

# PODIUM

Journal of Science and Technology in Physical Culture

---

**Volume 19**  
**Issue 1**

**2024**

University of Pinar del Río "Hermanos Saíz Montes de Oca"



*Translated from the original in spanish*

*Original article*

## *Extension of the use of expert judgment. Validity, consistency and reliability of scientific results*

*Extensión del empleo del criterio de expertos. Validez, consistencia y fiabilidad de los resultados científicos*

*Extensão do uso do julgamento de especialistas. Validade, consistência e confiabilidade dos resultados científicos*

Arcelio Ezequiel Fernández González<sup>1\*</sup> , Darmary Rodríguez Varis<sup>1</sup> ,  
Enilda Mariselis Jorrín Carbo<sup>1</sup> 

<sup>1</sup>University of Matanzas. Faculty of Physical Culture Sciences Matanzas, Cuba.

\*Corresponding author: [arcelio.fernandez@umcc.cu](mailto:arcelio.fernandez@umcc.cu)

*Received:* 27/03/2023.

*Approved:* 27/12/2023.

---

### **RESUMEN**

El criterio (juicio o consulta) de expertos, en las últimas décadas, ha sido ampliamente utilizado en las investigaciones cualitativas, puede decirse que para muchos investigadores constituye la "regla de oro" para validar sus hallazgos. El propósito del artículo consistió en proponer un procedimiento para el procesamiento estadístico de los datos cuando se emplea



---

una extensión del método convencional del criterio de expertos, para establecer la validez, consistencia y fiabilidad en los hallazgos científicos. El mismo fue aplicado en la toma de decisiones de una de las tres dimensiones de la variable de los componentes del ejercicio problémico interdisciplinario del proceso de enseñanza-aprendizaje. En el proceso investigativo se utilizaron métodos empíricos como la revisión de documentos, la encuesta y el criterio de expertos. Los tests estadísticos aplicados arrojaron significación estadística ( $P < 0.05$  hasta  $P < 0.001$ ) entre las comparaciones y/o asociaciones realizadas. Los hallazgos encontrados demostraron, al aplicar el procedimiento, una aproximación hacia la validez, consistencia y fiabilidad de los resultados científicos.

**Palabras clave:** consistencia, expertos, fiabilidad, procedimiento, validez.

---

#### ABSTRACT

In recent decades, expert judgment has been widely used in qualitative research, and it can be said that for many researchers it constitutes the "golden rule" for validating their findings. The purpose of the article was to propose a procedure for the statistical processing of data when an extension of the conventional method of expert judgment is used to establish the validity, consistency and reliability of scientific findings. It was applied in the decision making of one of the three dimensions of the variable of the components of the interdisciplinary problem-solving exercise of the teaching-learning process. Empirical methods such as document review, survey and expert judgment were used in the research process. The statistical tests applied showed statistical significance ( $P < 0.05$  to  $P < 0.001$ ) among the comparisons and/or associations made. The findings found demonstrated, when applying the procedure, an approach towards the validity, consistency and reliability of the scientific results.

**Keywords:** consistency, experts, reliability, procedure, validity.

---



---

## RESUMO

O julgamento de especialistas (julgamento ou consulta) tem sido, nas últimas décadas, amplamente utilizado em pesquisas qualitativas, indiscutivelmente o "padrão ouro" para muitos pesquisadores validarem suas descobertas. O objetivo do artigo foi propor um procedimento para o processamento estatístico de dados ao usar uma extensão do método convencional de julgamento de especialistas para estabelecer validade, consistência e confiabilidade em descobertas científicas. Ele foi aplicado na tomada de decisão de uma das três dimensões dos componentes variáveis do exercício interdisciplinar de solução de problemas do processo de ensino-aprendizagem. Métodos empíricos, como análise de documentos, pesquisa e julgamento de especialistas, foram usados no processo de pesquisa. Os testes estatísticos aplicados mostraram significância estatística ( $P < 0,05$  a  $P < 0,001$ ) entre as comparações e/ou associações feitas. Os resultados demonstraram, ao aplicar o procedimento, uma abordagem voltada para a validade, a consistência e a confiabilidade dos resultados científicos.

