

Diretriz para incorporação do método Delphi na avaliação de teorias de enfermagem


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: descrever uma diretriz de uso do método Delphi para avaliação de teorias de enfermagem, na perspectiva de validação interna. **Método:** estudo metodológico, dirigido ao desenvolvimento de uma diretriz para o uso do método Delphi na avaliação de teorias de enfermagem. **Resultados:** o método Delphi, os princípios da sabedoria coletiva e os níveis de proficiência são utilizados na produção de uma diretriz para organização, busca, seleção e coordenação de atividades de avaliadores teóricos em equipes. Distingue três fases para o processo de avaliação teórica: fase preparatória (FP); fase intermediária (FI) e fase de avaliação da teoria (AT), incorporando procedimentos de seleção do tipo de Delphi; busca, seleção e classificação de juízes/avaliadores da teoria; definição de critérios para a realização de rodadas e de manutenção ou descarte de unidades da teoria avaliada. **Conclusão:** a diretriz elaborada foi capaz de adequar os elementos do método Delphi como uma estratégia favorável à validação interna de teorias de enfermagem.

Descritores: Conhecimento; Modelos Teóricos; Teoria de Enfermagem; Estudo de Validação; Tomada de Decisão; Enfermagem.





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.

Como citar este artigo

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>.

Introdução

A construção teórica pela estratégia teoria-pesquisa é um processo iniciado na elaboração de uma teoria explícita nas fases de desenvolvimento conceitual e operacionalização, posteriormente, necessitando avançar para as fases de confirmação ou não confirmação, até alcançar as condições de aplicação prática⁽¹⁾. Nas disciplinas aplicadas, a validação empírica é enfatizada para produzir o julgamento sobre a utilidade e a aplicação.

O empreendimento total da construção teórica exige a realização de avaliações para estimar o valor de uma “boa” teoria, que é aquela rica em “virtudes teóricas”, como apontam filósofos da ciência. Dentre essas virtudes, podemos destacar a observância de singularidade, falseamento, parcimônia, predição, explicação, conservacionismo, capacidade de generalização, fecundidade, consistência interna, riqueza empírica e abstração⁽²⁾.

Na identificação da boa teoria, a validação pode ser orientada pela perspectiva interna, a avaliação dos elementos intrínsecos ou pela perspectiva externa, a avaliação mediante o teste empírico⁽³⁾. Na enfermagem, a validação interna é comumente denominada de “avaliação de teoria” (*theory evaluation*) que visa determinar a adequação do seu uso e a abordagem epistemológica⁽⁴⁾. A depender do critério formal a ser usado, a avaliação pode incorporar a análise ou a decomposição teórica.

Apesar da relevância da avaliação de teoria e da existência de dezenas de critérios estruturados e sistemáticos para a sua realização, ainda é incomum verificar na literatura internacional a aplicação desses critérios. São mais comuns as validações externas em estudos empíricos com análises estatísticas ou revisões da literatura⁽⁴⁾.

Há um contínuo interesse na produção de teorias de médio alcance para melhor descrever, explicar, predizer ou prescrever os fenômenos, os fatos, os eventos ou as intervenções com os quais a enfermagem lida no cotidiano. Entretanto, são essas teorias, junto às de situação específica as mais raramente avaliadas⁽⁴⁾.

Se por um lado esse movimento contínuo contribui para o avanço da disciplina, por outro, ele exige a disponibilização de instrumentos e diretrizes que promovam boas práticas de desenvolvimento teórico. Portanto, são indispensáveis o acesso aos recursos que possam auxiliar no programa de elaboração, validação, refinamento e aplicação teórica.

A validação interna (avaliação) de uma teoria exige ao mesmo tempo o conhecimento da teoria e um elevado nível de conhecimento metateórico. Por conta disso, é difícil encontrar metateoristas experientes e disponíveis para realizar essa tarefa. Quando possível, é complexa

a coordenação da tarefa ou, até mesmo, a identificação de um analista considerado proficiente, pela aplicação de critérios epistemologicamente consistentes.

Diante da dificuldade de localizar peritos metateóricos, os princípios e critérios da sabedoria coletiva ou sabedoria da multidão podem ser úteis para a construção de diretrizes, métodos ou técnicas que orientem a formação de uma equipe, capaz de desenvolver a tarefa de avaliação teórica com igual ou superior resultado, comparado ao de um único perito metateórico. Na teoria da sabedoria da multidão, critérios como a independência, a descentralização, a diversidade e a agregação orientariam a constituição de grupos, em que a decisão agregada superaria a do especialista, em separado⁽⁵⁾.

Desse modo, analistas, não obrigatoriamente peritos em metateoria, funcionam como juízes do conteúdo, da estrutura e de outros critérios a serem avaliados. Do julgamento agregado são alcançados resultados consistentes que permitem que a avaliação da teoria seja levada a cabo com sucesso. Entretanto, diretrizes, métodos ou técnicas com essa conformação não estão disponíveis para uso com teorias de enfermagem.

