Observational Study

Medicine

OPEN

Validation of the knowledge evaluation questionnaire of the cardiopulmonary resuscitation training program in high school students

Fernando Calvo-Francés, PhD[a], Lucía Cilleros-Pino, PhD[a], Paula Andrea Barraza-Illanes, PhD[b], Maximino Díaz-Hernández, PhD[a], José Enrique Hernández-Rodríguez, PhD[a,*]

Abstract

There is an urgent need for generalized training in cardiopulmonary resuscitation (CPR) techniques, starting with secondary education. Validated instruments for assessing the efficacy of such interventions are not yet available. This study aimed to validate an evaluation questionnaire of a CPR training program for high school students, to analyze the levels of readability, difficulty, reliability, and content validity, as well as the fit the purpose for which they were designed, the trait they are intended to measure. An instrumental study was conducted in 2 phases. In the first phase, an inter-judge validation was carried out with 11 experts in CPR accredited instructors of basic and advanced CPR by the American Heart Association. In the second phase, the psychometric properties were evaluated from the perspective of Item Response Theory. During May of the 2017/18 and 2018/19 academic years, 259 4th-year secondary school students from a high school in the southeastern area of the island of Gran Canaria (mean age: 15.78 years; 50.60% male) were surveyed anonymously using the questionnaire to be validated. The questionnaire was easily readable (74.12 Flesch-Szigriszt Index); the difficulty level (Easy/Very Easy) in the context of this secondary school level of education and the ability level of the respondents overlapped sufficiently and there was no gender bias. The questionnaire was able to discriminate between respondents of slightly more than 7 levels of expertise, from low knowledge of CPR to high knowledge of CPR (Separation Index 7.53). The model fit was excellent (infit = 1/outfit = 1.01). The content validity index was adequate. The separation index and reliability exceeded what was considered adequate for guaranteed use. The level of difficulty of the items and the level of ability of the respondents is in line with the educational level of the students. The questionnaire did not produce a gender bias in response probability. The questionnaire is easily understandable and can discriminate between different levels of ability without differential gender bias, and its reliability is outstanding, as it exceeds the minimum criteria.

Abbreviations: CPR = cardiopulmonary resuscitation, INFIT = information-weighted fit statistic, IRT = item response theory, MNSQ = mean square infit statistic, OHCA = out of-hospital cardiac, OUTFIT = outlier-sensitive fit statistic, PROCES = cardiopulmonary resuscitation program oriented to compulsory secondary schools [Programa de Reanimació Cardiopulmonar Orientat a Centres d'Ensenyament Secuandari], RMSE = root mean squared error, ZSTD = mean square fit statistic t standardized.

Keywords: assessment, item response theory, psychometrics, questionnaire, scholar cardiopulmonary resuscitation

1. Introduction

Out of-hospital cardiac arrest (OHCA) is a significant public health problem worldwide.[1,2] Globally, it is estimated that, on average, <10% of all patients with OHCA will survive.[2] obtained survival rates to hospital admission of 22.0% and survival rates to hospital discharge of 8.8%.[2] The critical determinant of survival is the timely performance of bystander cardiopulmonary resuscitation (CPR) before the arrival of the emergency service.[3,4] Most people who experience an OHCA event do not get this procedure.[5]

The greatest challenge for successful out of-hospital resuscitation is CPR education.[1,5,6] The consensus conference of the International Liaison Committee on Resuscitation and different experiences, recommended further development of education and training programs for the population.[7,8] Specifically,

The authors have no funding and conflicts of interest to disclose.

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Supplemental Digital Content is available for this article.

[a] Nursing Department, Faculty of Health Sciences, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, [b] Fernando Pessoa Canarias University, Department of Clinical Psychology and Neuropsychology, Santa María de Guía, Spain.

** Correspondence: José Enrique Hernández-Rodríguez, Nursing department, Faculty of Health Sciences, University of Las Palmas de Gran Canaria, C/ Paseo Blas Cabrera Felipe, s/n, Las Palmas de Gran Canaria 35016, Spain (e-mail: joseenrique.hernandez@ulpgc.es).*

How to cite this article: Calvo-Francés F, Cilleros-Pino L, Barraza-Illanes PA, Díaz-Hernández M, Hernández-Rodríguez JE. Validation of the knowledge evaluation questionnaire of the cardiopulmonary resuscitation training program in high school students. Medicine 2023;102:42(e34609).

