

## Guideline for incorporating the Delphi method in the evaluation of nursing theories

Manuela Campos Gomes Borel<sup>1</sup>

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

Rafael Oliveira Pitta Lopes<sup>2</sup>

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

Maira Buss Thofehrn<sup>1</sup>

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


Maria Miriam Lima Da Nóbrega<sup>3</sup>

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

Cristina Arreguy-Sena<sup>1</sup>

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

Marcos Antônio Gomes Brandão<sup>4</sup>

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

**Objective:** to describe a guideline for the use of the Delphi method to evaluate nursing theories, from the perspective of internal validation. **Method:** a methodological study, targeted at the development of a guideline for the use of the Delphi method in the evaluation of nursing theories. **Results:** the Delphi method, principles of collective wisdom and levels of proficiency are used in the production of a guideline for organizing, searching, selecting and coordinating the activities of theoretical evaluators in teams. It distinguishes three phases for the theoretical evaluation process: Preparatory Phase (PP); Intermediate Phase (IP) and Theory Evaluation (TE) phase, incorporating Delphi-type selection procedures; search, selection and classification of judges/evaluators for the theory; definition of criteria for carrying out rounds and maintenance or removal of units of the theory evaluated. **Conclusion:** the developed guideline was able to adapt the elements of the Delphi method as a favorable strategy for the internal validation of nursing theories.

**Descriptors:** Knowledge; Models, Theoretical; Nursing Theory; Validation Study; Decision Making; Nursing.





<sup>1</sup> Universidade Federal de Juiz de Fora, Faculdade de Enfermagem, Juiz de Fora, MG, Brazil.

<sup>2</sup> Universidade Federal do Rio de Janeiro, Campus Macaé, Macaé, RJ, Brazil.

<sup>3</sup> Universidade Federal da Paraíba, Centro de Ciências da Saúde, João Pessoa, PB, Brazil.

<sup>4</sup> Universidade Federal do Rio de Janeiro, Escola de Enfermagem Anna Nery, Rio de Janeiro, RJ, Brazil.

### How to cite this article

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>.

## Introduction

The theoretical construction through the theory-research strategy is a process initiated in the elaboration of an explicit theory in the phases of conceptual development and operationalization, later on, needing to proceed to the confirmation or non-confirmation phases, until reaching practical application conditions<sup>(1)</sup>. In the applied disciplines, empirical validation is emphasized to produce a judgment about usefulness and application.

The total undertaking of theoretical construction requires judgments to be carried out to estimate the value of a "good" theory, which is that rich in "theoretical virtues, as science philosophers point out. Among these virtues, we can highlight the observance of singularity, falsification, parsimony, prediction, explanation, conservatism, capacity for generalization, fecundity, internal consistency, empirical wealth and abstraction<sup>(2)</sup>.

In identifying the good theory, an internal perspective can guide validation, the judgment of the intrinsic elements or by the external perspective, the judgment using the empirical test<sup>(3)</sup>. In nursing, internal validation is commonly referred to as "theory evaluation" which aims to determine the appropriateness of its use and the epistemological approach<sup>(4)</sup>. Depending on the formal criteria to be used, the evaluation can incorporate analysis or theoretical breakdown.

Despite the relevance of theory evaluation and the existence of dozens of structured and systematic criteria for its realization, it is still unusual to verify the application of these criteria in international literature. External validations are more common in empirical studies with statistical analyses or literature reviews<sup>(4)</sup>.

There is a continuous interest in the production of middle-range theories to describe better, explain, predict or prescribe the phenomena, facts, events or interventions with which nursing deals in daily life. However, these theories, together with those of a specific situation, are the most rarely evaluated<sup>(4)</sup>.

If, on the one hand, this continuous movement contributes to the progress of the discipline, on the other hand, it requires the availability of instruments and guidelines that promote good theoretical development practices. Therefore, access to resources that can assist in the program of elaboration, validation, refinement and theoretical application is indispensable.

The internal validation (evaluation) of a theory requires, at the same time knowledge of the theory and a high level of meta-theoretical knowledge. For such reason, it is difficult to find experienced and available meta-theorists to perform this task. Whenever possible, the coordination of this task is difficult or, even, is a complicated procedure the identification of

an analyst considered proficient by the application of epistemologically consistent criteria.