Presumivelmente, o Delphi é um método adequado para conduzir a avaliação de uma teoria de enfermagem, suportado por critérios da sabedoria da multidão, demonstrando que grupos podem julgar, adequadamente, sob condições de incerteza, definindo conceitos fundamentais, julgando e agregando o valor coletivo das ideias⁽⁵⁻⁷⁾. Tem sido usado para lidar com questões não esclarecidas por abordagens experimentais, nas quais a opinião de um grupo tem valor para clarificá-las, portanto, sendo compatíveis com a validação interna⁽⁶⁾.

Entretanto, sua aplicação para essa finalidade é escassa. Identificou-se na literatura o seu uso, somente em uma teoria da interface educação-informática na avaliação dos critérios de importância, precisão e clareza, parcimônia ou simplicidade, compreensão, operacionalização, validade empírica, frutificação e aplicação⁽⁸⁾. A descrição metodológica no mencionado estudo não fornece suficientes elementos para seu uso na avaliação de teorias de enfermagem com os critérios formais, usualmente aplicados na disciplina⁽⁴⁾.

No Brasil, o método Delphi tem auxiliado na abordagem de problemas empíricos como indicação de tendências, obtenção de consenso sobre programa ou intervenção, opinião de especialistas para comparação de tratamentos e, mais extensamente, na construção de instrumentos avaliativos e criação e validação de instrumentos⁽⁹⁾. A adaptação do método para a avaliação de teorias de enfermagem segue como potencial ainda não explorado, mesmo diante de seu caráter inovador. O presente artigo foi elaborado frente à escassez de

pesquisas e as potencialidades advindas da elaboração de uma diretriz para lidar com o desafio apresentado.

O artigo tem como objetivo descrever uma diretriz de uso do método Delphi para avaliação de teorias de enfermagem, na perspectiva de validação interna.

Método

Estudo metodológico de desenvolvimento de uma diretriz para o uso do método Delphi na avaliação de teorias de enfermagem, indicando procedimentos de organização, busca, seleção e coordenação de atividades de avaliadores teóricos em equipes. Foram bases de referência os critérios da sabedoria coletiva e os níveis de proficiência⁽⁵⁾ e sua elaboração deu-se no Rio de Janeiro, RJ, Brasil, entre os meses de novembro e dezembro de 2019.

Foram incorporados os elementos usados em referenciais metodológicos de concepção, construção e testagem de diretrizes, destacando-se: seleção do tópico e escopo; adaptação de protótipo de diretriz de estratégia de avaliação teórica, usando o método Delphi; formação de grupo para desenvolvimento; busca sistemática de evidências; análise e síntese de evidências disponíveis e criação da recomendação⁽¹⁰⁾.

Os procedimentos específicos de desenvolvimento da diretriz foram: revisão simples da literatura de manuscritos do uso do método Delphi na avaliação de teorias e nas demais aplicações; interpretação dos métodos de avaliação de teorias de enfermagem^(4,11-12); seleção de material complementar do tema da sabedoria coletiva; compilação e interpretação dos resultados de uso de um protótipo de uma diretriz de avaliação teórica, desenvolvida em dissertação de mestrado de uma das autoras, incorporando características do método Delphi;

elaboração da diretriz, levando em conta os princípios de construção de diretrizes na saúde e as necessárias adaptações ao objeto teórico-filosófico; discussão e revisão pelos autores; elaboração final da diretriz com diagramação.

O protótipo desenvolvido na dissertação de mestrado possuía as seguintes etapas: (a) seleção dos *experts*; (b) contato com peritos e convite para participação dos selecionados; (c) envio do instrumento, por meio eletrônico, aos que aceitaram participar; (d) apreciação de itens de avaliação da teoria baseado em escala Likert de concordância; (e) recebimento das respostas; (f) análise qualitativa e quantitativa dos resultados; (g) adaptação do conteúdo para nova rodada de avaliação teórica; (h) reenvio com *feedback* contendo os dados que levaram à modificação ou manutenção dos itens para realizar nova avaliação; (i) recebimento das respostas do instrumento adaptado; (j) análise dos dados da segunda versão; (l) construção final pelo consenso; (m) revisão gramatical e ortográfica e (n) encerramento da avaliação teórica.

A dissertação de mestrado que incorporou o uso do protótipo avaliou a Teoria dos Vínculos Profissionais⁽¹³⁾ pela estratégia de avaliação teórica de Meleis⁽¹⁴⁾. O estudo que aplicou o protótipo da diretriz respeitou os princípios éticos de pesquisa contidos na Resolução 466/2012 do Conselho Nacional de Saúde, obtendo parecer de aprovação em Comitê de Ética em Pesquisa, sob o número 3.237.583.