Received: 17 April 2023 / Received in final form: 13 July 2023 / Accepted: 14 July 2023

http://dx.doi.org/10.1097/MD.0000000000034609

it is vital to teaching these techniques from an early age, in primary and especially in high school, according to Pekins et al and Cheng et al[8,9] In a study by Jones et al,[10] only 13 to 14-year-olds performed chest compression as well as adults in other reported studies. No year 5 pupil (age 9–10) was able to compress the manikin chest to the I dth recommended in guidelines, and only 19% of pupils in year 7 (age 11–12) achieved adequate compression depth.[10] The data support that nurses have the potential and training to implement both these types of educational and cardiovascular disease prevention interventions,[11] in addition to their high efficacy in carrying them out.[12,13] Many CPR training programs have been implemented in Spanish schools.[14] These are: cardiopulmonary resuscitation program oriented to compulsory secondary schools [Programa de Reanimació Cardiopulmonar Orientat a Centres d'Ensenyament Secuandari] "(PROCES)" in Cataluña, "RCP na Aula" in Galicia, "Alertante" in Madrid and "Plan Salva Vidas" in Andalucía.[15–18]

To evaluate the efficiency of any educational intervention, valid and reliable measuring instruments are required. However, the questionnaires used for performance evaluation in CPR programs have not been validated.[13,15–22] We believe that the PROCES[15] questionnaire meets all the requirements for assessing the knowledge acquired by high school students in CPR techniques. We aim to validate using the item response theory (IRT), the levels of readability, difficulty, trustworthiness, and content validity, as well as the fit of the model.

2. Objective

(1) To validate an evaluation questionnaire of a CPR training program for high school students.
(2) To analyze the levels of readability, difficulty, reliability, and content validity.
(3) To analyze whether the questionnaire items fit the purpose for which they were designed and the trait they are intended to measure.

3. Methods

3.1. Design

An instrumental study[23,24] was developed in 2 phases. In the first phase, an inter-judge validation was carried out with 11 experts in the CPR, American Heart Association accredited instructors of basic and advanced cardiopulmonary resuscitation, who work in hospital emergency departments and prehospital emergency care with extensive experience in cardiac emergency care and CPR education and training. In the second phase, the psychometric properties were evaluated from the perspective of Item Response Theory (Fig. 1).

Item Response Theory, within Psychometrics, attempts to measure latent traits in item characteristics using mathematical models, offering insight into the relationship between an individual's trait level (e.g., level of difficulty and ability to perform a procedure) and item characteristics. Relationship between an individual's trait level (e.g., level of difficulty and ability to perform a procedure) and item characteristics. This theory depends on 2 important assumptions, namely the unidimensionality and the local independence of the items. With the first assumption, items only specifically measure a single latent trait, whereas, with the second assumption, it is assumed that a subject's responses to any pair of items in the test are unrelated when the same trait level is considered, that is, when an ability is held constant. On the other hand, one of the salient features of this theory concerning classical test theory is that the characteristic parameters of the items and the test do not depend on the sample and its characteristics, as is the case with TCT. This provides an advantage that makes it advisable for test validation and test construction because the item parameters are assumed invariant in the population of individuals. This theory generates both unidimensional

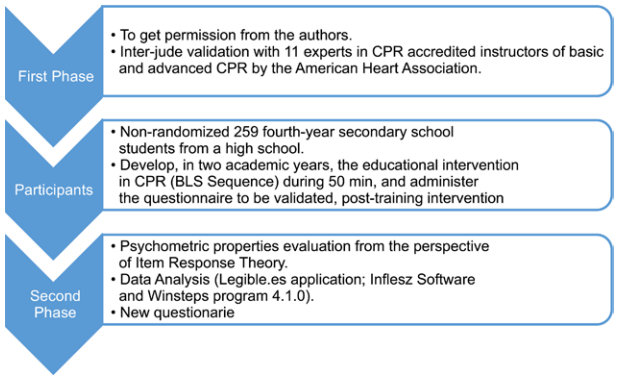

Figure 1. Phases of instrumental study development.

