

Rev. Latino-Am. Enfermagem
2021;29:e3387
DOI: 10.1590/1518-8345.4157.3387
www.eerp.usp.br/rlae




Artículo Original

Directriz para la incorporación del método Delphi en la evaluación de las teorías de enfermería


Manuela Campos Gomes Borel¹

 <https://orcid.org/0000-0003-0116-5029>


Rafael Oliveira Pitta Lopes²

 <https://orcid.org/0000-0002-9178-8280>


Maira Buss Thofehrn¹

 <https://orcid.org/0000-0002-0864-3284>


Maria Miriam Lima Da Nóbrega³

 <https://orcid.org/0000-0002-6431-0708>

Cristina Arreguy-Sena¹

 <https://orcid.org/0000-0002-5928-0495>

Marcos Antônio Gomes Brandão⁴

 <https://orcid.org/0000-0002-8368-8343>

Objetivo: describir una directriz para el uso del método Delphi en la evaluación de teorías de enfermería, desde la perspectiva de la validación interna. **Método:** estudio metodológico, dirigido a la elaboración de una directriz para el uso del método Delphi en la evaluación de teorías de enfermería. **Resultados:** el método Delphi, los principios de sabiduría colectiva y los niveles de competencia se utilizan en la elaboración de una directriz para organización, búsqueda, selección y coordinación de las actividades de los evaluadores teóricos en equipos. Distingue tres fases para el proceso de evaluación teórica: fase preparatoria (FP); fase intermedia (FI) y fase de evaluación teórica (ET), con la incorporación de procedimientos de selección tipo Delphi; búsqueda, selección y clasificación de jueces/evaluadores de la teoría; definición de criterios para la realización de rondas y mantenimiento o disposición de unidades de la teoría evaluada. **Conclusión:** la directriz desarrollada logró adaptar los elementos del método Delphi como estrategia favorable para la validación interna de las teorías de enfermería.

Descriptorios: Conocimiento; Modelos Teóricos; Teoría de Enfermería; Estudio de Validación; Toma de Decisiones; Enfermería.





¹ Universidade Federal de Juiz de Fora, Faculdade de Enfermagem, Juiz de Fora, MG, Brasil.

² Universidade Federal do Rio de Janeiro, Campus Macaé, Macaé, RJ, Brasil.

³ Universidade Federal da Paraíba, Centro de Ciências da Saúde, João Pessoa, PB, Brasil.

⁴ Universidade Federal do Rio de Janeiro, Escola de Enfermagem Anna Nery, Rio de Janeiro, RJ, Brasil.

Cómo citar este artículo

Borel MCG, Lopes ROP, Thofehrn MB, Nóbrega MML, Arreguy-Sena C, Brandão MAG. Guideline for incorporating the Delphi method in the evaluation of nursing theories. Rev. Latino-Am. Enfermagem. 2021;29:e3387. [Access   ]; Available in: . DOI: <http://dx.doi.org/10.1590/1518-8345.4157.3387>.

Introducción

La construcción teórica por la estrategia teoría-investigación es un proceso iniciado en la elaboración de una teoría explícita en las fases de desarrollo conceptual y operacionalización, que, con posterioridad, debe avanzar hacia las fases de confirmación o no confirmación, hasta alcanzar las condiciones prácticas de aplicación⁽¹⁾. En las disciplinas aplicadas, se enfatiza la validación empírica para producir juicios sobre la utilidad y la aplicación.

El emprendimiento total de construcción teórica requiere la realización de evaluaciones para estimar el valor de una "buena" teoría, que es aquella rica en "virtudes teóricas", como señalan los filósofos de la ciencia. Entre estas virtudes, podemos destacar la observancia de singularidad, falseamiento, parsimonia, predicción, explicación, conservacionismo, capacidad de generalización, fecundidad, consistencia interna, riqueza empírica y abstracción⁽²⁾.

En la identificación de una buena teoría, la validación puede estar guiada por la perspectiva interna, con la evaluación de los elementos intrínsecos o por la perspectiva externa, mediante la evaluación a través de la prueba empírica⁽³⁾. En enfermería, la validación interna se denomina comúnmente "evaluación de la teoría" (*theory evaluation*), que tiene como objetivo determinar la adecuación de su uso y el enfoque epistemológico⁽⁴⁾. Dependiendo de los criterios formales que se utilicen, la evaluación puede incorporar el análisis o el desglose teórico.

A pesar de la relevancia de la evaluación de teoría y de la existencia de decenas de criterios estructurados y sistemáticos para su realización, aún es poco común verificar la aplicación de estos criterios en la literatura internacional. Las validaciones externas son más comunes en estudios empíricos con análisis estadísticos o revisiones de la literatura⁽⁴⁾.

Se advierte un interés continuo en la producción de teorías de mediano alcance para describir, explicar, predecir o prescribir mejor los fenómenos, los hechos, los eventos o las intervenciones que la enfermería aborda en la vida diaria. Sin embargo, estas teorías, junto a las de situación específica, raramente son evaluadas⁽⁴⁾.

Si, por un lado, este movimiento continuo contribuye al avance de la disciplina, por otro, exige la disponibilidad de instrumentos y directrices que promuevan buenas prácticas de desarrollo teórico. Por tanto, es indispensable el acceso a los recursos que puedan ayudar en el programa de elaboración, validación, perfeccionamiento y aplicación teórica.

La validación interna (evaluación) de una teoría exige al mismo tiempo conocimiento de la teoría y un alto nivel de conocimiento metateórico. Debido a esto, es difícil

encontrar metateóricos experimentados y disponibles para realizar esta tarea. Cuando se logra, la coordinación de la tarea es compleja, como lo es la identificación de un analista considerado competente mediante la aplicación de criterios epistemológicamente consistentes.