**Palavras-chave:** consistência, especialistas, confiabilidade, procedimento, validade.

---

## INTRODUCTION

The criterion (judgment or consultation) of experts, in recent decades, has been widely used in the social, humanistic, economic, technological, medical sciences and sciences of physical culture practice and education to validate a hypothesis, proposal or component of scientific research in the field of qualitative research (Aguilar et al., 2022; Díaz et al., 2020; Jorrín et al., 2021; Marrero and Smith, 2022; Mora and Lao, 2021; Robles and Rojas, 2015 and Torres *et al.*, 2022).

In qualimetric research studies, three evaluation methodologies using expert criteria are considered: preference, peer comparison and Delphi or Delphos (Díaz et al., 2020). The latter has been the most widely used by the scientific community to validate its findings (Ibid.).



Despite the wide use of expert judgment in qualitative research, which for many scholars could be said to be the "golden rule" in the validation of scientific results and whose validation, for others, is achieved in a relatively short time compared to that invested in experimentation, it has been widely questioned in the scientific literature, due to the inherent subjective component in obtaining the data, its objectivity, internal and external validity, reliability, trustworthiness, consistency and applicability (Cruz, 2020; Okuda and Gómez, 2005 and Robles and Rojas, 2015).

It is considered that the weaknesses of the method do not consist in the fact that the information, as suggested, is usually presented in an imprecise manner on nominal or ordinal scales, but in the detractors mentioned above. Its use is justified in those cases in which it is not possible to use quantitative methods of research and/or experimentation.

With the purpose of mitigating the possible biases associated with the expert judgment method and thus, pretending to give internal validity to the results of the study, it is proposed to use them as alternatives to the triangulation and evaluation of expert judgment with a fuzzy approach (Carvajal *et al.*, 2023; Cruz, 2020; Marín *et al.*, 2021 and Okuda and Gómez, 2005).

On the other hand, validity has been defined as the degree to which an instrument measures what it really intends or serves the purpose for which it was constructed. And reliability is understood as the degree to which the instrument measures accurately discarding error and does so through consistency, temporal stability and agreement among experts, Arribas (cited by Robles and Rojas, 2015).

It is assumed that an adequate statistical processing of data can give credence to the validity, consistency and reliability of the scientific findings found in purely qualitative research. Based on these two concepts, the purpose of the article is to propose, for decision-making between dimensions (items, pedagogical categories, scientific or methodological criteria, processes, etc.) a procedure for statistical processing of data when the expert judgment method is used in qualitative research and to achieve validity, consistency and reliability of scientific findings.



---

## MATERIALS AND METHODS

In the development of the research, theoretical and empirical methods were used, among which the empirical methods were considered:

Document review: it was used with the purpose of analyzing the information related to expert judgment, its origin, statistical procedures for its use, and its use in the development of qualitative research related to different sciences of knowledge such as: educational, medical, industrial, economic, social and humanistic, in applied linguistics and technological, business and event management; it was also used in the review of contents related to non-parametric statistical procedures.

The survey: it was used for the diagnosis and was applied to 20 researchers (anonymously), with the purpose of determining whether they were aware of the possible use of conventional methods of expert judgment (Delphi, pairwise or preference-based comparison) as described in the scientific literature, for decision-making between dimensions (items, pedagogical categories, scientific or methodological criteria, processes, etc.), or whether they had observed in the literature consulted the application of some of the procedures proposed in this study.

Expert criteria: the conventional method of expert judgment (Delphi) was used to determine the competence coefficient (K) of the expert candidates and the proposed procedure (extension of the use of expert judgment); in addition, the experts were asked to evaluate on a scale of 1 to 10 (points), the most important dimension between the delimitation of previous knowledge of the new elements to be sought (NE), actions to solve the problem (AS) and interdisciplinary relationship with the profession (RI), and to assign the highest score to the dimension they considered the most important, in order to score the answers to the questions asked to the students about an interdisciplinary problem-solving exercise in the subject of Exercise Physiology in students of the Bachelor's Degree in Physical Culture and to take into account the relative importance of the three dimensions considered for this purpose (NE, AS and RI). In this way, the principle of the integral character of the grades of the answers to the questions asked was not lost.