Table 1

Rasch model parameters and their interpretation.

Parameter	Acceptable range*
INFIT OUTFIT MNSQ values (mean square fit statistic)†	An MNSQ close to 1 indicates that the model fits the data well and that the observed responses fit the model's expectations. MNSQ < 1: suggests that the model fits better than expected (higher precision of the estimate). MNSQ > 1: The model fits worse than expected, indicating a discrepancy between the observed responses and those expected by the model.
INFIT OUTFIT ZSTD values (mean square fit statistic *t* standardized)†	ZSTD close to 0: indicates a good fit of the model, which implies that the observed responses and the expected responses according to the model are in agreement. ZSTD less than −2 or >2: may indicate poor model fit.
P, SD (Population standard deviation)	P, SD values can vary depending on the characteristic being measured and the demographics of the population. Tests that measure specialized knowledge in a homogeneous population may have lower P, and SD values (which is our case).
S, SD (Sample standard deviation)	One way to assess the adequacy of S, SD values is to compare them to the Population Standard Deviation (P, SD). If the value of S, SD is similar to or close to P, SD, it may indicate that the selected sample is representative of the population and that the variability measured in the sample is similar to the variability in the general population.
RMSE (Root mean squared error)	In general, lower values of true RMSE and model RMSE indicate better accuracy and fit of the model to the observed data. A low RMSE means that the observed scores and the actual or expected scores are very similar, implying that the model has a good ability to predict the scores.

INFIT = information-weighted fit statistic, MNSQ = mean square infit statistic, OUTFIT = outlier-sensitive fit statistic, RMSE = root mean squared error, ZSTD = mean square fit statistic t standardized.

* It is important to note that the interpretation of these values should be done in conjunction with the other adjustment indicators and considering the specific context of the test and the measurement objectives. There is no single, absolute threshold for determining which values are appropriate. A common way to evaluate these values is to compare them with those found in other similar studies carried out in the same population or similar populations, but in the systematic review, we did not find any.

† For both parameters, a value of MNSQ (mean square infit statistic) = 1, indicates a perfect fit, whereas values between 0.8 and 1.3 are considered adequate. As for the standardized fit statistics ZSTD (Mean Square Fit Statistic t standardized) values, a good fit is between −2 and + 2.

and multidimensional models, with different response formats, dichotomous or polynomial. In IRT, the score of the aptitude, either trait, ability, or competence level is considered independent of the item so the subject's response to the item will depend only on that level. In this sense, the trait would be the independent variable, and the response, the dependent variable.[25–27]

4. Participants

The main inclusion criterion was to select all students in the 4th year of Compulsory Secondary Education from a high school in the southeastern area of the island of Gran Canaria (Spain). (ESO) between 15 and 18 years of age were surveyed anonymously using the questionnaire to be validated. They formed a nonrandom, voluntary, and convenient sample. Data were collected during the 18/19 and 19/20 school years before the start of confinement in March 2020. According to Nunnally[28] formula for the calculation of sample size in item response theory, for a confidence level of 95% and a margin of error of 5%, the sample should be at least 245 persons, and this requirement was met.

5. Procedure

We performed the process after obtaining permission from the authors. The eleven experts evaluated the wording and representativeness of the questionnaire items and suggested additions, modifications, or deletions to the questionnaire. The experts determined the content validity of the items, which was understood as the degree to which they adequately represented the construct to be measured. According to Tristan model, they were asked to classify items as essential, useful, or unnecessary. A value ≥ 0.58 was established as the cutoff point for maintaining the item, as proposed by Tristán-López, as the minimum value to be obtained when assessing content validity[29] After these calculations, we analyze the legibility and readability of the last version.[30,31] Finally, the questionnaire was administered to a nonrandom sample of Spanish-speaking high school students before and after a CPR training program. For this study, only the responses to the posttest were used, since this is the point at which knowledge may have been acquired and, therefore, can be evaluated. The PROCES questionnaire has been validated using IRT (Rasch analysis).[32]

6. Instruments

The PROCES program questionnaire has 20 multiple-choice questions with only 1 correct answer out of the 4 options. The questionnaire comprises 2 sections, notions of physiopathology and CPR practice, each of which is 10 questions long.[15]

Based on the experts recommendations, we add 4 items that account for another 3 steps to the steps to take in cardiac arrest: ensuring the scene is safe, moving the victim to the safety position, and using a defibrillator (items 17, 18,19, and 20). Additionally, question 14 was modified from the original to match the updated guidelines.[8,9] The final questionnaire to be validated consisted of 24 questions.