Ante la dificultad de ubicar expertos metateóricos, los principios y criterios de la sabiduría colectiva o la sabiduría de la multitud pueden ser útiles para la construcción de directrices, métodos o técnicas que guíen la formación de un equipo capaz de desarrollar la tarea de evaluación teórica con igual o superior resultado, en comparación con el de un solo experto metateórico. En la teoría de la sabiduría de la multitud, criterios como la independencia, la descentralización, la diversidad y la agregación orientarían la formación de grupos, en los que la decisión agregada superaría a la del especialista, por separado⁽⁵⁾.

De esta forma, los analistas, no necesariamente expertos en metateoría, funcionan como jueces del contenido, de la estructura y de otros criterios a evaluar. A partir del juicio agregado, se obtienen resultados consistentes que permiten llevar a cabo con éxito la evaluación de la teoría. Sin embargo, las directrices, métodos o técnicas con esta conformación no están disponibles para su uso en las teorías de enfermería.

Se presume que el método Delphi es adecuado para realizar la evaluación de una teoría de enfermería basada en criterios de sabiduría colectiva, demostrando que los grupos pueden juzgar, adecuadamente, en condiciones de incertidumbre, definir conceptos fundamentales, juzgar y agregar el valor colectivo de las ideas⁽⁵⁻⁷⁾. Se ha utilizado para tratar cuestiones no resueltas por enfoques experimentales, en las que la opinión de un grupo tiene valor para esclarecerlas siendo, por tanto, compatible con la validación interna⁽⁶⁾.

Sin embargo, su aplicación para este propósito es escasa. Su uso se identificó en la literatura, solamente en una teoría de la interfaz educación-informática en la evaluación de los criterios de importancia, precisión y claridad, parsimonia o sencillez, comprensión, operacionalización, validez empírica, fructificación y aplicación⁽⁸⁾. La descripción metodológica en el estudio mencionado no aporta elementos suficientes para su uso en la evaluación de las teorías de enfermería con criterios formales, generalmente aplicados en la disciplina⁽⁴⁾.

En Brasil, el método Delphi ha contribuido en el abordaje de problemas empíricos como la indicación de tendencias, la obtención de consenso sobre un programa o intervención, la opinión de expertos para comparar tratamientos y, más ampliamente, en la construcción de instrumentos evaluativos y creación y validación de instrumentos⁽⁹⁾. La adaptación del método para la evaluación de las teorías de enfermería sigue siendo un

potencial que aún no ha sido explorado, incluso dado su carácter innovador. Este artículo fue elaborado en vista de la escasez de investigación y las potencialidades advenidas de la elaboración de una directriz para enfrentar el desafío presentado.

El artículo tiene como objetivo describir una directriz para el uso del método Delphi en la evaluación de teorías de enfermería, desde la perspectiva de la validación interna.

Método

Estudio metodológico de elaboración de una directriz para el uso del método Delphi en la evaluación de teorías de enfermería, indicando procedimientos de organización, búsqueda, selección y coordinación de las actividades de los evaluadores teóricos en equipos. Los criterios de sabiduría colectiva y niveles de competencia fueron la base de referencia⁽⁵⁾ y su elaboración tuvo lugar en Río de Janeiro, RJ, Brasil, entre los meses de noviembre y diciembre de 2019.

Se incorporaron los elementos utilizados en los marcos metodológicos de concepción, construcción y prueba de directrices; destacándose: selección del tema y alcance; adaptación de un prototipo de directriz de estrategia de evaluación teórica, utilizando el método Delphi; formación de grupos para desarrollo; búsqueda sistemática de evidencia; análisis y síntesis de la evidencia disponible y creación de la recomendación⁽¹⁰⁾.

Los procedimientos específicos para el desarrollo de la directriz fueron: simple revisión de la literatura del manuscrito sobre el uso del método Delphi en la evaluación de teorías y otras aplicaciones; interpretación de los métodos de evaluación de la teoría de enfermería^(4,11-12); selección de material complementario sobre el tema de la sabiduría colectiva; recopilación e interpretación de los resultados del uso de un prototipo de una guía de

evaluación teórica, desarrollada en una tesis de maestría por uno de las autoras, incorporando características del método Delphi; elaboración de la directriz, teniendo en cuenta los principios de construcción de directrices en salud y las adaptaciones necesarias al objeto teórico-filosófico; discusión y revisión por los autores; elaboración final de la directriz con diagramación.

El prototipo desarrollado en la tesis de maestría tuvo los siguientes pasos: (a) selección de *experts* (b) contacto con expertos e invitación a participar de los seleccionados; (c) envío del instrumento, por vía electrónica, a quienes aceptaron participar; (d) apreciación de ítems para evaluar la teoría con base en la escala de acuerdo Likert; (e) recepción de respuestas; (f) análisis cualitativo y cuantitativo de los resultados; (g) adecuación del contenido para una nueva ronda de evaluación teórica; (h) reenvío de *feedback* conteniendo datos que llevaron a la modificación o mantenimiento de los ítems para realizar una nueva evaluación; (i) recepción de respuestas del instrumento adaptado; (j) análisis de datos de la segunda versión; (l) construcción final por consenso; (m) revisión gramatical y ortográfica y (n) cierre de la evaluación teórica.

La disertación de maestría que incorporó el uso del prototipo evaluó la Teoría de los Vínculos Profesionales⁽¹³⁾ por la estrategia de evaluación teórica de Meleis⁽¹⁴⁾. El estudio que aplicó el prototipo de la directriz, respetó los principios éticos de la investigación contenidos en la Resolución 466/2012 del Consejo Nacional de Salud, obteniendo dictamen de aprobación del Comité de Ética en Investigación, con el número 3.237.583.

La directriz que promueve el uso del método Delphi en la evaluación de las teorías de enfermería, tiene tres fases: fase preparatoria (FP), fase intermedia (FI) y fase de evaluación teórica (ET). El presente estudio detalla la fase intermedia como se expresa en la Figura 1.



Figura 1 - Esquema de la directriz para incorporación del método Delphi en la evaluación de las teorías de enfermería. Río de Janeiro, RJ, Brasil, 2019

En la fase preparatoria (FP) se decide sobre la teoría que será evaluada y se opta por la estrategia de evaluación a emplear de entre las alternativas disponibles en la literatura.