Consequently, a decision-making situation arose for which it was not possible to use the conventional methods of expert judgment (Delphi, pairwise or preference-based comparison) as described in the literature.

As a tool for testing the proposed procedure to the scores given by the experts to each of the dimensions under study, percentage values of the theoretical and real scores were estimated, and subsequently, a series of non-parametric statistical tests were applied to demonstrate the validity, consistency and reliability of the proposed procedure in the decision-making process; the SPSSPC version 25.0 statistical package was used for this purpose.

Practical validation of the proposed procedure (extension of the use of expert judgment): to determine (and resolve) which of the three dimensions NE, AS and RI (explained above) was the most important and the situations in which its solution was not possible with the application of the conventional methods of expert judgment, as described in the scientific literature.

The research was cross-sectional and qualitative and used tools from nonparametric statistical methods.

The selection of the sample responded to a non-probabilistic intentional sampling, in which the competence coefficient (K) related to the source of argumentation or substantiation of the subject under study was determined for 20 expert candidates, who were the same ones to whom the diagnostic survey was applied.

## RESULTS

### *Document review*

The review and analysis of the literature related to the topic studied made it possible to detect that although the existing conventional methods for the use of expert judgment (Delphi, peer and preference comparison) can be used to theoretically validate the proposal of a methodology, strategy, program, methods, items, pedagogical categories, scientific or



methodological criteria, processes, hypothesis, component of scientific research, etc., it was not possible to use it as described in the literature to make a decision between dimensions (items, pedagogical categories, scientific or methodological criteria, processes, etc.) of a methodology, strategy, program, methods, items, pedagogical categories, processes, etc.), it was not possible to use it, as described in the literature, to make a decision between dimensions (items, pedagogical categories, scientific or methodological criteria, processes, etc.) of a research or any component of it, so it was necessary to establish a new procedure for this purpose (Table 1).

### Survey

*Table 1. - Results of the applied survey*

Question	Yes (%)	No (%)
1	100	-
2	100	-
3	100	-
4	-	100
5	100	-
6	100	-
7a	-	100
7b	-	100
7c	-	100
7d	-	100
7e	-	100
8	-	100
9	-	100
10	-	100
11a	-	100
11b	-	100
11c	100	-

In Table 1, it can be observed that 100% of the respondents have a PhD degree, reviewed the scientific literature related to the expert criterion and applied it in their research (questions 1, 2 and 3, respectively), so they have knowledge of the subject under study.

Likewise, 100% of them state that the conventional methods of expert judgment (Delphi, pairwise or preference-based comparison) as described in the scientific literature cannot be applied to make decisions between dimensions (question 4).





100 % express that for the selection of the experts, the competence coefficient K, expertise index or their biograms are taken into account (question 5), and likewise, each expert (on a rating scale from 1 to 10) gave scores to each of the dimensions (question 6), as they are procedures used in the use of the conventional expert criteria.

It can be seen that 100% think that the total and real theoretical scores for the study dimensions are not estimated or calculated (questions 7a and 7b), nor the percentages of effective, theoretical and real scores for each dimension (questions 7c, 7d and 7e).

Likewise, for 100%, no statistical tests are applied for the a priori and a posteriori comparison of the means of the scores given, no statistical significance is sought for the estimated percentages, no statistical tests are applied for the comparison between dimensions of the percentages of the scores given, no statistics of central tendency and dispersion are estimated for each dimension under study, and the actual scores are not subjected to the goodness-of-fit test to normal distribution (questions 8, 9, 10, 11a, 11b, respectively). Although they did express that Kendall's coefficient of concordance is calculated to determine the association between the scores given by the experts to each dimension (question 11c).

Thus, the analysis of documents and the survey applied showed that, although the researchers use expert judgment as described in the scientific literature, it is not possible to use it for decision making among dimensions, since no procedure is described for this purpose, which justifies the development of the proposal entitled extension of the use of expert judgment.