Resultados

Englobando o método Delphi na avaliação de teorias de enfermagem, a diretriz possui três fases: fase preparatória (FP), fase intermediária (FI) e fase de avaliação da teoria (AT). O presente estudo detalha a fase intermediária como esquematizada na Figura 1.



Figura 1 - Esquematização da diretriz para incorporação do método Delphi na avaliação de teorias de enfermagem. Rio de Janeiro, RJ, Brasil, 2019

Na fase preparatória (FP) decide-se pela teoria a ser avaliada e seleciona-se a estratégia de avaliação a ser empregada dentre as alternativas disponíveis na literatura.

Na fase intermediária, nove procedimentos ligados ao uso do método Delphi são delineados. O primeiro procedimento é referente ao tipo de Delphi a ser usado, influenciado pelo nível da teoria a ser avaliada e por sua maturidade de aplicação.

No segundo procedimento, define-se o papel de coordenação da avaliação teórica que pode ser acumulado com a condição de avaliador primário. Esse pode ser exercido por membro da equipe de desenvolvimento da teoria ou por outro pesquisador com consistente conhecimento sobre ela, responsabilizando-se pela preparação e condução do processo de avaliação. A função do avaliador primário é fornecer ajuda aos avaliadores secundários no acesso aos materiais de análise, podendo produzir relatório preliminar com resultados da tarefa realizada que serve como um documento básico destinado à consulta pelos demais avaliadores. De outro modo, os secundários operam como juízes a partir do documento básico ou exercendo função avaliativa primária.

O terceiro procedimento é a localização e a seleção dos candidatos a avaliadores/juízes. A busca pode ser em plataformas que contenham currículos eletrônicos, sendo filtros: a nacionalidade, a formação acadêmica, o idioma e a atuação profissional, dentre outros necessários. Recomenda-se a apreciação de publicações, projetos ou pesquisas e a verificação de tempo de atuação na área para a escolha criteriosa do candidato a avaliador. A seleção de candidatos a avaliadores secundários geralmente é difícil, na inexistência de critérios classificatórios de uma presumível perícia metateórica.

Assim, tendo por base a sabedoria coletiva⁽⁵⁾, as equipes de avaliadores (juízes) devem possuir diversidade no nível de perícia. Para estimar a perícia dos juízes propõem-se critérios em cinco domínios: formação educacional em teorias de enfermagem; experiência profissional na área da teoria; experiência e conhecimento metateórico; difusão de conhecimento produzido no tema de metateorias ou teorias de enfermagem e reconhecimento da *expertise* no tema de metateorias ou teorias de enfermagem, por pares (Figura 2).

Domínios e critérios para categorização dos avaliadores/juízes da teoria	Escore
1. Formação educacional em teorias de enfermagem:	
a) Doutorado em Enfermagem ou áreas afins e tese teórica ou metateórica	4 pontos
b) Mestrado em Enfermagem ou áreas afins e tese teórica ou metateórica	3 pontos
c) Doutorado em enfermagem com tese em outros temas	2 pontos
2. Experiência profissional na área da teoria a ser avaliada:	
(a) Mais de quatro anos de experiência no ensino, pesquisa ou prática clínica na área/tema da teoria a ser avaliada e, pelo menos, um ano no ensino de teorias ou pesquisa teórica ou metateórica de enfermagem	4 pontos
b) Mais de quatro anos de experiência no ensino, pesquisa ou prática clínica na área/tema da teoria a ser avaliada	3 pontos
c) Entre dois a quatro anos de experiência no ensino, pesquisa ou prática clínica na área/tema da teoria a ser avaliada ou, pelo menos, um ano no ensino de teorias ou pesquisa teórica ou metateórica de enfermagem	2 pontos
3. Experiência e conhecimento metateórico:	
a) Elaboração ou orientação de mais de uma teoria ou metateoria de enfermagem	4 pontos
b) Elaboração ou orientação de uma teoria ou metateoria de enfermagem	3 pontos
c) Elaboração ou orientação de, pelo menos, uma teoria ou metateoria não-enfermagem	2 pontos
4. Difusão de conhecimento produzido no tema de metateorias ou teorias de enfermagem:	
a) Autoria em mais de um artigo publicado em periódico indexado de padrão internacional de alto fator de impacto (Para Brasil: Qualis A1) em tema de teoria ou metateoria de enfermagem	4 pontos
b) Autoria de um artigo publicado em periódico indexado de padrão internacional de alto fator de impacto (Para Brasil: Qualis A2 a A4) em tema de teoria ou metateoria de enfermagem	3 pontos
c) Autoria de, pelo menos, um artigo publicado em periódico indexado de padrão nacional com médio fator de impacto (Para Brasil: Qualis B1) em tema de teoria ou metateoria de enfermagem	2 pontos
5. Reconhecimento da <i>expertise</i> no tema de metateorias ou teorias de enfermagem, por pares:	
a) Mais de uma participação como convidado (palestrante, conferencista, comentarista, professor em curso ou minicurso) em evento científico para ministrar tema ligado a teorias ou metateorias de enfermagem	4 pontos
b) Uma participação como convidado (palestrante, conferencista, comentarista, professor em curso ou minicurso) em evento para ministrar tema ligado a teorias ou metateorias de enfermagem	3 pontos
c) Participação como ouvinte/participante/aluno em evento ou curso concluído sobre teorias ou metateorias de enfermagem	2 pontos