Given the difficulty in locating meta-theoretical experts, the principles and criteria of collective wisdom or crowd wisdom can be useful for the construction of guidelines, methods or techniques that guide the formation of a team, capable of developing the task of theoretical evaluation with the same or superior result, compared to that of a single meta-theoretical expert. In the crowd wisdom theory, criteria such as independence, decentralization, diversity and aggregation would guide the constitution of groups, in which the aggregate decision would surpass that of the specialist, separately<sup>(5)</sup>.

In this way, analysts, not necessarily experts in meta-theory, act as judges for the content, the structure and other criteria to be judged. From the aggregate judgment, consistent results are achieved that allow for the theory evaluation to be carried out successfully. However, guidelines, methods or techniques with this conformation are not available for use with nursing theories.

Presumably, the Delphi method is adequate to evaluate a nursing theory supported by the crowd wisdom criteria, demonstrating which groups can judge adequately under conditions of uncertainty, defining the fundamental concepts, judging and adding the collective value of the ideas<sup>(5-7)</sup>. It has been used to deal with issues not clarified by experimental approaches in which the opinion of a group has value to clarify them, therefore, being compatible with internal validation<sup>(6)</sup>.

However, its application for this purpose is scarce. Its use was identified in the literature only in a theory of the education-informatics interface in the evaluation of the criteria of importance, precision and clarity, parsimony or simplicity, understanding, operationalization, empirical validity, fruiting and application<sup>(8)</sup>. The methodological description in the study mentioned above does not provide enough elements for its use in the evaluation of nursing theories with formal criteria, usually applied in the discipline<sup>(4)</sup>.

In Brazil, the Delphi method has helped in addressing practical problems such as trend indication, obtaining consensus on a program or intervention, expert opinion for comparing treatments and, more widely, in the construction of tools for evaluation and in the creation and validation of instruments<sup>(9)</sup>. The adaptation of the method for nursing theories evaluation remains a potential that has not yet been explored, even given its innovative character. This article was prepared given the scarcity of research studies and the potential from the development of a guideline.

The article aims to describe a guideline for the use of the Delphi method to evaluate nursing theories, from the perspective of internal validation.

## Method

This research is a methodological study for developing a guideline for the use of the Delphi method in nursing theories evaluation, indicating procedures for organizing, searching, selecting and coordinating the activities of theoretical evaluators in teams. The criteria of collective wisdom and levels of proficiency<sup>(5)</sup> were the reference basis and its elaboration took place in Rio de Janeiro, RJ, Brazil, between the months of November and December 2019.

The elements used in the methodological frameworks for the design, construction and testing of guidelines were incorporated, highlighting the following: selecting the topic and scope; adapting a prototype of a theoretical evaluation strategy guideline, using the Delphi method; group formation for development; systematic search for evidence; analysis and synthesis of available evidence and elaboration of the recommendation<sup>(10)</sup>.

The specific procedures for developing the guideline were the following: a simple review of manuscripts on the use of the Delphi method in theories evaluation and other applications; interpreting nursing theory evaluation methods<sup>(4,11-12)</sup>; selection of the complementary material on the topic of collective wisdom; compiling and interpreting the results of using a prototype of a theoretical evaluation guideline developed in a master's thesis by one of the authors, incorporating features of the Delphi method; elaborating the guideline, taking into account the principles of construction of guidelines in health and the necessary adaptations to the theoretical-philosophical object; discussion and review by the authors; final elaboration of the guideline with diagramming interpretation of nursing theory evaluation methods.

The prototype developed in the master's thesis had the following stages: (a) selection of the *experts*; (b) contact with experts and invitation for participation by those selected; (c) electronically sending the instrument to those who agreed to participate; (d) appreciation of theory evaluation items based on an agreement Likert scale; (e) receiving the answers; (f) qualitative and quantitative analysis of the results; (g) adaptation of the content for a new round of theoretical evaluation; (h) forwarding with *feedback* containing the data that led to the modification or maintenance of the items to perform a new evaluation; (i) receiving the answers to the adapted instrument; (j) analysis of the data from the second version; (l) final construction by consensus; (m) grammatical and orthographic review and (n) closing the theoretical evaluation.