#### *Extension of the use of expert judgment*

The procedures that make up the extension of the use of expert judgment are described below:

1. To insist that the researcher must be interested in making a decision between dimensions in the research or some component of it, for which it is not possible to use the conventional methods of expert judgment as described in the literature.



2. Select, if it has not been done previously, the experts to be consulted (the objectives to be achieved with the work are correctly informed) and determine the competence coefficients (K), the expertise index (IE) as described by Marreo and Smith (2022) or if preferred their biographies (biographies of the experts, according to Robles and Rojas (2015).
3. To survey those experts who have high competence coefficients related to the subject of the study.
4. The experts must consider, compare and study each of the dimensions (items, pedagogical categories, scientific or methodological criteria, processes, etc.) and award a score on a scale of 1 to 10 points, in order of importance.
5. Estimate or calculate:

a). - The total theoretical score ( $P_{tet}$ ) that can be awarded by all the experts during the assessment process of the dimensions under study, that is Equation 1:

$$P_{tet} = N \times D \times 10, (1)$$

Where:

N: number of experts participating in the assessment of the dimensions (usually between 15 and 30).

D: number of dimensions to be evaluated.

10 = total number of points that each expert can award to each dimension.

b). - The actual total score: sum of the sum of the scores given by the experts to each of the dimensions.

c). - Percentage of effective score awarded by the experts Equation 2:

$$PPef = (Prt / P_{tet}) \times 100 \% (2)$$

Where:

PPef = percentage of effective score.

Prt = actual total score.



Ptet = total theoretical score.

d). - Percentage of theoretical score given by the experts to each dimension as Equation 3:

$$PPtd = (Tpd / Ptet) \times 100 \% (3),$$

Where:

PPtd = percentage of theoretical score given to each dimension.

Tpd = total points awarded to the dimension.

Ptet = total theoretical score.

e). - Percentage of actual score given by the experts to each dimension as Equation 4:

$$PPrd = (Tpd / Prt) \times 100 \% (4),$$

Where:

PPrd = percentage of actual score awarded to each dimension.

Tpd = total points awarded to the dimension.

Prt = actual total score.

6. Apply a statistical test for a priori and a posteriori comparison of means (Post Hoc). For this purpose, the Kruskal-Wallis test and the Nemenyi test (multiple comparison of means test), respectively, or other equivalent tests, are suggested.

7.- Search the proportion significance table of Folgueira (2003) based on the Critical Values of the Sign Test algorithm (Bukaè 1975) for the statistical significance of the estimated percentages.

8.- Apply a statistical test for the comparison between dimensions of the percentages of the scores given by the experts, such as the Student's t-test or an equivalent test.

9.- In addition, it can be done:

a). - Estimate the central tendency and dispersion statistics for the scores given by the experts to each dimension under study (mean, median, standard deviation, maximum and minimum).

b). - Test for goodness-of-fit to the normal distribution using the Kolmogorov-Smirnov test



or another test, the real scores given by the experts.

c). - Compare between dimensions the a priori mean scores, through the alternative Jonckheere-Terpstra test and the median test.

d). - Calculate Kendall's coefficient of concordance to determine the association between the scores given by the experts to each dimension.

*Practical validation of the extension of the use of expert judgment*

The proposed procedure was applied during the development of the thesis Methodology to favor the interdisciplinary problematic approach from the subject Physiology of Physical Exercise of the Bachelor's Degree in Physical Culture, presented in option to the scientific degree of Doctor in Sciences of Physical Culture (Rodríguez, 2022) and was used for the practical validation of the proposed procedure.

Emphasis was placed on the evaluation of three dimensions of the variable of the components of the interdisciplinary problem-solving exercise for the teaching-learning process of Physical Culture undergraduate students, in order to determine which of these dimensions was the most important (procedure 1).

The experts (20 possible candidates), all PhD and full professors with vast experience in higher education teaching (15 years of teaching experience on average), after the determination of the coefficient of competence related to the source of argumentation or substantiation of the subject of study (procedure 2), the 15 with the highest K coefficient (procedure 3) were the experts who anonymously evaluated the dimensions of the interdisciplinary problem-solving exercise component, in order to determine the most important dimension (proceed 4) among NE, AS and RI, without neglecting to evaluate the other two, so as not to lose the principle of comprehensiveness when rating the answers to the questions asked to the students.