Figura 2 – Critérios para classificação do nível de perícia presumida de avaliadores de teorias de enfermagem. Rio de Janeiro, RJ, Brasil, 2019

O candidato a avaliador/juiz tem seu nível de perícia hierarquizado pela pontuação obtida pelo somatório dos itens dos cinco domínios. A pontuação máxima para cada domínio é de 4 pontos e a mínima de 2. Quando o avaliador não atende a qualquer dos critérios de um domínio, recebe

pontuação 0. A pontuação total varia de 0 a 20, sendo 4 pontos o valor mínimo arbitrado para considerar o candidato apto para compor a equipe de avaliação teórica.

O nível de perícia presumida de um candidato a avaliador é estabelecido pelo escore obtido da avaliação

dos cinco domínios. Cinco níveis são propostos: iniciante, iniciante avançado, competente, proficiente e perito.

A classificação do nível de perícia presumida de um candidato é estabelecida a partir de sua pontuação total obtida (ver Figura 1).

A definição do número mínimo de avaliadores na equipe dependerá do somatório dos pontos individuais de cada avaliador e da composição mista que garanta membros de, ao menos, dois níveis diferentes de perícia (Figura 3).

Nível de perícia presumida do avaliador	Pontuação requerida para adequar-se ao nível	Crítérios para definição da equipe, em função dos pontos de perícia
Iniciante	somatório mínimo de 4 pontos	(a) Equipe de dois avaliadores: soma dos pontos dos avaliadores deve ser de no mínimo 35 pontos
Iniciante avançado	somatório entre 5 e 10 pontos	(b) Equipe de três avaliadores: soma dos pontos dos avaliadores deve ser de no mínimo 36 pontos
Competente	somatório entre 11 pontos e 14 pontos	(c) Equipe de quatro avaliadores: soma dos pontos dos avaliadores deve ser de no mínimo 48 pontos
Proficiente	somatório de pontuação entre 15 e 17 pontos	(d) Equipe de cinco avaliadores: soma dos pontos dos avaliadores deve ser de no mínimo 60 pontos
Perito	somatório de pontuação maior ou igual a 18 pontos	(e) Equipe de seis ou mais avaliadores: aplicar a seguinte equação:

Fonte: Brandão, MA, 2019.

Figura 3 - Classificação do nível de perícia presumida dos avaliadores de acordo com o escore total obtido e definição do número de avaliadores da equipe. Rio de Janeiro, RJ, Brasil, 2019

Por exemplo, pela aplicação da equação proposta, a composição de uma equipe de sete avaliadores pelo cálculo dos pontos totais requeridos do somatório dos avaliadores irá exigir, aproximadamente, 74 pontos. Exemplificando, uma configuração adequada incluiria: seis avaliadores iniciantes avançados de 10 pontos cada e um avaliador competente, com 14 pontos. Obviamente, podem ser aplicadas outras configurações que respeitem a pontuação mínima para a equipe. Além da pontuação total, a exigência de equipes mistas em termos de níveis de perícia objetiva garantir o critério de diversidade de juízes.

Conhecendo-se que são comuns as perdas diante dos convites e durante as rodadas iniciais do Delphi, é recomendável a seleção de número superior de juízes necessários para a composição mínima das equipes, buscando-se manter proporcionalidade entre os níveis.

O quarto procedimento é o convite aos candidatos selecionados, sendo baseado em comunicação por via escrita ou eletrônica e respeitando princípios éticos de pesquisa.

O quinto procedimento envolve a reaplicação dos critérios de categorização dos avaliadores/juízes pelo nível de perícia e ajustes à composição nas equipes, considerando recusas aos convites de participação.

O sexto procedimento é o planejamento de rodadas e critérios de interrupção. Esse planejamento considera o nível de abstração, o número de conceitos e a complexidade da teoria a ser avaliada. Destaca-se, também, o número de juízes que desenvolvem as avaliações e, também, a consideração em relação à quantidade de critérios a serem avaliados na teoria. A interrupção das rodadas de avaliação deverá ser amparada no julgamento explícito dos avaliadores com alcance de

consenso avaliativo ou constituição de multiplicidade de ideias no dissenso. Outra decisão é estabelecer ou não, *a priori*, um máximo de rodadas. Essa decisão refere-se mais a uma variável tempo disponível para avaliação do que ao julgamento avaliativo.

