos antiguos y en una concepción de la validez alineada con el periodo de fragmentación⁽⁹¹⁻⁹²⁾. Esto sugiere que el concepto de validez aplicado a los diagnósticos de enfermería se estaba desarrollando en una etapa anterior a la teoría de la validez discutida en psicometría.

Se propusieron diferentes métodos cualitativos y cuantitativos supuestamente aplicables a la validación de los diagnósticos de enfermería que incluyeron la teoría fundamentada⁽⁹³⁻⁹⁴⁾, la etnografía y la fenomenología⁽⁹²⁾, el análisis de concepto⁽⁹⁵⁾, la triangulación⁽⁸⁸⁾, los métodos epidemiológicos⁽⁹⁶⁾ y las técnicas multivariantes clásicas^(87,89-90,97). A pesar de la amplia variedad metodológica señalada en esta reunión, los estudios de validación de diagnósticos de enfermería permanecieron, en gran medida, centrados en el enfoque de la década de 1990⁽¹³⁾.

Solo a partir de la década de 2000 los grupos latinoamericanos comenzaron a adoptar estrategias diferenciadas, como el uso de pruebas diagnósticas basadas en un panel de jueces⁽¹¹⁾, el modelado por la teoría de respuesta al ítem⁽⁹⁸⁾, el análisis de clases latentes⁽¹⁴⁾ y el desarrollo de teorías de mediano alcance⁽¹²⁾. Este periodo denota la búsqueda de una ontología y una epistemología para la teoría de la validez de los diagnósticos de enfermería. Este proceso parece estar en desarrollo y estar influenciado por las nuevas tecnologías de la información, incluyendo la inteligencia artificial y el desarrollo de sistemas de registros electrónicos de salud.

Esta evolución histórica infunde la necesidad de una mejora teórica y metodológica por parte de los investigadores del área con el objetivo de revisar y proponer estructuras diagnósticas que permitan una interpretación adecuada y coherente de las respuestas humanas expresadas por quienes necesitan cuidados de enfermería. En este sentido, se observa que, aunque temporalmente desfasados respecto a la evolución de la teoría de la validez psicométrica, los procesos de validación de los diagnósticos de enfermería se han desarrollado progresivamente, incorporando métodos más sofisticados. Sin embargo, persiste una brecha importante: una discusión profunda sobre el significado de los resultados producidos en estos procesos.

CONCLUSIÓN

A pesar de que la enfermería ha avanzado progresivamente en la incorporación de métodos más sofisticados, persiste un desajuste conceptual en relación con las discusiones contemporáneas sobre la validez, especialmente en lo que respecta a la interpretación de los resultados diagnósticos y sus consecuencias. Ante esto, se recomienda que los investigadores expliquen con mayor rigor los supuestos teóricos que guían sus estudios de validación, definan claramente las interpretaciones y usos previstos de los diagnósticos, y organicen la evidencia empírica de manera coherente con estos objetivos. En el ámbito de la enseñanza, es fundamental am-

pliar la enseñanza de los diagnósticos de enfermería más allá de la reproducción de estructuras clasificatorias, incorporando la evolución histórica de los métodos y sus implicaciones para el razonamiento clínico y la toma de decisiones.

CONFLICTO DE INTERESES

Los autores declaran no tener conflictos de intereses.