En la fase intermedia, se describen nueve procedimientos relacionados con el uso del método Delphi. El primer procedimiento está relacionado con el tipo de Delphi a utilizar, influenciado por el nivel de la teoría a evaluar y por la madurez de su aplicación.

En el segundo procedimiento se define el rol coordinador de la evaluación teórica, que se puede acumular con la condición de evaluador primario. Esto puede ser realizado por un miembro del equipo de desarrollo de la teoría o por otro investigador con conocimiento consistente al respecto, que será responsable de la preparación y conducción del proceso de evaluación. El papel del evaluador primario es ayudar a los evaluadores secundarios a acceder a los materiales de análisis, pudiendo producir un informe preliminar con los resultados de la tarea realizada que sirve como documento básico destinado a la consulta de los demás evaluadores. De lo contrario, los secundarios operan como jueces desde el documento básico o realizando una función de evaluación primaria.

El tercer procedimiento es la ubicación y selección de candidatos para evaluadores/jueces. La búsqueda puede ser en plataformas que contengan *curriculum*s currículos electrónicos, siendo los filtros: la nacionalidad, la formación académica, el idioma y la actuación profesional, entre otros. Se recomienda la apreciación de publicaciones, proyectos o investigaciones y la verificación del tiempo de experiencia en el área para la cuidadosa selección del candidato a evaluador. La selección de candidatos para evaluadores secundarios suele ser difícil, en ausencia de criterios de clasificación para una supuesta experiencia metateórica.

Así, con base en la sabiduría colectiva⁽⁵⁾, los equipos de evaluadores (jueces) deben tener diversidad en el ámbito de pericia. Para estimar la pericia de los jueces se proponen criterios en cinco dominios: formación educativa en teorías de enfermería; experiencia profesional en el campo de la teoría; experiencia y conocimiento metateórico; difusión del conocimiento producido sobre el tema de las metateorías o teorías de enfermería y reconocimiento de la experiencia en el tema de las metateorías o teorías de enfermería, por parte de los pares (Figura 2).

Dominios y criterios para categorización de los evaluadores/jueces de teoría	Puntaje
1. Formación educativa en teorías de enfermería:	
a) Doctorado en Enfermería o áreas afines y tesis teórica o metateórica	4 puntos
b) Maestría en Enfermería o áreas afines y tesis teóricas o metateóricas	3 puntos
c) Doctorado en enfermería con tesis sobre otros temas	2 puntos
2. Experiencia profesional en el área de la teoría a evaluar:	
(a) Más de cuatro años de experiencia en docencia, investigación o práctica clínica en el área/tema de la teoría a ser evaluado y, al menos, un año en docencia de teorías o investigación de enfermería teórica o metateórica	4 puntos
b) Más de cuatro años de experiencia en la docencia, investigación o práctica clínica en el área/tema de la teoría a ser evaluada	3 puntos
c) Entre dos y cuatro años de experiencia en docencia, investigación o práctica clínica en el área/tema de la teoría a evaluar o, al menos, un año en docencia de teorías o investigación de enfermería teórica o metateórica	2 puntos
3. Experiencia y conocimientos metateóricos:	
a) Elaboración u orientación de más de una teoría o metateoría de enfermería	4 puntos
b) Elaboración u orientación de una teoría o metateoría de enfermería	3 puntos
c) Elaboración u orientación de, al menos, una teoría o metateoría no enfermería	2 puntos
4. Difusión del conocimiento producido sobre el tema de las metateorías o teorías de enfermería:	
a) Autoría en más de un artículo publicado en una revista indexada de estándar internacional con alto factor de impacto (Para Brasil: Qualis A1) sobre el tema de la teoría o metateoría de enfermería	4 puntos
b) Autoría de un artículo publicado en una revista indexada de estándar internacional con alto factor de impacto (Para Brasil: Qualis A2 a A4) sobre el tema de teoría o metateoría de enfermería	3 puntos
c) Autoría de, al menos, un artículo publicado en una revista indexada a nivel nacional con un factor de impacto medio (Para Brasil: Qualis B1) sobre el tema de la teoría o metateoría de enfermería	2 puntos
5. Reconocimiento de la experiencia en el tema de las metateorías o teorías de enfermería, por pares:	
a) Más de una participación como invitado (conferenciante, disertante, comentarista, docente en curso o curso corto) en un evento científico para impartir un tema relacionado con las teorías o metateorías de enfermería	4 puntos
b) Participación como invitado (conferenciante, disertante, comentarista, docente en curso o curso corto) en un evento para impartir una temática relacionada con las teorías o metateorías de enfermería	3 puntos
c) Participación como oyente/participante/alumno en un evento o curso completo sobre teorías o metateorías de enfermería	2 puntos

Figura 2 - Criterios para clasificar el nivel presunto de experiencia de los evaluadores de las teorías de enfermería. Rio de Janeiro, RJ, Brasil, 2019

El candidato a evaluador/juez tiene su nivel de experiencia jerarquizado por la puntuación obtenida por la suma de los ítems de los cinco dominios. La puntuación máxima para cada dominio es de 4 puntos y la mínima es de 2. Cuando el evaluador no cumple con ninguno de los criterios de un dominio, recibe una puntuación de 0. La puntuación total varía de 0 a 20, siendo 4 puntos el valor mínimo arbitrado para considerar al candidato apto para componer el equipo evaluador teórico.

El nivel de supuesta experiencia de un candidato a evaluador se establece mediante el puntaje obtenido en

la evaluación de los cinco dominios. Se proponen cinco niveles: principiante, principiante avanzado, competente, proficiente y experto.

La clasificación del supuesto nivel de pericia de un candidato se establece a partir de su puntaje total obtenido (ver Figura 1).

La definición del número mínimo de evaluadores en el equipo dependerá de la suma de los puntos individuales de cada evaluador y de la composición mixta que garantiza miembros de, al menos, dos niveles diferentes de experiencia (Figura 3).