The master's dissertation that incorporated the use of the prototype evaluated the Theory of Professional Links<sup>(13)</sup> by Meleis' theoretical evaluation strategy<sup>(14)</sup>. The study that applied the prototype of the guideline respected the ethical principles of research contained in Resolution 466/2012 of the National Health Council, obtaining an approval opinion from the Research Ethics Committee, under number 3.237.583.

## Results

Encompassing the Delphi method in the nursing theories evaluation, the guideline has three phases: Preparatory Phase (PP), Intermediate Phase (IP) and Theory Evaluation (TE) phase. This study details the intermediate phase, as shown in Figure 1.

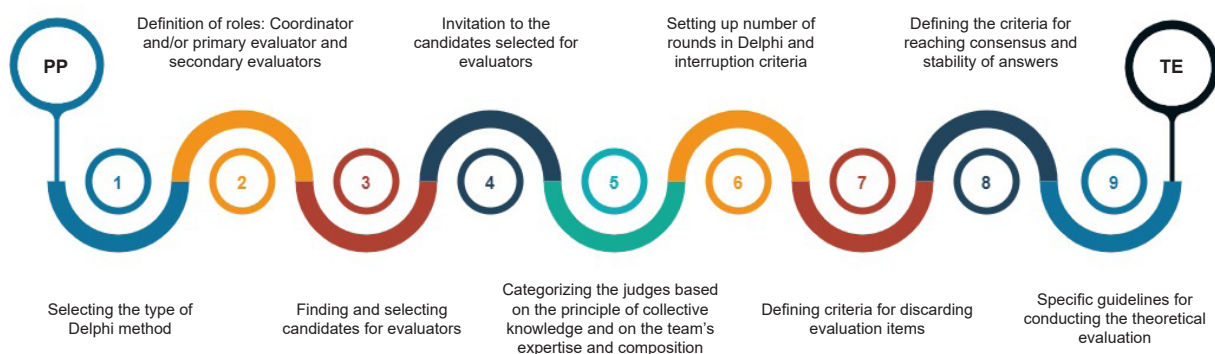


Figure 1 – Outline of the guideline for incorporating the Delphi method in nursing theories evaluation. Rio de Janeiro, RJ, Brazil, 2019

In the preparatory phase (PP) the theory to be evaluated is chosen and the strategy to be employed is selected from the alternatives available in the literature.

In the intermediate phase, nine procedures related to the use of the Delphi method are outlined. The first procedure is related to the type of Delphi to be used,

influenced by the level of the theory to be evaluated and by its application maturity.

In the second procedure, the coordination role of the theoretical evaluation is defined, which can be accumulated with the condition of a primary evaluator. This can be done by a member of the theory development team or by another researcher with consistent knowledge about it, being responsible for preparing and conducting the evaluation process. The role of the primary evaluator is to provide help to the secondary evaluators in accessing the analysis materials, with the ability to produce a preliminary report with results of the task performed that serves as a primary document intended for consultation by the other evaluators. Otherwise, the secondary operate as judges from the primary document or by performing a primary evaluative function.

The third procedure is the location and selection of candidates for evaluators/judges. The search can be performed in platforms that contain electronic resumes,

the following being used as filters: nationality, academic background, language and professional performance, among others necessary. Appraisal of publications, projects or research studies is recommended, as well as verification of the performance time in the area for the careful selection of the candidate for evaluator. The selection of candidates for secondary evaluators is usually difficult, in the absence of classification criteria for a presumed meta-theoretical expertise.

So, based on collective wisdom<sup>(5)</sup>, the teams of evaluators (judges) must have diversity in their expertise levels. The expertise of the judges is analyzed by criteria in five domains, namely: Educational training in nursing theories; Professional experience in the theoretical area; Meta-theoretical experience and knowledge; Dissemination of knowledge produced on the topic of nursing meta-theories or theories and Peer-recognition of expertise on the topic of nursing meta-theories or theories (Figure 2).