7. Data analysis

We analyze the readability of the last version of the questionnaire through the Web Legible.es application and INFLESZ software.[30,31]For Rasch analysis, we used the Winsteps program 4.1.0.[32,33] We carried out an analysis considering the correctness and accuracy of the answers to the questions. Compared to classical test theory, IRTs are less dependent on the sample and items used, and normality is not mandatory.[25–27,34]

The main parameters to obtain through Winsteps are Table 1:

Information-weighted fit statistic (INFIT) and outlier-sensitive fit statistic (OUTFIT): These are 2 types of statistics used to assess the fit of the data to the IRT models. Both statistics are based on the comparison of the observed responses and the expected responses according to the model. The INFIT and OUTFIT measure the discrepancy between observed and expected responses relative to the expected accuracy of the model. They are used to detect unusual or anomalous response patterns that could indicate poor model fit.

Mean square fit statistic (MNSQ): It is a statistic used to assess the fit of the model in the IRT. Calculates the discrepancy between the observed and expected responses based on the model and averages them across all test items. An MNSQ value close to 1 indicates a good fit of the model, while values greater or <1 indicate a poor fit.

ZSTD (Mean Square Fit Statistic *t* Standardized): It is a standardized version of the MNSQ statistic. It is calculated by dividing the MNSQ by its standard error and provides a measure of relative fit that can be compared between different tests or groups.

Population Standard Deviation: P refers to the population standard deviation, which is a measure of the dispersion of scores in an entire population. It represents the variability of the characteristic measured by the items in the general population.

Sample Standard Deviation: S refers to the sample standard deviation, which measures the dispersion of scores in a population sample. It represents the variability of the characteristic measured by the items in the selected population sample.

Root mean squared error (RMSE): It measures the average difference between the observed values and the estimated values in a model. In the context of IRT, the RMSE is used to assess the accuracy of the model in estimating the abilities of individuals. A lower RMSE value indicates better model accuracy.[34–36]

The Wright measure map was used to interpret the estimated measures in the distribution of the items and respondents, and the Mantel–Haenszel test was used to rule out differential bias by gender. The significance level was set up as $P < .05$. The Wright–Person Measure Map is a visual representation used in IRT. This map shows the relationship between individual skill levels and item difficulty on a common scale (DIGIT units). It allows to identify the location of the individuals concerning the items and vice versa, which helps to understand the suitability of the items to measure different levels of ability. On the right of

Table 2

Readability indices.

Index	Value	Difficulty
Readability	80.27	Easy
Perspicuity	75.48	Easy
Legibility (Flesch-Szigriszt Index)	74.12	Pretty easy
Grade level	1.6	
Estimated time of reading	4.2 min	

Table 3

Standardized residual variance in eigenvalues and variance explained.

	Eigenvalue	Observed	Expected
Total raw variance in observations.	28.11	100.0%	100.0%
Raw variance explained by measurements	8.11	28.9%	28.5%
Raw variance explained by persons	2.53	9.0%	8.9%
Raw variance explained by the items	5.58	19.9%	19.6%
Unexplained raw variance (total)	20.00	71.1%	71.5%
Unexplained variance in the 1st contrast	1.95	6.9%	9.8%

Table 4

Reliability, separation and INFIT OUFIT adjustment of the items to the model.