Nivel supuesto de experiencia del evaluador	Puntuación requerida para adaptarse al nivel	Criterios para definición el equipo, según los puntos de experiencia
Principiante	suma mínima de 4 puntos	(a) Equipo de dos evaluadores: la suma de los puntos de los evaluadores debe ser de al menos 35 puntos
Principiante avanzado	suma entre 5 y 10 puntos	(b) Equipo de tres evaluadores: la suma de los puntos de los evaluadores debe ser de al menos 36 puntos
Competente	suma entre 11 puntos y 14 puntos	(c) Equipo de cuatro evaluadores: la suma de los puntos de los evaluadores debe ser de al menos 48 puntos
Proficiente	suma de puntuaciones entre 15 y 17 puntos	(d) Equipo de cinco evaluadores: la suma de los puntos de los evaluadores debe ser de al menos 60 puntos
Experto	suma de puntuaciones mayores o iguales a 18 puntos	(e) Equipo de seis o más evaluadores: aplicar la siguiente ecuación:

Fuente: Brandão, 2019.

Figura 3 - Clasificación del nivel de presunta experiencia de los evaluadores según el puntaje obtenido y definición del número de evaluadores del equipo. Rio de Janeiro, RJ, Brasil, 2019

Por ejemplo, por la aplicación de la ecuación propuesta, la composición de un equipo de siete evaluadores por el cálculo del total de puntos requeridos de la suma de los evaluadores exigirá, aproximadamente, 74 puntos. Ejemplificando, una configuración adecuada incluiría: seis evaluadores principiantes avanzados de 10 puntos cada uno y un evaluador competente, con 14 puntos. Evidentemente, se pueden aplicar otras configuraciones que respeten la puntuación mínima del equipo. Además de la puntuación total, la exigencia de equipos mixtos en términos de niveles de competencia objetiva garantizar el criterio de diversidad de jueces.

Considerando que las bajas ante invitaciones y durante las rondas iniciales de Delphi son comunes, se recomienda seleccionar un número mayor de jueces para la composición mínima de los equipos, buscando mantener la proporcionalidad entre los niveles.

El cuarto procedimiento es la invitación a los candidatos seleccionados, basada en la comunicación por medios escritos o electrónicos y respetando los principios éticos de la investigación.

El quinto procedimiento implica la reaplicación de los criterios para la categorización de los evaluadores/jueces por el nivel de experiencia y para ajustes en la composición de los equipos, teniendo en cuenta las negativas a participar.

El sexto procedimiento es la planificación de rondas y criterios de interrupción. Esta planificación considera el nivel de abstracción, el número de conceptos y la complejidad de la teoría a evaluarse. También se destaca el número de jueces que desarrollan las evaluaciones y, también, la consideración en relación con la cantidad de criterios a ser evaluados en teoría. La interrupción de las rondas de evaluación debe estar sustentada en el juicio explícito de los evaluadores con el alcance de un consenso evaluativo o constitución de una multiplicidad de ideas en el disenso. Otra decisión es establecer o no, *a priori*, un máximo de rondas. Esta decisión se refiere más a la variable tiempo disponible para la evaluación que al juicio de evaluación.

Posteriormente, para orientar a los evaluadores, se deben establecer criterios explícitos de descarte de ítems en cada ronda. Los elementos de una teoría sujetos a evaluación son sus componentes como conceptos, presupuestos, suposiciones, enunciados y esquemas modulares. Así, los evaluadores deben tener claro que la decisión de exclusión viene impulsada por la estrategia seleccionada y, no solo, por sus opiniones personales, libremente emitidas. Su función es juzgar un determinado ítem de la teoría contra los criterios de evaluación establecidos en la estrategia.

El octavo procedimiento implica la definición del alcance del consenso y la estabilidad de las respuestas. El consenso se puede verificar mediante medidas formales de acuerdo, medidas de tendencia central, porcentaje de acuerdo, medida de tendencia central dentro de un intervalo específico, entre otras⁽¹⁵⁾. Se puede planificar el empleo de una escala Likert de cinco puntos para dos propósitos: (1) verificar el acuerdo de los evaluadores secundarios con el resultado de la evaluación primaria o (2) organizar las evaluaciones secundarias en declaraciones que serán sometidas al análisis de acuerdo en una ronda posterior.

Incluso cuando se aplican escalas, se recomienda garantizar campos de edición libres para que los evaluadores/jueces puedan expresar sus sugerencias, recomendaciones y valoraciones detalladas.

El noveno procedimiento es el fomento de una directriz específica sobre la estrategia de evaluación teórica. Cuando el coordinador o evaluador primario lo considere necesario, se puede realizar capacitaciones complementarias y específicas sobre el contenido de la estrategia.

Discusión

La evaluación teórica es capaz de aportar elementos sobre una "buena" teoría, con varios criterios formales y sistemáticos disponibles en la literatura^(2,11). Sin embargo, no siempre se dispone de recursos humanos con la competencia y conocimiento necesario para desarrollar, adecuadamente, el proceso de juzgar las virtudes teóricas. Y, en este sentido, el método Delphi, con la utilización del principio de sabiduría colectiva por consenso o disenso, puede multiplicar la experiencia del grupo, aun ampliando el universo de alternativas para la evaluación⁽¹⁵⁻¹⁶⁾. Asimismo, asiste en la coordinación del trabajo de evaluación.

A través de la evaluación, se advierten relaciones y vínculos de conceptos que permiten al revisor la verificación de punto de fortalezas y limitaciones de la teoría; identificación de la necesidad de nuevos elementos de la teoría o perfeccionamiento de los existentes y, como meta final, la determinación del potencial aporte de la teoría evaluada para el conocimiento científico⁽¹¹⁾.

A diferencia del análisis teórico (*theory analysis*) que descompone una teoría para examinar sus partes o componentes⁽⁴⁾, la evaluación teórica también los juzga. Sin embargo, incluso una teoría considerada como "buena" puede resultar inadecuada en su valor descriptivo, explicativo, predictivo o prescriptivo a partir de la confirmación o aplicación. Esto posiciona a la validación interna como un paso relevante, aunque no final, en un programa de desarrollo teórico.