Domains and criteria for categorizing theory evaluators/judges	Score
1. Educational training in nursing theories:	
a) PhD in Nursing or related areas and theoretical or meta-theoretical thesis	4 points
b) Master's degree in Nursing or related areas and theoretical or meta-theoretical thesis	3 points
c) PhD in Nursing with a thesis on other topics	2 points
2. Professional experience in the area of the theory to be evaluated:	
(a) More than four years of experience in teaching, research or clinical practice in the area/theme of the theory to be evaluated and, at least, one year in teaching theories or theoretical or meta-theoretical nursing research	4 points
b) More than four years of experience in teaching, research or clinical practice in the area/theme of the theory to be evaluated	3 points
c) Between two and four years of experience in teaching, research or clinical practice in the area/theme of the theory to be evaluated or, at least, one year in teaching theories or theoretical or meta-theoretical research in nursing	2 points
3. Meta-theoretical experience and knowledge:	
a) Elaboration or orientation of more than one nursing theory or meta-theory	4 points
b) Elaboration or orientation of a nursing theory or meta-theory	3 points
c) Elaboration or orientation of, at least, one non-nursing theory or meta-theory	2 points
4. Dissemination of knowledge produced on the topic of nursing meta-theories or theories:	
a) Authorship in more than one article published in an international standard indexed journal with a high impact factor (for Brazil: Qualis A1) on the theme of nursing theory or meta-theory	4 points
b) Authorship of an article published in an international standard indexed journal with a high impact factor (for Brazil: Qualis A2 to A4) on nursing theory or meta-theory	3 points
c) Authorship of, at least, one article published in a national indexed journal with a medium impact factor (for Brazil: Qualis B1) on the theme of nursing theory or meta-theory	2 points
5. Peer-recognition of expertise on the topic of nursing meta-theories or theories:	
a) More than one participation as a guest (lecturer, speaker, commentator, an instructor/professor in a course or short course) in a scientific event to teach a theme related to nursing theories or meta-theories	4 points
b) Participation as a guest (lecturer, speaker, commentator, an instructor/professor in a course or short course) in an event to teach a theme related to nursing theories or meta-theories	3 points
c) Participation as a listener/participant/student in a completed event or course on nursing theories or meta-theories	2 points

Figure 2 – Criteria for classifying the presumed level of expertise of evaluators of nursing theories. Rio de Janeiro, RJ, Brazil, 2019

The candidate for evaluator/judge has their level of expertise ranked by the score obtained by the sum of the items of the five domains. The maximum score for each domain is 4 points and the minimum is 2. When the evaluators do not meet any of the criteria for a domain, they receive a zero score. The total score ranges from

0 to 20, 4 points being the minimum arbitrated value to consider the candidate suitable to be a member of the theoretical evaluation team.

The level of presumed expertise for a candidate for evaluator is established by the score obtained from the analysis of the five domains. Five levels are proposed,

namely: beginner; advanced beginner; competent; proficient and expert.

The classification of a candidate's presumed expertise level is established from their total score obtained (see Figure 1).

The definition of the minimum number of evaluators in the team will depend on the sum of the individual points of each evaluator and on the mixed composition that guarantees members with at least two different levels of expertise (Figure 3).

Assumed level of expertise of the evaluator	Score required to be fit to the level	Criteria for defining the team, according to the expertise points
Beginner	Minimum sum of 4 points	(a) Team of two evaluators: The sum of the evaluators' points must be at least 35 points (b) Team of three evaluators: The sum of the evaluators' points must be at least 36 points (c) Team of four evaluators: The sum of the evaluators' points must be at least 48 points (d) Team of five evaluators: The sum of the evaluators' points must be at least 60 points (e) Team of six or more evaluators: Apply the following equation:
Advanced beginner	Sum between 5 and 10 points	
Competent evaluator	Sum between 11 and 14 points	
Proficient	Sum of scores between 15 and 17 points	
Expert	Sum of scores greater than or equal to 18 points	

Source: Brandão, MA, 2019

Figure 3 – Classification of the level of presumed expertise for the evaluators according to the total score obtained and definition of the number of evaluators in the team. Rio de Janeiro, RJ, Brazil, 2019

For example, by applying the proposed equation, composing a team of seven evaluators by calculating the total required points of the sum of the evaluators will require approximately 74 points. Exemplifying, an appropriate configuration would include six advanced beginner evaluators with 10 points each and a competent evaluator with 14 points. Obviously, other configurations that respect the minimum score for the team can be applied. In addition to the total score, the requirement for mixed teams in terms of expertise levels aims to guarantee the criterion on diversity of judges.