Posteriormente, para orientar os avaliadores, devem ser estabelecidos critérios explícitos para descarte de itens em cada rodada. Os itens de uma teoria sujeitos à avaliação são seus componentes como conceitos, pressupostos, suposições, afirmativas e esquemas modelares. Portanto, os avaliadores devem ter clareza de que a decisão pela exclusão é dirigida pela estratégia selecionada e não, apenas, por suas opiniões pessoais, livremente emitidas. Sua função é julgar um dado item da teoria em função de critérios de avaliação estabelecidos na estratégia.

O oitavo procedimento engloba a definição do alcance do consenso e da estabilidade das respostas. O consenso pode ser verificado por medidas formais de concordância, medidas de tendência central, percentagem de concordância, medida de tendência central dentro de um intervalo específico, dentre outras⁽¹⁵⁾. Pode ser planejado o emprego de uma escala Likert de cinco pontos com duas finalidades: (1) verificar a concordância de avaliadores secundários com o resultado da avaliação primária ou (2) organizar as avaliações secundárias em assertivas que serão submetidas à análise de concordância em rodada posterior.

Mesmo aplicando-se escalas é recomendável garantir campos de livre edição para que avaliadores/juízes possam expressar suas sugestões, recomendações e apreciações detalhadas.

O nono procedimento é o fornecimento de orientações específicas acerca da estratégia de avaliação

teórica. Quando o coordenador ou avaliador primário julgar necessário, podem ser realizadas capacitações complementares e específicas do conteúdo da estratégia.

Discussão

A avaliação teórica é capaz de fornecer elementos sobre uma “boa” teoria, com diversos critérios formais e sistemáticos disponíveis na literatura^(2,11). Porém, nem sempre estão disponíveis os recursos humanos com competência e conhecimento requerido para desenvolver, apropriadamente, o processo de julgar as virtudes teóricas. E, nesse aspecto, o Delphi, ao empregar o princípio da sabedoria coletiva pelo consenso ou dissenso, pode multiplicar a perícia do grupo, ainda, ampliando o universo de alternativas para a avaliação⁽¹⁵⁻¹⁶⁾. De igual forma, auxilia na coordenação do trabalho de avaliação.

Pela avaliação são percebidas relações e vínculos de conceitos, permitindo ao revisor a verificação de pontos fortes e limitações da teoria; identificação da necessidade de novos elementos da teoria ou de aperfeiçoamento dos já existentes e, como meta final, a determinação da contribuição potencial da teoria avaliada para o conhecimento científico⁽¹¹⁾.

Ao contrário da análise teórica (*theory analysis*) que decompõe uma teoria para exame de suas partes ou componentes⁽⁴⁾, a avaliação teórica também os julga. Contudo, mesmo uma teoria julgada como “boa”, pode se mostrar inadequada em seu valor descritivo, explicativo, preditivo ou prescritivo a partir da confirmação ou da aplicação. Isso coloca a validação interna como etapa relevante, porém, não terminal de um programa de desenvolvimento teórico.

Teorias que violem as virtudes da “boa” teoria são mais difíceis de refutar e tendem a não contribuir, de fato, com o conhecimento⁽¹⁷⁾, pois a inadequação de elementos e construções torna difícil a avaliação e testagem teórica. Assim, tem fundamental importância planejar a validação interna e/ou externa como parte de um programa mais abrangente. O uso da diretriz apresentada pode evitar dispêndio de recursos, ao colaborar na identificação de teorias que não disponham de virtudes suficientes que sustentem uma validação por pesquisa de campo.

São incertos os motivos do uso reduzido de estratégias de avaliação de teorias de enfermagem por critérios sistemáticos formais⁽⁴⁾. Contudo, pode-se atribuir influência à dificuldade de obtenção de avaliadores com suficiente autoridade epistêmica para julgar os itens metateóricos da validação interna. Supõe-se que as estratégias vinculadas à sabedoria coletiva possam contornar este problema da dependência do “perito” com vantagens substanciais⁽¹⁸⁾.

O método Delphi fundamenta-se nas suposições de John Dewey, enfatizando comunicação anônima entre indivíduos com *expertise* em dado tópico, com a meta de buscar a opinião de peritos de um modo iterativo e estruturado e, usualmente, buscando atingir um posicionamento consensual^(15,19). A liberdade e a observância das opiniões pessoais dos juizes garantem o critério de independência da sabedoria coletiva⁽⁵⁾.

Quanto ao uso nas pesquisas, embora seja utilizado predominantemente nas mistas e nas quantitativas, tem sua aplicação qualitativa e até mesmo na construção de teorias práticas, no contexto de organização da comunidade⁽¹⁵⁾. A avaliação de uma teoria é um empreendimento qualitativo permeado pela subjetividade e por normas, condutas e códigos do avaliador⁽⁸⁾.

O método Delphi pode coordenar estas características qualitativas do processo de avaliação, lidando com variáveis pessoais do critério de independência, retirando o máximo do trabalho em grupo. Pode ser usado para interpretação, predições e obtenção de recomendações da avaliação desenvolvida⁽⁸⁾.