Las teorías que violen las virtudes de "buena" teoría son más difíciles de refutar y tienden a no contribuir, realmente, al conocimiento⁽¹⁷⁾, pues la insuficiencia de elementos y construcciones dificulta la evaluación y prueba teórica. Por lo tanto, es de fundamental importancia planificar la validación interna y/o externa como parte de un programa más completo. El uso de la directriz presentada puede evitar el gasto de recursos, al colaborar en la identificación de teorías que no dispongan de virtudes suficientes para sostener una validación por investigación de campo.

Las razones del uso reducido de las estrategias de evaluación de la teoría de enfermería mediante criterios sistemáticos formales son inciertas⁽⁴⁾. Sin embargo, la influencia se puede atribuir a la dificultad de obtener evaluadores con suficiente autoridad epistémica para juzgar los ítems metateóricos de validación interna. Supone que las estrategias vinculadas a la sabiduría colectiva puedan superar este problema de dependencia del "experto" con ventajas sustanciales⁽¹⁸⁾.

El método Delphi se basa en los supuestos de John Dewey, que enfatiza la comunicación anónima entre individuos con experiencia en un tema determinado, con el objetivo de buscar la opinión de expertos de manera iterativa y estructurada, que, en general, intenta lograr una posición consensuada^(15,19). La libertad y la observancia de las opiniones personales de los jueces garantiza el criterio de independencia de la sabiduría colectiva⁽⁵⁾.

En cuanto a su uso en investigaciones, aunque sea predominantemente usado en las mixtas y cuantitativas, tiene aplicación cualitativa e incluso en la construcción de teorías prácticas, en el contexto de la organización comunitaria⁽¹⁵⁾. La evaluación de una teoría es un emprendimiento cualitativo marcado por la subjetividad y por los estándares, conducta y códigos del evaluador⁽⁸⁾.

El método Delphi puede coordinar estas características cualitativas del proceso de evaluación, tratando con variables personales desde el criterio de independencia y aprovechar el trabajo en grupo. Puede utilizarse para interpretación, predicción y obtención de recomendaciones de la evaluación desarrollada⁽⁸⁾.

Al elegir el método Delphi, el enfoque más común es el tradicional, también llamado normativo o de consenso. Su objetivo es reducir la variación en las estimaciones y los sesgos entre expertos. Ya el *Policy Delphi*, o disenso, busca obtener una amplia gama de opiniones, pero sin buscar consensos⁽¹⁶⁾.

Para la evaluación teórica, el proceso de consenso por el método Delphi es la indicación más probable; sin embargo, el empleo del disenso puede recomendarse para teorías de alta originalidad, densidad conceptual, complejidad y abstracción teórica o cuando sea difícil determinar los criterios de consenso. Además, el objetivo

de la evaluación puede ser explorar lo contradictorio en la producción de definiciones o proposiciones teóricas.

En cuanto a las características de la teoría, el método Delphi para obtención de consenso puede ser indicado para rango micro o medio con conceptualización descripta en más de un estudio empírico o para evaluar una teoría parcialmente difundida, probada o utilizada.

Supuestamente, para el Delphi de consenso la composición de equipos con un alto número de evaluadores solo se justifica cuando es difícil obtener evaluadores con niveles más altos de experiencia, porque en grupos de muchos componentes es un reto lograr el consenso. Por otro lado, se asume que el enfoque de disenso se beneficia de la composición de equipos más grandes y con una amplia gama de niveles de competencia, tendiendo a ampliar el debate desde diferentes perspectivas y aportar elementos originales, que difieren de la teoría original y de la evaluación primaria.

Paneles con más participantes tienden a tener tasas de respuesta más bajas, con una reducción estimada de 0,08 puntos porcentuales por cada participante agregado⁽²⁰⁾. Se indican de 5 a 20 expertos si se trata de una recomendación basada exclusivamente en las características del método Delphi⁽²⁰⁾. Los estudios sobre el desarrollo y la aplicación de *core outcome set* (COS) han utilizado el método Delphi para determinar qué resultados medir y en ellos predominan los paneles Delphi de hasta 50 personas⁽²⁰⁾.

En la evaluación teórica, es un desafío poder establecer un número mínimo y máximo de evaluadores/jueces, debido a su carácter filosófico y a la naturaleza epistemológica abstracta inherente a la teorización. Por ejemplo, en el caso de teorías nuevas o poco difundidas, puede resultar difícil contar con muchos evaluadores secundarios con la experiencia adecuada. Por otro lado, grandes equipos de evaluadores principiantes pueden no disponer de conocimientos de naturaleza metateórica, provocando una dispersión de perspectivas que dificultaría la agregación de ideas. En este caso, la directriz busca eludir los límites combinando un equilibrio entre el criterio de diversidad del principio de sabiduría colectiva y la experiencia requerida para la evaluación teórica⁽⁴⁻⁵⁾.

En el prototipo de la directriz, se incorporaron cuatro evaluadores de tres niveles diferentes de experiencia, siendo tres evaluadores secundarios que en conjunto alcanzaron 36 puntos (14, 13 y 9 puntos individuales). Los criterios de definición del equipo, según los puntos de especialización, fueron útiles para la composición de este pequeño grupo, ya que la configuración de menos participantes garantizaba la máxima tasa de respuesta, como se ha esperado para este tamaño de panel⁽²⁰⁾. Las diferencias de niveles y narración de capacitación de los evaluadores aseguraron

el criterio de descentralización⁽⁵⁾. Sin embargo, siempre que sea posible, se recomienda formar equipos con cinco o más jueces.

El modelo de Patricia Benner⁽²¹⁻²²⁾ con sus cinco niveles de adquisición de competencias fue la base para crear los criterios de pericia de los jueces en la directriz y buscaba reconocer la experiencia profesional como un componente esencial para la validación. La amplia difusión de los estudios de estos autores y sus criterios ayudaron en la definición. Se advierten modelos más complejos de reglas de agregación para definir la composición del equipo, como el *Contribution Weighted Model* (CWM), que pondera los pronósticos con base en el desempeño relativo de cada juez y la precisión del grupo^(18,23).