The questionnaire used to assess the importance of the dimensions of the components of the interdisciplinary problem-solving exercise is presented below:



Dear expert, we need your expertise to answer the following questionnaire in your opinion. In the class with an interdisciplinary problem-solving approach, the student must give an answer to an interdisciplinary teaching problem. In order to offer an integral evaluation to the answers given by the student, the teacher must analyze which of the following three dimensions is more important, without neglecting to evaluate the other two. Please mark (x) in the following evaluative scale and give the highest score to the dimension you consider most important (Table 2).

*Table 2. - Survey*

Dimension	Score									
	1	2	3	4	5	6	7	8	9	10
1 Precedent knowledge of the new elements to be searched (NE).										
2 Problem solving actions (SA).										
3 Interdisciplinary relationship with the profession (IR).										

Thank you very much for your cooperation.

The rest of the procedures (5 to 9) of the proposed procedure were completed.

*Results of the practical validation of the extension of the use of expert judgment*

The total theoretical ( $P_{tet}$ ) and total actual ( $P_{rt}$ ) scores were 450 and 391 points, respectively (procedures 5a and 5b).

Table 3 shows the statistical significance levels of the estimated percentages of effective score (procedure 5c) and within each dimension (both theoretical and actual, procedures 5d and 5e, respectively) (Tabla 3).

*Table 3. - Statistical significance for the percentages of the effective, theoretical and actual scores given by the experts within each dimension*

Percentage of effective score (%)	86.88 *		
	NE	AS	RI
Percentage of theoretical score (%)	24.44 ns	32.89 ns	29.55 ns
Percentage of actual score (%)	28.13 ns	37.85 ns	34.01 ns



NE: previous knowledge of the new elements to be searched; AS: actions to solve the problem; RI: interdisciplinary relationship with the profession. ns: not significant \*:  $P < 0.05$

The estimated percentages of theoretical and actual scores given by the experts within each dimension did not prove to be significant; however, the actual score did achieve levels of statistical significance (86.88 %,  $P < 0.05$ ), which evidences the high percentage of scores given to the study dimensions, which expresses the importance attributed to them in the evaluation process of the pedagogical tests to be subsequently applied to the students.

As can be seen, for the experts, the three dimensions were important, but AS stood out, with 32.89 % theoretical score and 37.85 % real score respectively (Table 2).

Although the percentages of theoretical and actual scores were not significant, Student's t test (procedure 8), whose values ranged from 2.808 to 2.897, yielded significant differences when comparing these percentages between dimensions ( $P < 0.05$ , Table 3). It also made it possible to consider the AS dimension as the most important for the evaluation process of the pedagogical tests, since it had the highest percentages of significant scores ( $P < 0.05$ ), as shown in Table 4.

**Table 4.** - Comparison between dimensions of the percentages of the scores given by the experts

	NE	AS	RI
Percentage of theoretical score (%)	24.44 a	32.89 b	29.55 c
Percentage of actual score (%)	28.13 a	37.85 b	34.01 c

NE: previous knowledge of the new elements to be searched; AS: actions to solve the problem RI: interdisciplinary relationship with the profession. ns: not significant \*:  $P < 0.05$  \*\*:  $P < 0.01$ . Percentages with different letters differ at  $P < 0.05$ .

The above result became consistent when comparing the actual mean values of the scores between dimensions, through the Kruskal-Wallis test, (proceed 6, table 4), this demonstrated the existence of highly significant differences (Chi-square equal 31.156,  $P < 0.001$ ), as did the alternative Jonckheere-Terpstra test (J-T equal to 441,500) and the median test (Chi-square equal to 29.400),  $P < 0.05$  or  $P < 0.001$  (Table 5).