Na escolha do método Delphi, a abordagem mais comum é a tradicional, também denominada normativa ou de consenso. Ela tem por meta reduzir a variância das estimativas e vieses entre peritos. Já o *Policy Delphi* ou de dissenso, busca obter uma ampla gama de opiniões, mas sem almejar um consenso⁽¹⁶⁾.

Para a avaliação teórica, o método Delphi de consenso é a indicação mais provável; porém, o emprego do dissenso pode ser recomendado para teorias de alta originalidade, densidade conceitual, complexidade e abstração teórica ou quando for difícil a determinação dos critérios de consenso. Também, pode ser uma meta da avaliação explorar o contraditório na produção de definições ou proposições teóricas.

Em relação às características da teoria, o Delphi de consenso pode ser indicado para as de micro ou de médio alcance com conceitualização descrita em mais de um estudo empírico ou para avaliar uma teoria parcialmente difundida, testada ou usada.

Supostamente, para o Delphi de consenso a composição de equipes com um número elevado de avaliadores só se justifica quando for difícil obter avaliadores com níveis mais altos de perícia, pois em grupos de muitos componentes é desafiador obter consenso. Por outro lado, presume-se que a abordagem de dissenso beneficie-se da composição de equipes maiores e com ampla diversidade de níveis de proficiência, tendendo a ampliar o debate de perspectivas diferentes e trazer elementos originais, diferindo da teoria original e da avaliação primária.

Painéis com mais participantes tendem a ter menores taxas de resposta, estimando-se redução de 0,08 pontos

percentuais por cada participante adicionado⁽²⁰⁾. São indicados de 5 a 20 peritos se for recomendação baseada exclusivamente nas características do método Delphi⁽²⁰⁾. Estudos sobre o desenvolvimento e a aplicação de *core outcome set* (COS) têm utilizado o Delphi para determinar quais resultados medir e neles predominam painéis Delphi de até 50 pessoas⁽²⁰⁾.

Na avaliação teórica, é desafiante estabelecer um número mínimo e máximo de avaliadores/juízes, por conta de seu caráter filosófico e natureza epistemológica abstrata inerente à teorização. Por exemplo, para teorias novas ou com pouca difusão pode ser difícil dispor de muitos avaliadores secundários com perícia adequada. Por outro lado, grandes equipes de avaliadores iniciantes podem não dispor de conhecimentos de natureza metateórica, causando uma dispersão de perspectivas que dificultaria a agregação de ideias. Nesse caso, a diretriz busca contornar os limites combinando um equilíbrio entre o critério de diversidade do princípio da sabedoria coletiva e a perícia necessária à avaliação teórica⁽⁴⁻⁵⁾.

No protótipo da diretriz foram incorporados quatro avaliadores de três níveis de perícia diferentes, sendo três avaliadores secundários que em conjunto alcançaram 36 pontos (14, 13 e 9 pontos individuais). Os critérios para definição da equipe, em função dos pontos de perícia, foram úteis para a composição deste grupo de tamanho reduzido, assim como a configuração de menos participantes garantiu a máxima taxa de respostas, como esperado para este tamanho de painel⁽²⁰⁾. As diferenças de níveis e estórias de formação dos avaliadores garantiram o critério de descentralização⁽⁵⁾. No entanto, sempre que possível, recomenda-se formar equipes com cinco ou mais juízes.

O modelo de Patricia Benner⁽²¹⁻²²⁾ com seus cinco níveis de aquisição de competência foi a base para a criação dos critérios de perícia dos juízes na diretriz e buscou reconhecer a experiência profissional como componente essencial para a validação. A ampla difusão dos estudos desses autores e seus critérios auxiliaram na definição. Existem modelos mais complexos de regras de agregação para definir a composição da equipe, como por exemplo, o *Contribution Weighted Model* (CWM) que pondera os prognósticos baseando-se na performance relativa de cada juiz e na acurácia do grupo^(18,23).

As contribuições de avaliadores/juízes são carregadas de conhecimentos, experiências e pontos de vistas particulares na avaliação da teoria. O processo iterativo do método Delphi pode permitir que nas rodadas os critérios obscuros da avaliação possam ser clarificados ou modificados, por meio de uma cuidadosa interpretação das respostas dos avaliadores secundários, por parte do coordenador. A composição de equipes com diferentes

níveis de competência garante o critério de diversidade da sabedoria coletiva⁽⁵⁾.

Comumente, os estudos aplicam de duas a três rodadas para o método Delphi⁽¹⁹⁾. Porém, os múltiplos critérios a avaliar, o alto número e a diversidade de perfis dos avaliadores podem demandar mais rodadas para o alcance do consenso. É desejável planejar um mínimo em função da quantidade de avaliadores, para garantir que não recaia sobre o coordenador do Delphi um esforço excessivo de gerenciamento dos resultados da avaliação, pondo em risco a qualidade do trabalho.