Los aportes de los evaluadores/jueces están cargados de conocimientos, experiencias y puntos de vista particulares en la evaluación de la teoría. El proceso interactivo del método Delphi puede permitir que en las rondas, los criterios oscuros de la evaluación puedan ser aclarados o modificados, mediante una cuidadosa interpretación de las respuestas de los evaluadores secundarios, por parte del coordinador. La composición de equipos con diferentes niveles de competencia garantiza el criterio de diversidad de sabiduría colectiva⁽⁵⁾.

Los estudios suelen aplicar de dos a tres rondas en el método Delphi⁽¹⁹⁾. Sin embargo, los múltiples criterios a evaluar, el elevado número y la diversidad de perfiles de los evaluadores pueden requerir más rondas para llegar a un consenso. Es deseable planificar un mínimo en función de la cantidad de evaluadores, para asegurar que un esfuerzo excesivo para gestionar los resultados de la evaluación no recaiga sobre el coordinador del Delphi, poniendo en riesgo la calidad del trabajo.

El nivel de exhaustividad de las teorías puede influir en la definición de los criterios a ser evaluados por los jueces; por ejemplo, cuando una determinada teoría de rango medio se evalúa como modelo, se pueden utilizar criterios aún más específicos y empíricos⁽¹²⁾. Sin embargo, esto no interfiere, directamente, con la naturaleza del método Delphi como estrategia.

La decisión de consensuar entre los jueces es un tipo de mecanismo para cumplir con el criterio de agregar sabiduría colectiva, transformando los juicios individuales en una decisión en equipo⁽⁵⁾. La decisión consensuada puede partir de la propia opinión de los evaluadores de que se llegó a un consenso; sin embargo, se recomienda que esto no opere de manera automática después de la realización de la técnica Delphi⁽¹⁹⁾.

Es necesario especificar, claramente, cuáles son las condiciones para llegar a un consenso, cuando la decisión es cualitativa. Cuando se adoptan procedimientos de medición cuantitativa, es necesario disponer las medidas y los puntos de corte que se utilizarán para determinar

el grado de acuerdo o desacuerdo, compatible con el consenso o disenso⁽¹⁹⁾.

No existen reglas obligatorias para establecer el consenso, pero la escala Likert de cinco puntos es la más común entre las escalas utilizadas para estimar los desacuerdos o concordancias⁽²⁴⁾. La misma permite la comprobación del grado de concordancia de cada artículo o conjunto. Para las estimaciones del consenso alcanzado, utilizando la escala Likert, se pueden aplicar medidas formales de acuerdo, como la estadística Kappa, para verificar el juicio concordante de los jueces sobre los elementos de la teoría.

En general, los porcentajes de concordancia adoptan el valor de 0,8 u 80% como punto de corte mínimo⁽¹⁵⁾. Sin embargo, el investigador puede considerar otros puntos de corte, respaldados por evidencia o una recomendación consistente. Se formula una declaración explícita sobre el alcance del consenso con una indicación de las razones que se consideraron en la toma de decisiones.

Otra medida útil que se puede utilizar sobre los datos obtenidos por la escala Likert es el coeficiente de validez de contenido. El coeficiente de Aiken y su tabla de probabilidad caudal se pueden utilizarse para indicar la validez de un ítem en particular evaluado por varios jueces, estimando un consenso. También se puede aplicar para juzgar la validez, por un solo juez, del contenido de todos los ítems de la teoría. El rango de coeficientes varía de 0 a 1 y los valores más altos indican validación⁽²⁵⁾.

A pesar de que la literatura, en general, recomienda el uso de escalas cuantitativas para señalar el consenso, se deben agregar justificaciones cualitativas, especialmente cuando la recomendación se debe a la invalidez del ítem^(24,26). La simple exclusión de un ítem puede hacer que toda la teoría sea incoherente o ilógica. Esto se debe a que las unidades de una teoría desempeñan roles y tienen diferente relevancia en la estructura teórica.

Por ejemplo, la exclusión de un presupuesto puede desnaturalizar la teoría como un todo, ya que esta tipología de elementos funciona como premisas no dadas a la prueba empírica. Así, su retirada niega las ideas que guiaron al propio teórico en la construcción de la teoría. Por otro lado, las unidades del tipo "proposiciones" se presentan, precisamente, para generar hipótesis de prueba en estudios de validación empírica, por lo que, naturalmente, están sujetas a exclusión o mantenimiento después de las evidencias obtenidas de la experimentación o de la investigación de campo⁽¹⁾.

Es de destacar que, a partir de la evaluación de la Teoría de los Vínculos Profesionales⁽²⁷⁻²⁸⁾, factores emergentes demandaron cambios de criterios que no fueron detallados cuando se elaboró el prototipo, lo que contribuyó a profundizar en la elaboración de la directriz presentada en este artículo.

El estudio no está exento de limitaciones. El enfoque de cualquier investigación que utilice el Delphi será siempre la obtención de respuestas de alta calidad de una selección de personas expertas⁽²⁹⁾. Sin embargo, la validación interna de una teoría lidia con criterios teórico-filosóficos que pueden dificultar que, un evaluador secundario, produzca o juzgue la calidad de las respuestas, por la naturaleza del objeto evaluado y por la propiedad del juicio a realizar. Por ejemplo, la definición conceptual es uno de los elementos de una teoría, evaluado en semántica, lógica y contexto⁽¹⁴⁾. En particular, puede ser difícil hacer un juicio de "buena respuesta" para un constructo tan complejo, frente a tales propiedades.