REFERENCIAS

1. Fabrigar LR, Wegener DT, Petty RE. A Validity-Based Framework for Understanding Replication in Psychology. *Pers Soc Psychol Rev.* 2020;24(4):316-344. <https://doi.org/10.1177/1088868320931366>. PMID: 32715894.
2. Borsboom D, Markus KA. Truth and Evidence in Validity Theory. *J Educ Meas.* 2013;50(1):110-114. <https://doi.org/10.1111/jedm.12006>.
3. Borsboom D, Mellenbergh GJ, van Heerden J. The Concept of Validity. *Psychol Rev.* 2004;111(4):1061-1071. <https://doi.org/10.1037/0033-295x.111.4.1061>. PMID: 15482073.
4. Sireci S, Faulkner-Bond M. Validity evidence based on test content. *Psicothema.* 2014;1(26):100-107. <https://doi.org/10.7334/psicothema2013.256>. PMID: 24444737.
5. Pasquali L. Validade dos testes psicológicos: será possível reencontrar o caminho?. *Psic., Teor. Pesq.* (Online). 2007;23(spe):99-107. <https://doi.org/10.1590/s0102-37722007000500019>.
6. Markus KA, Borsboom D. *Frontiers of test validity theory: measurement, causation, and meaning.* London: Routledge; 2013.
7. Pasquali L. *Psicometria: teoria dos testes na psicologia e na educação.* 4th ed. Petrópolis: Vozes; 2004.
8. Gordon M. *Nursing diagnosis: process and application.* 3rd ed. St Louis: Mosby; 1993.
9. Gordon M, Sweeney MA. Methodological Problems and Issues in Identifying and Standardizing Nursing Diagnoses. *ANS Adv Nurs Sci.* 1979;2(1):1-16. <https://doi.org/10.1097/00012272-197910000-00002>. PMID: 114094.
10. Borsboom D, Wijsen LD. Frankenstein's validity monster: the value of keeping politics and science separated. *Assess Educ.* 2016;23(2):281-283. <https://doi.org/10.1080/0969594x.2016.1141750>.
11. Lopes MV de O, Silva VM da, Araujo TL de. Methods for Establishing the Accuracy of Clinical Indicators in Predicting Nursing Diagnoses. *Int J Nurs Knowl.* 2012;23(3):134-139. <https://doi.org/10.1111/j.2047-3095.2012.01213.x>. PMID: 23043652.
12. Lopes MV de O, Silva VM da, Herdman TH. Causation and Validation of Nursing Diagnoses: A Middle Range Theory. *Int J Nurs Knowl.* 2017;28(1):53-59. <https://doi.org/10.1111/2047-3095.12104>. PMID: 26095430.
13. Fehring RJ. Methods to validate nursing diagnoses. *Heart Lung.* 1987;16(6 Pt 1):625-629. PMID: 3679856.
14. Lopes MV de O, Silva VM da. Métodos avançados de validação de diagnósticos de enfermagem. In: Herdman TH, Napoleão AA, Lopes CT, Silva VM da, editores. PRONANDA: Programa de atualização em Diagnósticos de Enfermagem. Ciclo 4, v. 3. Porto Alegre: Artmed Panamericana; 2016. p. 31-74.
15. Creason NS. Clinical Validation of Nursing Diagnoses. *Int J Nurs Terminol Classif.* 2004;15(4):123-132. <https://doi.org/10.1111/j.1744-618x.2004.tb00009.x>. PMID: 15712860.
16. Lopes MV de O, Silva VM da, Araujo TL. Métodos de pesquisa para validação clínica de conceitos diagnósticos. In: Herdman TH, editor. PRONANDA: Programa de atualização em Diagnósticos de Enfermagem. Conceitos básicos. Porto Alegre: Artmed Panamericana; 2013. p. 87-132.
17. Lunney M. Critical Thinking and Accuracy of Nurses' Diagnoses. *Int J Nurs Terminol Classif.* 2003;14(3):96-107. <https://doi.org/10.1111/j.1744-618x.2003.00096.x>. PMID: 14649031.
18. Mazalová L, Marečková J. Types of validity in the research of NANDA International components. *Profese online.* 2012;5(2):11-15. <https://doi.org/10.5507/pol.2012.011>.
19. Tanner CA, Hughes AM. Nursing diagnosis: issues in clinical practice research. *Top Clin Nurs.* 1984;5(4):30-38. PMID: 6558994
20. Whitley GG. Processes and Methodologies for Research Validation of Nursing Diagnoses. *Int J Nurs Terminol Classif.* 1999;10(1):5-14. <https://doi.org/10.1111/j.1744-618x.1999.tb00016.x>. PMID: 10358519.
21. Haig BD. Construct Validation and Clinical Assessment. *Behav Change.* 1999;16(1):64-73. <https://doi.org/10.1375/bech.16.1.64>.
22. DuBois PH. *A history of psychological testing.* Boston: Allyn and Bacon; 1970.
23. Gregory RJ. The history of psychological testing. In: Gregory RJ, editor. *Psychological testing: history, principles, and applications.* 6th ed. London: Pearson Education; 2010. p. 56-81.
24. Newton PE, Shaw SD. *Validity in educational & psychological assessment.* London: Sage; 2014.
25. Borsboom D. Zen and the art of validity theory. *Assess Educ.* 2016;23(3):415-421. <https://doi.org/10.1080/0969594x.2015.1073479>.
26. Galton F. *Hereditary genius.* London: Macmillan and Co.; 1869.
27. Cattell JM. V.—Mental Tests and Measurements. *Mind.* 1890;15(59):373-381. <https://doi.org/10.1093/mind/os-xv.59.373>.
28. Spearman C. "General Intelligence," Objectively Determined and Measured. *Am J Psychol.* 1904;15(2):201-292. <https://doi.org/10.2307/1412107>.
29. Burt CL. Historical sketch of the development of psychological tests. In: Board of Education, editor. *Report of the consultative committee on psychological tests of educable capacity and their possible use in the*

FINANCIACIÓN

El presente trabajo fue realizado con el apoyo del Consejo Nacional de Desarrollo Científico y Tecnológico. Procesos n.º 310378/2023-0 y 313495/2021-1.