Knowing that the losses in the face of invitations and during initial Delphi rounds are common, it is recommended to select a higher number of judges required for the minimum composition of the teams, seeking to maintain proportionality between the levels.

The fourth procedure is the invitation to the selected candidates, based on written or electronic communication and respecting ethical research principles.

The fifth procedure involves the reapplication of the criteria for categorizing the evaluators/judges by their expertise levels and adjustments to the composition of the teams considering refusals to participate.

The sixth procedure is the planning of rounds and interruption criteria. This planning considers the level of abstraction, the number of concepts and the complexity of the theory to be evaluated. The number of judges developing the evaluations is also noteworthy, as is the consideration regarding the number of criteria to be evaluated in the theory. The interruption of the rounds must be supported by the explicit judgment on the part of the evaluators after reaching an evaluative consensus or constitution of a multiplicity of ideas in the dissent.

Another decision is to establish or not, *a priori*, a maximum number of rounds. This decision refers more to time available for the task than to the evaluative judgment.

Subsequently, to guide the evaluators, explicit criteria must be established for discarding items in each round. The items of a theory subjected to evaluation are its components like concepts, assumptions, suppositions, statements, and model schemes. Therefore, the evaluators must be certain that the decision to exclude is driven by the selected strategy and not, only, by their freely-expressed personal opinions. Their function is to judge a given theory item against the evaluation criteria established in the strategy.

The eighth procedure encompasses the definition of the consensus scope and of the stability of the answers. The consensus can be verified by formal measures of agreement, measures of central tendency, percentage of agreement, and measure of central tendency within a specific interval, among others<sup>(15)</sup>. The use of a five-point Likert scale can be planned with two purposes: (1) to verify the agreement of secondary evaluators with the result of the primary evaluation or (2) to organize the secondary evaluations in assertions that will be submitted to agreement analysis in a later round.

Even when scales are applied, it is recommended to guarantee free editing fields so that the evaluators/judges can express their suggestions, recommendations and detailed appraisals.

The ninth procedure is to provide specific guidance on the theoretical evaluation strategy. When the coordinator or primary evaluator deems it necessary, complementary and specific training on the content of the strategy can be carried out.



## Discussion

Theoretical evaluation is able to provide elements about a "good" theory, with several formal and systematic criteria available in the literature<sup>(2,11)</sup>. However, human resources with the competence and knowledge required to properly develop the process of judging theoretical virtues are not always available. And, in this respect, by using the principle of collective wisdom by consensus or dissent, the Delphi method can multiply the group's expertise, further expanding the universe of alternatives for the evaluation<sup>(15-16)</sup>. Likewise, it assists in the coordination of the process.

Through the evaluation, relationships and links of concepts are perceived, allowing the reviewer to verify the theory's strengths and limitations; identifying the need for new elements of the theory or improving the existing ones and, as a final goal, determining the potential contribution of the evaluated theory for the scientific knowledge<sup>(11)</sup>.

Unlike the theory analysis that decomposes a theory to examine its parts or components<sup>(4)</sup>, theoretical evaluation also judges them. However, even a theory judged to be "good" can prove to be inadequate in its descriptive, explanatory, predictive or prescriptive value from its confirmation or application. This places internal validation as a relevant stage, although not terminal of a theoretical development program.

Theories that violate the virtues of a "good" theory are more difficult to refute and tend not to, actually, contribute to knowledge<sup>(17)</sup>. The inadequacy of elements and constructions hinders theoretical evaluation and testing. Thus, it is fundamentally important to plan internal and/or external validation as part of a more comprehensive program. Using the guideline herein presented may avoid the expenditure of resources, when collaborating in the identification of theories that do not have sufficient virtues to support validation by field research.

The reasons for the reduced use of nursing theory evaluation strategies through formal systematic criteria are uncertain<sup>(4)</sup>. However, influence can be attributed to the difficulty in obtaining evaluators with sufficient epistemic authority to judge the meta-theoretical items of internal validation. It is supposed that the strategies linked to collective wisdom can overcome this problem of dependence on the "expert" with substantial advantages<sup>(18)</sup>.