**Table 5.** - Kendall, Kruskal-Wallis, Jonckheere - Terpstra, median tests for association and comparison between dimensions of the actual mean scores given by the experts

	NE	AS	RI
<b>Average (points)</b>	7.333 a	9.866 b	8.866 c
<b>Minimum (points)</b>	5.000	9.000	8.000
<b>Maximum (points)</b>	9.000	10.000	10.000
<b>Standard deviation</b>	1.175	0.351	0.639
<b>Median</b>	7.000	10.000	9.000
<b>N</b>	15	15	15
<b>D máx.</b>		1,598 **	
<b>W</b>		0.794 ***	
<b>Chi- cuadrado(K-W)</b>		31.156 ***	
<b>Chi- cuadrado (P-M)</b>		29.400 ***	
<b>J -T</b>		441,500 *	
<b>GL</b>		2	

NE: previous knowledge of the new elements to be searched; AS: actions to solve the problem; RI: interdisciplinary relationship with the profession. W: Kendall's contrast statistic; K-W: Kruskal-Wallis; P-M: median test; J - T: Jonckheere - Terpstra statistic; \*:  $P < 0.05$ ; \*\*:  $P < 0.01$  \*\*\*:  $P < 0.001$ . GL: degrees of freedom.

The Nemeyi test of multiple comparison of means posteriori or Post Hoc (procedure 7) showed that these differences were significant ( $P < 0.05$ ) among all the mean scores and the one with the highest mean score (9.866) was the problem-solving actions (AS) dimension, as can be seen in Table 5.

In addition to the above, Kendall's coefficient of concordance ( $W = 0.794$ ) showed a significant association ( $P < 0.001$ , Table 5) between the actual scores given to each of the study dimensions, which demonstrates a high degree of consistency in the criteria issued by the experts.

These findings made it possible to determine that the highest percentages of theoretical, actual (32.89 % and 37.85 %) and significant ( $P < 0.05$ ) scores were found for the AS dimension. Its higher mean score value (9.866) and median (10.000), when compared with the other dimensions (Tables 4 and 5) and the consistency in the scores given by the experts made it possible to consider this dimension as the most important of the interdisciplinary



problem approach variable for the purpose of scoring the questions in the students' evaluations.

Table 5 also shows the central tendency and dispersion statistics for the scores given by the experts within each dimension and the goodness-of-fit test to the normal curve. In the latter case, the Kolmogorov-Smirnov test statistic ( $D_{max} = 1.598$ ) proved to be highly significant ( $P < 0.01$ ), and justified the use of nonparametric statistics in this study, since the scores awarded by the experts to the dimensions under study (AS, NE and RI) do not follow a normal distribution.

Although the subjective component inherent to expert judgment was not eliminated, the findings demonstrate the validity, reliability and consistency of the scientific results for making the decision under study, although the procedure can be used in other studies where expert judgment is applied to validate a hypothesis or a component of a qualitative scientific research.

## DISCUSSION

Contreras and Palau (2020) and Herrera *et al.* (2022) consider that the quality of the results of the application of the expert judgment method depends, to a large extent, on the care taken in the preparation of the questionnaire and the choice of the experts consulted, a criterion shared by the authors, based on the results obtained.

Díaz *et al.* (2020) consider that the expert judgment method can be used in any part of the research that does not allow easy modeling. In the present article, the extension of the method (proposed procedure) was used for decision making, but it can be used to seek consistency of scientific results on the basis of the scores given.

Díaz *et al.* (2020) also used the method for decision-making on the financing of sustainable development projects, as well as Marreo and Smith (2022) for decision-making in maintenance planning in business management and Mora and Lao (2021) in the validation of the procedure for event management in Cuban hotels; but in none of the cases, decision-





making was based on the statistical processing of the scores given by the experts to the dimensions or components of the study research such as the one conducted.

Decision-making with the application of expert criteria is based on demonstrating the scientific proposal as Adequate (with its other value categories) or Not Adequate, and validating it. In this article, the decision was made as a result of the application of non-parametric statistics and the use of new valuation criteria such as the effective, theoretical and real percentages in general and within each dimension, derived from the scores given by the experts. This made it possible to approach the validation, consistency and reliability of scientific results in the field of qualitative or subjective research and thus, as Cruz (2020) considers, attenuate the inherent biases associated with these investigations, derived from the high degree of subjectivity in obtaining the data.