O nível de abrangência de teorias pode influenciar a definição dos critérios a serem avaliados por juízes; por exemplo, quando uma dada teoria de médio alcance for avaliada como um modelo, podem ser utilizados critérios ainda mais específicos e de natureza empírica⁽¹²⁾. Contudo, isso não interfere, diretamente, com a natureza do método Delphi como estratégia.

A decisão pelo alcance do consenso entre juízes é um tipo de mecanismo para atender ao critério de agregação da sabedoria coletiva, transformando os julgamentos individuais em uma decisão da equipe⁽⁵⁾. A decisão consensual pode partir da própria opinião dos avaliadores de que um consenso foi atingido; entretanto, é recomendável que isso não ocorra automaticamente ao fim da realização da técnica Delphi⁽¹⁹⁾.

Há que se especificar, claramente, quais são as condições para o alcance do consenso, quando a decisão for qualitativa. Quando são adotados procedimentos de mensuração quantitativa, deve-se estabelecer as medidas e os pontos de corte que serão usados para determinar o grau de concordância ou discordância, compatíveis com o consenso ou o dissenso⁽¹⁹⁾.

Não há regras mandatórias para o estabelecimento de consenso, porém a escala Likert de cinco pontos é a mais usual entre as escalas utilizadas para estimar discordâncias ou concordâncias⁽²⁴⁾. Oportuniza a verificação do grau de concordância de cada um dos itens ou do conjunto. Para estimativas de alcance de consenso com uso da Likert, podem ser aplicadas medidas formais de concordância como a estatística Kappa, para verificar o julgamento concordante dos juízes sobre os elementos da teoria.

Usualmente, os percentuais de concordância adotam como ponto de corte mínimo o valor de 0,8 ou 80%⁽¹⁵⁾. Entretanto, o pesquisador pode considerar outros pontos de corte, sustentado em evidência ou em recomendação consistente. Indica-se o pronunciamento explícito sobre o alcance do consenso com a indicação dos motivos que foram considerados na tomada de decisão.

Outra medida útil que pode ser utilizada sobre os dados obtidos pela escala Likert é o coeficiente de validade de conteúdo. O coeficiente de Aiken e sua

tabela de probabilidade caudal podem ser usados para indicar a validade de um item em particular avaliado por vários juízes, estimando um consenso. Também pode ser aplicado para julgar a validade, por um único juiz, do conteúdo de todos os itens da teoria. O intervalo do coeficiente varia de 0 a 1, sendo que valores mais altos indicam a validação⁽²⁵⁾.

Apesar da literatura, de forma geral, recomendar a utilização de escalas quantitativas para sinalizar o consenso, as justificativas qualitativas devem ser adicionadas, principalmente quando a recomendação for pela não validade do item^(24,26). A simples exclusão de um item pode tornar incoerente ou ilógica toda a teoria. Isso porque unidades de uma teoria desempenham papéis e possuem relevância diferente na estrutura teórica.

Por exemplo, a exclusão de um pressuposto pode descaracterizar a teoria como um todo, visto que esta tipologia de elemento funciona como premissas não dadas ao teste empírico. Assim, a sua retirada nega as ideias que guiaram o próprio teorista na construção da teoria. Por outro lado, as unidades do tipo “proposições” são apresentadas, justamente, para gerar hipóteses de testagem em estudos de validação empírica; portanto estão, naturalmente, sujeitas a exclusão ou a manutenção após as evidências obtidas da experimentação ou da pesquisa de campo⁽¹⁾.

Destaca-se que a partir da avaliação da Teoria dos Vínculos Profissionais⁽²⁷⁻²⁸⁾ fatores emergentes exigiram modificações de critérios não detalhados quando da elaboração do protótipo, o que colaborou com o aprofundamento para a criação da diretriz apresentada neste artigo.

O estudo não está livre de limitação. O foco de qualquer pesquisa que utilize o Delphi será sempre obter respostas de alta qualidade de uma seleção de indivíduos peritos⁽²⁹⁾. Contudo, a validação interna de uma teoria lida com critérios teórico-filosóficos que podem tornar difícil, para um avaliador secundário, produzir ou julgar a qualidade das respostas, pela natureza do objeto avaliado e pela propriedade de julgamento a ser realizado. Por exemplo, a definição conceitual é um dos elementos de uma teoria, avaliado na semântica, lógica e contexto⁽¹⁴⁾. Notadamente pode ser difícil fazer um julgamento de “boa resposta” para um construto tão complexo, diante de tais propriedades.

A limitação da subjetividade do julgamento do juiz na avaliação de teorias deve ser confrontada com a raiz filosófica do teórico e do avaliador. Raízes crítico-sociais, hermenêuticas ou do novo pragmatismo tendem a lidar com maior fluidez frente as diferentes perspectivas, inclusive, explorando-as no consenso ou dissenso. Por outro lado o pós-positivismo, ao requerer maior objetividade da realidade, exige critérios mais

estáveis, generalizáveis ou de mensuração⁽³⁰⁾. Nessa última raiz filosófica, métodos como modelagem de equação estrutural, análise fatorial e regressões múltiplas podem ser a melhor escolha para a avaliação da teoria, obviamente com critérios mais aproximados da validação externa⁽⁴⁾.