The Delphi method is based on the John Dewey's assumptions, emphasizing anonymous communication between individuals with expertise in a given topic, with the goal of seeking the opinion of experts in an iterative and structured way and usually seeking to achieve a consensual position<sup>(15,19)</sup>. The freedom and

observance of the judges' personal opinions guarantees the independence criterion of collective wisdom<sup>(5)</sup>.

Regarding the use in research studies, although it is used predominantly in mixed and quantitative, it has its qualitative application and even in the construction of practical theories, in the context of community organization<sup>(15)</sup>. Theory evaluation is a qualitative process permeated by subjectivity and by standards, conducts and codes of the evaluator<sup>(6)</sup>.

The Delphi method can coordinate these qualitative characteristics of the evaluation process, dealing with personal variables of the independence criterion, making the most of group work. It can be used for interpretation, for predictions and for obtaining recommendations of the evaluation developed<sup>(6)</sup>.

In choosing the Delphi method, the most common approach is the traditional one, also being referred to as normative or of consensus. It aims to reduce variance in the estimates and biases among experts. However the Delphi Policy or Policy of dissent, seeks to obtain a wide range of opinions, but without seeking consensus<sup>(16)</sup>.

For the theoretical evaluation, consensus Delphi is the most likely indication; however, the use of dissent can be recommended for theories of high originality, conceptual density, complexity and theoretical abstraction or when it is difficult to determine the consensus criteria. Additionally, one of the goals of the evaluation can be to explore the contradictions in the production of definitions or theoretical proposals.

Regarding the characteristics of the theory, consensus Delphi can be indicated for those of micro- or middle-range with conceptualization described in more than one empirical study or to evaluate partially disseminated, tested or used theories.

Supposedly, for consensus Delphi the composition of teams with a high number of evaluators is only justified when it is difficult to obtain evaluators with higher levels of expertise, because it is challenging to obtain consensus in groups of many components. On the other hand, it is assumed that the dissent approach benefits from the composition of larger teams and with a wide range of proficiency levels, tending to broaden the debate from different perspectives and to bring original elements that differ from the original theory and from the primary evaluation.

Panels with more participants tend to have lower answer rates, with an estimated reduction of 0.08 percentage points for each added participant<sup>(20)</sup>. A number of 5 to 20 experts are indicated if it is a recommendation based exclusively on the characteristics of the Delphi method<sup>(20)</sup>. Studies on the development and application of Core Outcome Set (COS) have used the

Delphi method to determine which results to measure, with the predominance of Delphi panels of up to 50 people<sup>(20)</sup>.

In the theoretical evaluation it is challenging to establish a minimum and maximum number of evaluators/judges, due to its philosophical character and abstract epistemological nature inherent to theorization. For example, for new or poorly disseminated theories, it can be difficult to have many secondary evaluators with adequate expertise. On the other hand, large teams of beginner evaluators may not have knowledge of a meta-theoretical nature, causing a dispersion of perspectives that would hinder the aggregation of ideas. In this case, the guideline seeks to circumvent the limits by combining a balance between the criterion of diversity of the principle of collective wisdom and the expertise required for theoretical evaluation<sup>(4-5)</sup>.

The prototype of the guideline included four evaluators with three different expertise levels, and three secondary evaluators who together collectively summed 36 points (14, 13, and 9 individual points). According to the expertise points, the criteria for defining the team were useful for the composition of this small group, as the configuration of fewer participants guaranteed the maximum answer rate, as expected for this panel size<sup>(20)</sup>. The differences in training levels and stories of the evaluators ensured the decentralization criterion<sup>(5)</sup>. However, whenever possible, it is recommended to assemble teams with five or more judges.

Patricia Benner's model<sup>(21-22)</sup> with its five levels of competence acquisition was the basis for creating the judges' criteria of expertise in the guideline and sought to recognize the professional experience as an essential component for validation. The wide dissemination of studies by these authors and their criteria helped in the definition. There are more sophisticated models of aggregation rules to define the composition of the team, for example, the Contribution Weighted Model (CWM) that weighs the prognosis based on the relative performance of each judge and the accuracy of the group<sup>(18,23)</sup>.