Okuda and Gómez (2005) and Cruz (2020) have proposed triangulation and fuzzy-focused expert judgment to increase the validity and consistency of the findings; however, the latter variant fails to eliminate the high subjective component of the expert judgment method and is somewhat complex to use when compared to the procedure proposed in the present study.

Herrera *et al.* (2022) suggest that the degree of agreement between experts can be calculated by estimating the Kappa statistic (when the variables are given on a nominal scale) and the Kendall's coefficient of agreement. In this case, this agreement was achieved by using Kendall's coefficient of concordance. Likewise, García *et al.* (2023) measured the degree of correlation and internal consistency of the variables among experts in an intervention strategy, using a method similar to the one used in the study presented.

## CONCLUSIONS

The review and analysis of the literature related to the topic studied, as well as the results of the survey applied, made it possible to detect that, although the existing conventional methods on the use of expert judgment can be used to theoretically validate a scientific



proposal, it is not possible to use them, as described in the scientific literature, to carry out a decision-making process such as the one addressed in this article.

The application of the proposed procedure, based on the evaluation of the three dimensions (delimitation of previous knowledge of the new elements to be sought, actions to solve the problem and interdisciplinary relationship with the profession) of the interdisciplinary problem-solving approach variable, to determine which should be considered as the most important for the evaluation and grading of the pedagogical tests on an interdisciplinary problem-solving exercise of the Physical Exercise Physiology subject in Physical Culture undergraduate students, made it possible to conclude, due to its higher and significant percentages of theoretical and real scores, its higher and significant mean score and median value (when compared to the other dimensions studied) and the consistency in the scores given by the experts, that the dimension of actions for solving the problem was the most important, which validates the proposed procedure to be used in decision making.

Although the subjective component inherent in the use of the proposed procedure (extension of the use of expert judgment) could not be eliminated, the findings demonstrated the validity, consistency and reliability of the scientific results and can be used, provided that the conventional methods of expert judgment are applied, to seek consistency of the scientific results based on the statistical processing of the scores given.

#### **ACKNOWLEDGMENTS**

The authors wish to express their sincere thanks to the experts consulted, without whose collaboration this work would not have been possible.



## BIBLIOGRAPHIC REFERENCES

- Aguilar, J., Jódar, E., Brañas, F., Gómez, C., González, Y., Malouf, J., Sánchez, R., Segura, J., Suárez, J., & Valdés, C. (2022). Consenso Delphi sobre Estrategias Terapéuticas y de Prevención Sanitaria de la hipovitaminosis D. *Rev Osteoporos Metab Miner.*, 14(4), 115-124. [http://www.revistadeosteoporosisymetabolismomineral.com/pdf/articulos/14\\_4\\_4.pdf](http://www.revistadeosteoporosisymetabolismomineral.com/pdf/articulos/14_4_4.pdf).
- Bukaè, J. (1975). Critical Values of the Sign Test. Algorithm AS 85. *Applied Statistics*, 24(2), 1-12.
- Carvajal, B., González, F. e Ibarra, L. (2023). Triangulación de métodos en ciencias sociales como fundamento en la investigación universitaria en Latinoamérica. *Revista científica de humanidades y artes*, 11(2), 1-16. DOI: <https://doi.org/10.5281/zenodo.8140907>.  
<https://revistas.uclave.org/index.php/mayeutica>.
- Contreras, N., & Palau, M. (2020). Diseño y Validación de un Cuestionario para Evaluar el Clima Organizacional Hospitalario. *Salud y Administración*, 7(19), 3-11. <https://revista.unsis.edu.mx/index.php/saludyadmon/article/view/165/133>.
- Cruz, M. (2020). *Un procedimiento de evaluación basado en el criterio de expertos con enfoque difuso*. *Focus*. Ediciones UDG, Universidad de Holguín, 1 15. <https://revistas.udg.co.cu/index.php/roca/article/view/1684>.
- Díaz, F., Cruz, M., Pérez, M. y Ortiz, T. (2020). El método criterio de expertos en las investigaciones educacionales: visión desde una muestra de tesis doctorales. *Revista Cubana de Educación Superior*, 39 (1. Enero - Abril), 1- 15. <http://scielo.sld.cu/57-4314-rces-39-01 -e18.pdf>
- García, R., Ayup, D., Mendoza, N., Milián, P. y Castaneda, I. (2023). Validación Delphi: Estrategia de intervención para mejorar el clima organizacional en centros diagnostico integrales venezolanos. *Universidad y Sociedad*. *Revista Científica de la Universidad*