- public system of education. London: His Majesty's Stationery Office; 1924. p. 1-61.
30. Meyer M. The Grading of Students. *Science*. 1908;28(712):243-250. <https://doi.org/10.1126/science.28.712.243>. PMID: 17820846.
 31. Ruch GM. The objective or new-type examination: an introduction to educational measurement. Chicago: Scott, Foresman and Company; 1929.
 32. Buckingham BR. Intelligence and its measurement: A symposium--XIV. *J Educ Psychol*. 1921;12(5):271-275. <https://doi.org/10.1037/h0066019>.
 33. Ruch GM. The improvement of the written examination. Chicago: Scott, Foresman; 1924.
 34. Beckstead JW. Content validity is naught. *Int J Nurs Stud*. 2009;46(9):1274-1283. <https://doi.org/10.1016/j.ijnurstu.2009.04.014>. PMID: 19486976.
 35. Bingham WV. Aptitudes and aptitude testing. New York: Harper & Bros; 1937.
 36. Kelley TL. Interpretation of educational measurements. New York: Macmillan; 1927.
 37. Thurstone LL. The reliability and validity of tests: derivation and interpretation of fundamental formulae concerned with reliability and validity of tests and illustrative problems. Ann Arbor, Mich.: Edwards Brothers; 1932.
 38. Sireci SG. The Construct of Content Validity. *Soc Indic Res*. 1998;45(1-3):83-117. <https://doi.org/10.1023/a:1006985528729>.
 39. Rulon PJ. On the validity of educational tests. *Harv Educ Rev*. 1946;16:290-296.
 40. Kane MT. Current Concerns in Validity Theory. *J Educ Meas*. 2001;38(4):319-342. <https://doi.org/10.1111/j.1745-3984.2001.tb01130.x>.
 41. Shepard LA. Chapter 9: Evaluating Test Validity. *Rev Res Educ*. 1993;19(1):405-450. <https://doi.org/10.3102/0091732x019001405>.
 42. Bennett GK, Gordon HP. Personality test scores and success in the field of nursing. *J Appl Psychol*. 1944;28(3):267-278. <https://doi.org/10.1037/h0056001>.
 43. Gordon HP. Study Habits Inventory scores and scholarship. *J Appl Psychol*. 1941;25(1):101-107. <https://doi.org/10.1037/h0058849>.
 44. Hawkins NG. A Measure of Empirical Validity. *Nurs Res*. 1959;8(1):13-17. <https://doi.org/10.1097/00006199-195908010-00004>.
 45. Cronbach LJ. Essentials of psychological testing. New York: Harper & Brothers; 1949.
 46. American Psychological Association. Committee on Test Standards. Technical recommendations for psychological tests and diagnostic techniques: preliminary proposal. *Am Psychol*. 1952;7(8):461-475. <https://doi.org/10.1037/h0056631>.
 47. American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques. *Psychol Bull*. 1954;51(2, Pt.2):1-38. <https://doi.org/10.1037/h0053479>. PMID: 13155713.
 48. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull*. 1955;52(4):281-302. <https://doi.org/10.1037/h0040957>. PMID: 13245896.
 49. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: APA; 1974.
 50. Chamorro IL, Davis ML, Green D, Kramer M. Development of an instrument to measure premature infant behavior and caretaker activities: time-sampling methodology. *Nurs Res*. 1973;22(4):300-309. <https://doi.org/10.1097/00006199-197307000-00004>. PMID: 4489647.
 51. Feldman H. Validity of a feedback system for evaluation of pediatric nursing care. *Nurs Res*. 1965;14(3):257-261.
 52. Rodgers B, Ferhott J, Cooper CL. A Screening Tool to Detect Psychosocial Adjustment of Children with Cystic Fibrosis. *Nurs Res*. 1974;23(5):420-425. <https://doi.org/10.1097/00006199-197409000-00009>.
 53. Bergman R, Edelstein A, Rotenberg A, Melamed Y. Psychological tests: their use and validity in selecting candidates for schools of nursing in Israel. *Int J Nurs Stud*. 1974;11(2):85-109. [https://doi.org/10.1016/0020-7489\(74\)90008-x](https://doi.org/10.1016/0020-7489(74)90008-x). PMID: 4496609.
 54. Johnson DM, Wilhite MJ. Reliability and Validity of Subjective Evaluation of Baccalaureate Program Nursing Students. *Nurs Res*. 1973;22(3):257-261. <https://doi.org/10.1097/00006199-197305000-00012>.
 55. Mowbray JK, Taylor RG. Validity of Interest Inventories for the Prediction of Success in a School of Nursing. *Nurs Res*. 1967;16(1):78-80. <https://doi.org/10.1097/00006199-196701610-00016>.
 56. Seither FG. A Predictive Validity Study of Screening Measures Used to Select Practical Nursing Students. *Nurs Res*. 1974;23(1):60-62. <https://doi.org/10.1097/00006199-197401000-00013>.
 57. Yu Chao YM. [Validity and reliability of scoring performed nursing procedures]. *Hu Li Za Zhi*. 1974;21(4):50-60. Chinese. PMID: 4498580.
 58. Zedeck S, Baker HT. Nursing performance as measured by behavioral expectation scales: A multitrait-multirater analysis. *Organ Behav Hum Perform*. 1972;7(3):457-466. [https://doi.org/10.1016/0030-5073\(72\)90029-3](https://doi.org/10.1016/0030-5073(72)90029-3).
 59. Carlson CE, Vernon DTA. Measurement of informativeness of hospital staff members. *Nurs Res*. 1973;22(3):198-204. <https://doi.org/10.1097/00006199-197305000-00002>.
 60. Munson FC, Heda SS. An Instrument for Measuring Nursing Satisfaction. *Nurs Res*. 1974;23(2):159-165. <https://doi.org/10.1097/00006199-197403000-00013>.
 61. Cronbach LJ. Test validation. In: Thorndike R, editor. Educational measurement. 2nd ed. Washington, DC: American Council on Education; 1971. p. 443-507.
 62. Messick S. The standard problem: Meaning and values in measurement and evaluation. *Am Psychol*. 1975;30(10):955-966. <https://doi.org/10.1037/0003-066x.30.10.955>.
 63. Messick S. Test validity and the ethics of assessment. *Am Psychol*. 1980;35(11):1012-1027. <https://doi.org/10.1037/0003-066x.35.11.1012>.
 64. Messick S. Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educ Res*. 1989;18(2):5-11. <https://doi.org/10.3102/0013189x018002005>.
 65. Messick SJ. Validity. In: Linn RL, editor. Educational measurement. 3rd ed. New York: American Council on Education/Macmillan; 1989. p. 13-103.
 66. Wolming S, Wikström C. The concept of validity in