The contributions of evaluators/judges have knowledge, experiences, and particular points of view in the evaluation of the theory. The iterative process of the Delphi method can allow that, in the rounds, the obscure criteria of the evaluation can be clarified or modified, through a careful interpretation of the answers of the secondary evaluators, by the coordinator. The composition of teams with different levels of competence guarantees the diversity criterion of collective wisdom<sup>(5)</sup>.

The studies commonly apply two to three rounds for the Delphi method<sup>(19)</sup>. However, the multiple criteria to be evaluated, the high number and diversity of profiles of the evaluators may require more rounds to reach consensus. It is desirable to plan a minimum according

to the number of evaluators, to ensure that an excessive effort to manage the task results does not fall on the Delphi coordinator, compromising their quality.

The scope level of theories can influence the definition of criteria to be evaluated by judges; for example, when a given middle-range theory is evaluated as a model, even more specific and empirical criteria can be used<sup>(12)</sup>. However, this does not, directly, interfere with the nature of the Delphi method as a strategy.

The decision to reach consensus among judges is a type of mechanism to meet the criteria of aggregating collective wisdom, transforming individual judgments into a team's decision<sup>(5)</sup>. The consensual decision can start from the evaluators' own opinion that a consensus was reached; however, it is recommended that this does not happen automatically after completion of the Delphi technique<sup>(19)</sup>.

It is necessary to specify which conditions are required for reaching consensus when the decision is qualitative. When quantitative measurement procedures are adopted, establishing the measures and cut-off points will be used to establish the degree of agreement or disagreement, compatible with the consensus or dissent<sup>(19)</sup>.

There are no mandatory rules for consensus building, but the five-point Likert scale is the most common among the scales used to estimate disagreements or agreements<sup>(24)</sup>. It makes it possible to check the degree of agreement for each item or set. For consensus reach estimates using the Likert scale, formal agreement measures such as the Kappa statistic can be applied, to verify the judges' concordant judgment on the elements of the theory.

Usually, the percentages of agreement adopt the value of 0.8 or 80% as a minimum cut-off point<sup>(15)</sup>. However, the researcher can consider other cut-off points, supported by evidence or by a consistent recommendation. An explicit statement on the reach of consensus is indicated with an indication of the reasons that were considered in decision-making.

Another useful measure that can be used on the data obtained by the Likert scale is the content validity coefficient. The Aiken coefficient and its caudal probability table can be used to indicate the validity of a particular item evaluated by several judges, estimating a consensus. It can also be applied to judge the validity, by a single judge, of the content for all the items of the theory. The coefficient range varies from 0 to 1, with higher values indicating validation<sup>(25)</sup>.

Despite the literature generally recommending the use of quantitative scales to signal consensus, qualitative justifications must be added, especially when the recommendation is for the item's invalidity<sup>(24,26)</sup>. The

simple exclusion of an item can make the whole theory incoherent or illogical. This is because units of a theory play roles and have different relevance in the theoretical structure.

For example, the exclusion of an assumption can de-characterize the theory as a whole, since this typology of element functions as premises not given to the empirical test. Thus, its removal negates the ideas that guided the theorists themselves in constructing the theory. On the other hand, "proposal" type units are submitted, precisely, to generate testing hypotheses in empirical validation studies; therefore, they are naturally subjected to exclusion or maintenance after evidence obtained from experimentation or field research<sup>(1)</sup>.

It is highlighted that, from the evaluation of the Theory of Professional Links<sup>(27-28)</sup>, emerging factors demanded changes to criteria not detailed when the prototype was elaborated, which contributed to the deepening for the creation of the guideline presented in this article.

The study is not limitation-free. The focus of any research using the Delphi method will always be obtaining high-quality answers from a selection of expert individuals<sup>(29)</sup>. However, the internal validation of a theory deals with theoretical-philosophical criteria that can make it difficult, for a secondary evaluator, to produce or judge the quality of the answers by the nature of the object evaluated and by the judgment property to be performed. For example, the conceptual definition is one of the elements of a theory, evaluated in its semantics, logic, and context<sup>(14)</sup>. Notably, it can be difficult to make a "good answer" judgment for such a complex construct, given such properties.