- de Cienfuegos, 15 (1), 723-734.  
<https://rus.ucf.edu.cu/index.php/rus/article/view/3589>
- Herrera, J., calero, J, González, M., Collazo, M. y Travieso, Y. (2022). El método de consulta a expertos en tres niveles de validación. *Revista Habanera de Ciencias Médicas*, 21(1), 1-11. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1729-519X2022000100014](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1729-519X2022000100014)
- Jorrín, E.; Quintana, D. y Kessel, J. (2021). Estudio preliminar de la orientación del contenido estadístico durante el proceso de formación del profesional de Cultura Física. *PODIUM-Revista de Ciencia y Tecnología en la Cultura Física* 16(2), 576-592. <https://podium.upr.edu.cu/index.php/podium/article/view/994>.
- Marrero, R. y Smith, A. (2022). Diseño del grupo de expertos para contribuir a la gestión de la planificación del mantenimiento. *Revista Universidad Sociedad*, 14 (S1), 97 109. <https://rus.ucf.edu.cu/index.php/rus/article/view/2615/2564>.
- Marín, F., Pérez, J., Senior, A. y García, J. (2021). Validación del diseño de una red de cooperación científica utilizando el coeficiente K para la selección de expertos. *Información Tecnológica*, 32(2), 79-88. <http://dx.doi.org/10.4067/So718-07642021000200079>.
- Mora, J. y Lao, Y. (2021). Factibilidad del método criterio de expertos para validar el procedimiento para la gestión de eventos en hoteles cubanos. *Anuario Facultad de Ciencias Económicas y Empresariales*, 2(1), 1-10. <https://anuarioeco.uo.edu.cu/index.php/aeco/article/view/5216>.
- Okuda, M. y Gómez, C. (2005). Métodos en investigación cualitativa: triangulación. *Revista Colombiana de Psiquiatría*, 3(1), 119-124. <https://www.redalyc.org/pdf/806/80628403009.pdf>.
- Robles, P. y Rojas, M. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en Lingüística aplicada. *Revista Nebrija de Lingüística Aplicada*, 18(1), 1-18.



[https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo\\_55002aca89c37.pdf](https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_55002aca89c37.pdf).

Torres J., Vera V., Zuzunaga, F., Talavera, J. y De la Cruz, J. (2022). Validez de contenido por juicio de expertos de un instrumento para medir conocimiento, actitudes y práctica al consumo de sal en la población peruana. *Revista Fac. Hum.*, 22(2), 273-279. Doi.: 10.25176/RFMH, v 22i2, 4678. <http://revistas.urp.edu.pe/index.php/RFMH>.

***Conflict of interest:***

The authors declare that they have no conflicts of interest

***Authors' contribution:***

**Arcelio Ezequiel Fernández González:** conception of the idea, literature search and review, preparation of instruments, application of instruments, collection of information, results of the instruments applied, preparation of database, statistical analysis, preparation of tables, general advice on the topic addressed, drafting of the original (first version), revision and final version of the article, correction of the article, authorship coordinator, translation of terms or information obtained, revision of the bibliographic standard applied.

**Darmary Rodríguez Varis:** literature review, preparation of instruments, application of instruments, collection of information, results of the instruments applied, preparation of tables, preparation of database, proofreading and writing, final version of the article, revision of the bibliographic standard applied.

**Enilda Mariselis Jorrín Carbó:** review of literature, preparation of instruments, application of instruments, collection of information, results of the instruments applied, preparation of tables, preparation of database, proofreading and writing, final version of the article, revision of the bibliographic standard applied.





This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license.