- theory and practice. *Assess Educ.* 2010;17(2):117-132. <https://doi.org/10.1080/09695941003693856>.
67. Mehrens WA. The Consequences of Consequential Validity. *Educational Measurement.* 1997;16(2):16-18. <https://doi.org/10.1111/j.1745-3992.1997.tb00588.x>.
 68. Popham WJ. Consequential validity: Right Concern-Wrong Concept. *Educational Measurement.* 1997;16(2):9-13. <https://doi.org/10.1111/j.1745-3992.1997.tb00586.x>.
 69. McDonald BR. Validation of Three Respiratory Nursing Diagnoses. *Nurs Clin North Am.* 1985;20(4):697-710. [https://doi.org/10.1016/s0029-6465\(22\)01915-6](https://doi.org/10.1016/s0029-6465(22)01915-6).
 70. Munns DC. A Validation of the Defining Characteristics of the Nursing Diagnosis "Potential for Violence". *Nurs Clin North Am.* 1985;20(4):711-722. [https://doi.org/10.1016/s0029-6465\(22\)01916-8](https://doi.org/10.1016/s0029-6465(22)01916-8). PMID: 3852301.
 71. Pokorny BE. Validating a Diagnostic Label. *Nurs Clin North Am.* 1985;20(4):641-655. [https://doi.org/10.1016/s0029-6465\(22\)01911-9](https://doi.org/10.1016/s0029-6465(22)01911-9).
 72. Smith JK. Social reality as mind-dependent versus mind-independent and the interpretation of test validity. *Journal of Research & Development in Education.* 1985;19(1):1-9.
 73. Vincent KG. The Validation of a Nursing Diagnosis. *Nurs Clin North Am.* 1985;20(4):631-640. [https://doi.org/10.1016/s0029-6465\(22\)01910-7](https://doi.org/10.1016/s0029-6465(22)01910-7).
 74. Voith AM, Smith DA. Validation of the Nursing Diagnosis of Urinary Retention. *Nurs Clin North Am.* 1985;20(4):723-729. [https://doi.org/10.1016/s0029-6465\(22\)01917-x](https://doi.org/10.1016/s0029-6465(22)01917-x).
 75. Baird SE. Development of a Nursing Assessment Tool to Diagnose Altered Body Image in Immobilized Patients. *Orthop Nurs.* 1985;4(1):47-54. <https://doi.org/10.1097/00006416-198501000-00009>. PMID: 3844713.
 76. Gebbie KM, Lavin MA, editores. Classification of nursing diagnoses: proceedings of the first national conference. St. Louis: C.V. Mosby; 1975.
 77. Feinstein AR. Clinical judgment. Baltimore: Williams & Wilkins; 1967.
 78. Gordon M, Sweeney MA. Methodological Problems and Issues in Identifying and Standardizing Nursing Diagnoses. *ANS Adv Nurs Sci.* 1979;2(1):1-16. <https://doi.org/10.1097/00012272-197910000-00002>. PMID: 114094.
 79. Farias JN, Nóbrega MML, Pérez VLAB, Coler MS. Diagnóstico de Enfermagem: uma abordagem conceitual e prática. João Pessoa: Santa Marta; 1990.
 80. Hoskins LM. Clinical validation methodologies for nursing diagnosis research. In: Carroll-Johnson RM, editor. Classification of nursing diagnoses: proceedings of the eighth conference. Philadelphia: Lippincott; 1989. p. 126-131.
 81. Fehring RJ. The Fehring model. In: Carroll-Johnson RM, Paquette M, editores. Classification of nursing diagnoses: proceedings of the tenth conference. Philadelphia: Lippincott; 1994. p. 55-62.
 82. Kane MT. An argument-based approach to validity. *Psychol Bull.* 1992;112(3):527-535. <https://doi.org/10.1037/0033-2909.112.3.527>.
 83. Kane MT. Validating the interpretations and uses of test scores. In: Lissitz RW, editor. The concept of validity. Charlotte: Information Age Publishing; 2009. p. 39-64.
 84. Cizek GJ. Defining and distinguishing validity: Interpretations of score meaning and justifications of test use. *Psychol Methods.* 2012;17(1):31-43. <https://doi.org/10.1037/a0026975>. PMID: 22268761.
 85. Newton PE. Clarifying the Consensus Definition of Validity. *Measurement (Mahwah N J).* 2012;10(1-2):1-29. <https://doi.org/10.1080/15366367.2012.669666>.
 86. Chang BL, Kim MJ, Jones P, McFarlane E, editores. Proceedings of the Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989 Abril 28-30. Palm Springs, CA. St. Louis: NANDA; 1989.
 87. Chang BL. Reliability and construct validity. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 217-232.
 88. Creason NS. Clinical validation of Nursing diagnoses. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 278-304.
 89. Padilla GV. Tool development: reliability and related statistics. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 177-184.
 90. Schroeder MA. Tool development: validity related to nursing diagnoses. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 147-165.
 91. Gordon M. Concept development: response to Knafl. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 64-74.
 92. Miller JF. Differentiating phenomenologic and ethnographic research approaches: response to Dreher. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 98-113.
 93. Strauss A, Corban JM. Grounded Theory's applicability to nursing diagnostic research. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 4-24.
 94. Thomas MD. Grounded Theory and Nursing Diagnosis. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 25-36.
 95. Knafl KA. Concept development. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating

- Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 37-63.
96. Alexander C. Epidemiologic approaches to validation of nursing diagnoses. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 121-136.
97. Fritzmaurice J. Tool development: validity related to nursing diagnoses – response to Schroeder. In: Chang BL, Kim MJ, Jones P, McFarlane E, editores. Invitational Conference on Research Methods for Validating Nursing Diagnoses; 1989; Palm Springs. California: NANDA; 1989. p. 166-176.
98. Oliveira-Kumakura AR de S, Caldeira S, Simão TP, Camargo-Figuera FA, Cruz D de ALM da, Carvalho EC de. The Contribution of the Rasch Model to the Clinical Validation of Nursing Diagnoses: Integrative Literature Review. *Int J Nurs Knowl*. 2018;29(2):89-96. <https://doi.org/10.1111/2047-3095.12162>. PMID: 27781407.

CONTRIBUCIÓN DE LOS AUTORES

Diseño del proyecto: Lopes MVO, Silva VM.

Obtención de datos: Lopes MVO, Silva VM.

Análisis de datos: Lopes MVO, Silva VM.

Interpretación de datos: Lopes MVO, Silva VM.

Todos los autores son responsables de la redacción textual y la revisión crítica del contenido intelectual, de la versión final publicada y de todos los aspectos éticos, legales y científicos relacionados con la exactitud e integridad del estudio.



Copyright © 2026 Online Brazilian Journal of Nursing

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.