La limitación de la subjetividad del juicio del juez en la evaluación de teorías debe confrontarse con la raíz filosófica del teórico y del evaluador. Raíces crítico-sociales, hermenéuticas o del nuevo pragmatismo tienden a lidiar con una mayor fluidez frente a las diferentes perspectivas, incluso explorarlas en consenso o disenso. Por otro lado, el pospositivismo, al requerir mayor objetividad en la realidad, demanda criterios más estables, generalizables o de medición⁽³⁰⁾. En esta última raíz filosófica, métodos como el modelado de ecuaciones estructurales, el análisis factorial y las regresiones múltiples pueden ser la mejor opción para la evaluación de la teoría, obviamente con criterios más cercanos a la validación externa⁽⁴⁾.

Entre los aportes al avance del conocimiento científico, el estudio agrega una dimensión inexplorada de la incorporación de evaluadores de diferentes niveles de experiencia metateórica en la tarea de evaluación teórica, incluida la incorporación de directrices para las fases de este proceso. Ante el crecimiento en el desarrollo de las teorías de enfermería de rango medio y micro y de una situación específica, con la consecuente formación de nuevos teóricos, la directriz puede facilitar el proceso de validación de nuevas teorías y forjar una base sólida de conocimiento disciplinario^(4,31).

Por otro lado, los niveles más altos de experiencia metateórica tienden a requerir largos años de formación en este campo, siendo más común verificarlos en medio académico y entre los investigadores veteranos. Al explorar el principio de diversidad de niveles de competencia para la evaluación teórica, el estudio fomenta la creación de equipos de diferentes niveles de competencia y promueve la cooperación y circulación del conocimiento entre los involucrados en la evaluación.

Finalmente, la aplicación del método Delphi en la evaluación de una teoría de enfermería debe entenderse claramente como diferente a la búsqueda de consensos sobre eventos, fenómenos, hechos, tecnologías, conductas o cualquier otro elemento fundamentalmente empírico. En el método Delphi empírico, la exclusión de un ítem puede tener una implicación mínima; sin

embargo, en la teoría, la evaluación de la insuficiencia de las suposiciones o conceptos centrales puede catalogar a toda la teoría como inadecuada. Obviamente, el objetivo principal de la evaluación es identificar una “buena teoría”, lo que implica juzgar la adecuación de sus componentes; sin embargo, este procedimiento debe ser realizado con extrema cautela por parte de los evaluadores, entendiéndolo que en una teoría existe jerarquía y relación entre los elementos.

Conclusión

La directriz elaborada logró adecuar los elementos del método Delphi como recurso favorable para la validación interna de las teorías de enfermería, potenciándolo con la incorporación de jueces con diferentes visiones del mundo, experiencias, conocimientos científicos y creatividad. Los criterios presentados en la directriz adaptan y articulan los niveles de competencia de los evaluadores con el principio de la sabiduría colectiva, sirviendo de guía para la selección y la composición de los equipos de jueces, además de facilitar la coordinación del trabajo de evaluación teórica. Debido a su carácter innovador, la directriz puede instrumentar a los metateóricos de enfermería y, posiblemente, acelerar el proceso de aplicación de las teorías en la práctica.

El uso de un prototipo de directriz, en la evaluación de una teoría de rango medio de enfermería, la Teoría de los Vínculos Profesionales, arrojó resultados satisfactorios que presumen su factibilidad y señalaron caminos de refinamiento.

Se entiende que es importante que otros investigadores repliquen su uso en la evaluación de teorías de gran y micro alcance para futuros ajustes y actualizaciones de la directriz, además de adoptar estrategias de evaluación por criterios formales diferentes al utilizado en el prototipo.

Referencias

1. Lynham SA. The General Method of Theory-Building Research in Applied Disciplines. *Adv Dev Hum Resour.* 2002 Aug 1;4(3):221-41. doi: 10.1177/1523422302043002
2. Naor M, Bernardes ES, Coman A. Theory of constraints: is it a theory and a good one? *Int J Prod Res.* 2013;51(2):542-54. doi: 10.1080/00207543.2011.654137
3. Wallis SE. Validation of Theory: Exploring and Reframing Popper's Worlds. *Integr Rev.* [Internet]. 2008 Dec [cited Jan 20, 2020];4(2):71-91. Available from: <https://ssrn.com/abstract=2322989>
4. Im EO. The current status of theory evaluation in nursing. *J Adv Nurs.* 2015 May;71(10):2268-78. doi: 10.1111/jan.12698
5. Yousefnezhad M, Huang S-J, Zhang D. WoCE: A framework for Clustering Ensemble by Exploiting the Wisdom of Crowds Theory. *IEEE Trans Cybern.* 2018;48(2):486-99. doi: 10.1109/TCYB.2016.2642999
6. Fleenor JW. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economics, Societies and Nations. *Pers Psychol.* 2006;59(4):982-5. doi: 10.1111/j.1744-6570.2006.00060_10.x
7. Jorm AF. Using the Delphi expert consensus method in mental health research. *Aust New Zeal J Psychiatry.* 2015 Aug 21;49(10):887-97. doi: 10.1177/0004867415600891
8. Holton EF, Lowe JS. Toward a general research process for using Dubin's theory building model. *Hum Resour Dev Rev.* 2007 Sep 1;6(3):297-320. doi: 10.1177/1534484307304219
9. Revorêdo LDS, Maia RS, Torres GDV, Chaves Maia EM. O uso da técnica Delphi em saúde: uma revisão integrativa de estudos brasileiros. *Rev Arq Ciências Da Saúde.* 2015 Abr-jun;22(2):16-21. doi: 10.17696/2318-3691.22.2.2015.136
10. Chen Y, Yang K, Marušić A, Qaseem A, Meerpohl JJ, Flottorp S, et al. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. *Ann Intern Med.* 2017 Jan;166(2):128. doi: 10.7326/M16-1565
11. McEween M, Wills EM. *Theoretical Basis for Nursing.* 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2014.
12. Risjord M. Middle-range theories as models: New criteria for analysis and evaluation. *Nurs Philos.* 2019;20(1):e12225. doi: 10.1111/nup.12225
13. Thofehrn MB, Leopardi MT. Teoria dos vínculos profissionais: um novo modo de gestão em enfermagem. *Texto Contexto Enferm.* 2006 jul/set;15(3):409-17. doi: 10.1590/S0104-07072006000300004
14. Ramalho Neto JM, Marques DKA, Fernandes MGM, Nóbrega MML. Meleis' Nursing Theories Evaluation: integrative review. *Rev Bras Enferm.* 2016 Jan/Feb;69(1):162-8. doi: 10.1590/0034-7167.2016690123i
15. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol.* 2014 Feb 26;67(4):401-9. doi: 10.1016/j.jclinepi.2013.12.002
16. Waggoner J, Carline JD, Durning SJ. Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future Consensus Research. *Acad Med.* 2016 May;91(5):663-8. doi: 10.1097/ACM.0000000000001092
17. Wacker JG. A definition of theory: research guidelines for different theory-building research methods in operations management. *J Oper Manag.* 1998 Nov 25;16(4):361-85. doi: 10.1016/S0272-6963(98)00019-9