The limitation for the subjectivity of the judge's judgment in theories evaluation must be confronted with the philosophical root of the theorist and of the evaluator. Critical-social, hermeneutic or new pragmatism roots tend to deal with greater fluidity in the face of different perspectives, including exploring them in consensus or in dissent. On the other hand, as it requires greater objectivity of reality, post-positivism requires more stable, generalizable or measurement criteria<sup>(30)</sup>. In this last philosophical root, methods such as structural equation modeling, factorial analysis and multiple regressions may be the best choice for theory evaluation, obviously with criteria closer to external validation<sup>(4)</sup>.

Among the contributions for the advancement of scientific knowledge, the study adds an unexplored dimension of the incorporation of evaluators of different levels of meta-thematic expertise in the task of theoretical evaluation, including incorporating guidelines for the phases of this process. Given the growth in developing middle- and micro-range nursing theories and of a specific

situation, with the consequent training of new theorists, the guideline can facilitate the validation process for the new theories, making up a solid base of disciplinary knowledge<sup>(4,31)</sup>.

On the other hand, higher levels of meta-theoretical expertise tend to require long years of training in this field, being more common to be verified in academia and among senior researchers. By exploring the principle of diversity of expertise levels for theoretical evaluation, the study encourages the creation of teams of different expertise levels, promoting cooperation and the circulation of knowledge to those involved in this process.

Finally, the application of the Delphi method in nursing theory evaluation must be clearly understood as different from the search for consensus on events, phenomena, facts, technologies, conducts or any other fundamentally empirical elements. In the empirical Delphi method, the removal of an item can have minimal implication; however, in the theory, the judgement of the inadequacy of central suppositions or concepts can place the whole theory in the condition of inadequate. Obviously, the main goal of the evaluation is to identify a "good theory", which implies judging the adequacy of its components; however, this procedure must be performed with extreme caution by the evaluators, understanding that, in a theory, there is hierarchy and relationship between the elements.

## Conclusion

The guideline developed was able to adapt the elements of the Delphi method as a favorable resource for the internal validation of nursing theories, enhancing it with the incorporation of judges with different views of the world, experiences, scientific knowledge, and creativity. The criteria displayed in the guideline adapt and articulate the proficiency levels of the evaluators with the principle of crowd wisdom, serving as a guide for the selection and composition of teams of judges, as well as facilitating the coordination of the theoretical evaluation work. Due to its innovative character, the guideline can instrumentalize nursing meta-theorists and, possibly, speed up the process of applying theories in practice.

The use of a guideline prototype in the evaluation of middle-range nursing theory, the Theory of Professional Links, brought satisfactory results that presume its feasibility and pointed out ways for refinement.

It is understood that it is essential that other researchers replicate its use in the evaluation of grand- and micro-range theories for future adjustments and updates of the guideline, also adopting evaluation strategies by formal criteria different from the one used in the prototype.



## References

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](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 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

---

### Authors' Contribution:

**Study concept and design:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Marcos Antônio Gomes Brandão. **Obtaining data:** Manuela Campos Gomes Borel. **Data analysis and interpretation:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Marcos Antônio Gomes Brandão. **Drafting the manuscript:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Maria Miriam Lima Da Nóbrega, Cristina Arreguy-Sena, Marcos Antônio Gomes Brandão. **Critical review of the manuscript as to its relevant intellectual content:** Manuela Campos Gomes Borel, Rafael Oliveira Pitta Lopes, Maira Buss Thofehrn, Maria Miriam Lima Da Nóbrega, Cristina Arreguy-Sena, Marcos Antônio Gomes Brandão

**All authors approved the final version of the text.**

**Conflict of interest: the authors have declared that there is no conflict of interest.**

Received: Feb 10<sup>th</sup> 2020

Accepted: Jul 4<sup>th</sup> 2020

Associate Editor:

Maria Lúcia do Carmo Cruz Robazzi

**Copyright © 2021 Revista Latino-Americana de Enfermagem**

This is an Open Access article distributed under the terms of the Creative Commons (CC BY).


This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

---

Corresponding author:

Marcos Antônio Gomes Brandão

E-mail: marcosantoniogbrandao@gmail.com

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