Dentre as contribuições ao avanço do conhecimento científico, o estudo adiciona uma dimensão não explorada da incorporação de avaliadores de diferentes níveis de perícia metateórica na tarefa de avaliação teórica, inclusive incorporando diretrizes para as fases deste processo. Frente ao crescimento no desenvolvimento de teorias de enfermagem de médio e micro alcances e de situação específica, com a consequente formação de novos teóricos, a diretriz pode facilitar o processo de validação das novas teorias, formando uma sólida base de conhecimento disciplinar^(4,31).

Por outro lado, níveis mais elevados de perícia metateórica tendem a exigir longos anos de formação neste campo, sendo mais comum de serem verificados em meio acadêmico e entre pesquisadores sêniores. Ao explorar o princípio da diversidade de níveis de perícia para a avaliação teórica, o estudo incentiva a criação de equipes de diferentes níveis de perícia, promovendo a cooperação e a circulação do conhecimento dos envolvidos na avaliação.

Finalmente, a aplicação do método Delphi na avaliação de uma teoria de enfermagem deve ser claramente compreendida como diferente da busca de consenso sobre eventos, fenômenos, fatos, tecnologias, condutas ou quaisquer outros elementos fundamentalmente empíricos. No método Delphi empírico, a queda de um item pode ter implicação mínima; contudo, na teoria a avaliação sobre a inadequação de suposições ou conceitos centrais pode colocar toda a teoria na condição de inadequada. Obviamente a principal meta da avaliação é identificar uma “boa teoria”, o que implica julgar a adequação de seus componentes; porém, este procedimento deve ser realizado com extrema cautela pelos avaliadores, compreendendo que em uma teoria há hierarquia e relação entre os elementos.

Conclusão

A diretriz elaborada foi capaz de adaptar os elementos do método Delphi como recurso favorável à validação interna de teorias de enfermagem, potencializando-o com a incorporação de juízes com diferentes visões de mundo, experiências, conhecimento científico e criatividade. Os critérios apresentados na diretriz adequam e articulam os níveis de proficiência dos avaliadores com o princípio da sabedoria da multidão, servindo de guia para a seleção e a composição de equipes de juízes, bem como facilitam

a coordenação do trabalho de avaliação teórica. Por seu caráter inovador, a diretriz pode instrumentalizar metateóricos de enfermagem e, possivelmente, acelerar o processo de aplicação das teorias na prática.

O uso de um protótipo de diretriz, na avaliação de uma teoria de médio alcance de enfermagem, a Teoria dos Vínculos Profissionais, trouxe resultados satisfatórios que presumem sua exequibilidade e apontou caminhos para refinamento.

Entende-se ser importante que outros pesquisadores repliquem seu uso na avaliação de teorias de grande e micro alcances para futuras adequações e atualizações da diretriz, também, adotando estratégias de avaliação por critérios formais diferentes do utilizado no protótipo.


Referências

- 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
- 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
- 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>
- Im EO. The current status of theory evaluation in nursing. *J Adv Nurs.* 2015 May;71(10):2268-78. doi: 10.1111/jan.12698
- 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
- Fleener 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
- 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
- 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
- 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
- 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
- McEween M, Wills EM. *Theoretical Basis for Nursing.* 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2014.
- Risjord M. Middle-range theories as models: New criteria for analysis and evaluation. *Nurs Philos.* 2019;20(1):e12225. doi: 10.1111/nup.12225
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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 profesionales 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

Contribuição dos autores:

Concepção e desenho da pesquisa: Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Marcos Antônio Gomes Brandão. **Obtenção de dados:** Manuela Campos Gomes Borel. **Análise e interpretação dos dados:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Marcos Antônio Gomes Brandão. **Redação do manuscrito:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Maria Miriam Lima Da Nóbrega, Cristina Arreguy-Sena, Marcos Antônio Gomes Brandão.

Autor correspondente:
Marcos Antônio Gomes Brandão
E-mail: marcosantoniogbrandao@gmail.com
 <https://orcid.org/0000-0002-8368-8343>

Revisão crítica do manuscrito quanto ao conteúdo 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 os autores aprovaram a versão final do texto.

Conflito de interesse: os autores declararam que não há conflito de interesse.

Recebido: 10.02.2020

Aceito: 04.07.2020

Editora Associada:
Maria Lúcia do Carmo Cruz Robazzi

Copyright © 2021 Revista Latino-Americana de Enfermagem
Este é um artigo de acesso aberto distribuído sob os termos da Licença Creative Commons CC BY.

Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original. É a licença mais flexível de todas as licenças disponíveis. É recomendada para maximizar a disseminação e uso dos materiais licenciados.