18. Budescu DV, Chen E. Identifying Expertise to Extract the Wisdom of Crowds. *Manage Sci.* 2015 Feb;61(2):267-80. doi: 10.1287/mnsc.2014.1909
19. Brady SR. Utilizing and Adapting the Delphi Method for Use in Qualitative Research. *Int J Qual Methods.* 2015 Dec 5;14(5):1-6. doi: 10.1177/1609406915621381
20. Gargon E, Crew R, Burnside G, Williamson PR. Higher number of items associated with significantly lower response rates in COS Delphi surveys. *J Clin Epidemiol.* 2019;108:110-20. doi: 10.1016/j.jclinepi.2018.12.010
21. Benner P. From Novice to Expert. *Am J Nurs.* [Internet]. 1982 Mar [cited Jan 20, 2020];82(3):402-7. Available from: https://journals.lww.com/ajnonline/Citation/1982/82030/From_Novice_To_Expert.4.aspx
22. Bathish M, Wilson C, Potempa K. Deliberate practice and nurse competence. *Appl Nurs Res.* 2018 Apr;40:106-9. doi: 10.1016/j.apnr.2018.01.002
23. Chen E, Budescu D V., Lakshmikanth SK, Mellers BA, Tetlock PE. Validating the Contribution-Weighted Model: Robustness and Cost-Benefit Analyses. *Decis Anal.* 2016 May;13(2):128-52. doi: 10.1287/deca.2016.0329
24. Marques JBV, Freitas D de. Método DELPHI: caracterização e potencialidades na pesquisa em educação. *Pro-Posições.* 2018 Mai/Ago;29(2):389-415. doi: 10.1590/1980-6248-2015-0140
25. Nair GG, Hellsten L-AM, Stamler LL. Accumulation of Content Validation Evidence for the Critical Thinking Self-Assessment Scale. *J Nurs Meas.* 2017;25(1):156-70. doi: 10.1891/1061-3749.25.1.156
26. Toma C, Picioreanu I. The Delphi Technique: Methodological Considerations and the Need for Reporting Guidelines in Medical Journals. *Int J Public Health Res.* [Internet]. 2016 Dec [cited Apr 20, 2020];4(6):47-59. Available from: <http://www.openscienceonline.com/journal/archive2?journalId=718&paperId=3586>
27. Jacondino MB, Martins CL, Thofehrn MB, Garcia BL, Fernandes HN, Joner LR. Vínculos profissionais en el trabajo de enfermería: elemento importante para el cuidado. *Enfermería Glob.* 2014 Abr;13(2):160-71. doi: 10.6018/eglobal.13.2.159881
28. Thofehrn MB, Quadros LDCM, Dias DG, Joner LR, Porto AR, Garcia BL. Teoria dos vínculos profissionais: visão dos enfermeiros que a implementaram no Brasil. *Enferm Foco.* 2013;4(2):127-9. doi: 10.21675/2357-707X.2013.v4.n2.528
29. Devaney L, Henchion M. Who is a Delphi 'expert'? Reflections on a bioeconomy expert selection procedure

- from Ireland. *Futures.* 2018 May;99:45-55. doi: 10.1016/j.futures.2018.03.017
30. Corry M, Porter S, McKenna H. The redundancy of positivism as a paradigm for nursing research. *Nurs Philos.* 2019;20(1):e12230. doi: 10.1111/nup.12230
31. Brandão MAG, Barros ALBL, Primo CC, Bispo GS, Lopes ROP. Nursing theories in the conceptual expansion of good practices in nursing. *Rev Bras Enferm.* 2019 Mar/Apr;72(2):577-81. doi: 10.1590/0034-7167-2018-0395

Contribución de los autores:

Concepción y dibujo de la pesquisa: Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Marcos Antônio Gomes Brandão. **Obtención de datos:** Manuela Campos Gomes Borel. **Análisis e interpretación de los datos:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Marcos Antônio Gomes Brandão. **Redacción del manuscrito:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Maria Miriam Lima Da Nóbrega, Cristina Arreguy-Sena, Marcos Antônio Gomes Brandão. **Revisión crítica del manuscrito en cuanto al contenido intelectual importante:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Maria Miriam Lima Da Nóbrega, Cristina Arreguy-Sena, Marcos Antônio Gomes Brandão

Todos los autores aprobaron la versión final del texto.

Conflicto de intereses: los autores han declarado que no existe ningún conflicto de intereses.

Recibido: 10.02.2020

Aceptado: 04.07.2020

Editora Asociada:
Maria Lúcia do Carmo Cruz Robazzi

Copyright © 2021 Revista Latino-Americana de Enfermagem


Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia Creative Commons CC BY.

Esta licencia permite a otros distribuir, mezclar, ajustar y construir a partir de su obra, incluso con fines comerciales, siempre que le sea reconocida la autoría de la creación original. Esta es la licencia más servicial de las ofrecidas. Recomendada para una máxima difusión y utilización de los materiales sujetos a la licencia.

Autor de correspondencia:

Marcos Antônio Gomes Brandão

E-mail: marcosantoniogbrandao@gmail.com

 <https://orcid.org/0000-0002-8368-8343>