						INFIT		OUTFIT
	Total score	Count	Measure	STANDARD ERROR MODEL	MNSQ	ZSTD	MNSQ	ZSTD
Media	138.90	259	.00	.15	1.00	–.1	1.01	–.1
P.SD	55.50	.0	1.18	.02	.06	1.1	.09	1.0
S.SD	56.90	.0	1.21	.03	.06	1.1	.09	1.1
MAX.	239	259	2.14	.24	1.11	1.9	1.17	1.5
MIN.	40	259	–2.55	.13	.89	–2.5	.87	–2.3
REAL RMSE	.16	TRUE SD	1.17	SEPARATION	7.53	RELIABILITY	.98	
RMSE MODEL	.15	TRUE SD	1.17	SEPARATION	7.63	RELIABILITY	.98	

INFIT = information-weighted fit statistic, MNSQ = mean square fit statistic, OUTFIT = outlier-sensitive fit statistic, P SD = population standard desviation (When the sample is the whole population), RMSE = Root mean squared error, SD = standard desviation, S SD = sample standard desviation (When sample represents population), ZSTD = mean square fit statistic *t* standardized.

the map, the difficulty of the items is represented, where the easiest items are located at the bottom and the most difficult items at the top. The ability of individuals is represented on the left of the map, where individuals with lower skill levels are on the bottom, and individuals with higher skill levels are on the top.[35,36]

8. Human subjects approval statement

The Board of the Secondary School where the study was carried out, as well as the Ethics and Research Committees of the reference hospitals in the area, approved the study Ethics Committee Hospital Universitario Dr. Negrín (CEImHUGC DrNegrin:2019-168-1). The participants signed an informed consent form through their legal guardians, who were duly informed through the communication channels of the educational center.

9. Results

9.1. Participants

Finally, the participants were 259 students. The mean age was 15.78 years (SD = 0.86), and 50.6% were male. Only 1 student did not complete the questionnaire because he was absent from one of the data collection interventions.

10. Content validity

In line with the cutoff point for content validity (value ≥ 0.58), 4 original items (4, 5,10, and 11) were removed. Finally, the questionnaire contained 20 items (see Supplemental Digital Content, http://links.lww.com/MD/J413, which contains the final validated items). The total content validity index was 0.83, which was considered appropriate.[29]

11. Legibility

As the readability parameters indicate, the final version of the 20 items (Table 2) shows that the text is fully comprehensible for the target population.[30,31]

12. Validation based on item response theory

We verified that the polarity items (biserial correlations) were positive (from.16 to.50), but 2 items (1 and 17) were slightly below the desirable level ≥ .20.[36] Since the data fit the model (Table 3), showing the success of the Rasch estimate, no item was removed from the analysis.[35] In Table 3 shows that the eigenvalue of the first contrast (1.95) was < 3, and the percentage of unexplained variance (6.9%) was almost 3 times lower (2.88) than the explained variance of the items (19.9%). As for the explained variance versus the unexplained variance, unidimensionality is accepted if the Rasch measurement shows a percentage variance of ≥ 20%; henceforth, with 28.9%, we have an appropriately explained total variance. All criteria are fulfilled, so we accept unidimensionality.[25,34,35] We accept local independence as all the residual correlations, apart from 3, were negative (between -.16 and -.28), and the positive ones (.22,.24, and.36) were < .50, which is acceptable in the general framework.

After verifying the unidimensionality and the local independence, we evaluated the model fit through the INFIT (Information-Weighted Fit Statistic) and OUTFIT (Outlier-Sensitive Fit Statistic) values (Table 4). Both (see Table 4) the overall INFIT (MNSQ = 1.00. ZSTD = –.1) and OUFIT (MNSQ = 1.01. ZSTD = –.1) values were good. Only item 14 showed INFIT and OUTFIT ZSTD values outside the acceptable range (INFIT = 2.5; OUTFIT = –2.3), but not MNSQ parameters (INFIT = .89; OUTFIT = .87). No decision needs to be made on this, as there is no correspondence of an MNSQ > 1.50.[35]

The separation and reliability indexes are appropriate (Table 4), with a value of 7.53 and.98, respectively. Both indices indicate that the items comprise a well-defined variable and that the reliability of the location of the items on the scale is good. Specifically, the Separation Index tells us that the questionnaire is capable of perfectly discriminating between more than 7 levels of knowledge. Table 4 also shows that the Population Standard Deviation (55.5) and S, S D (56.9) values are very similar, which indicates that the selected sample is representative of the population and that the variability measured in the sample is similar to the variability in the population in general. In turn, the RMSE values, both those referring to the actual values (.21) and the values predicted by the model (.20), are low and similar, which implies that the model has a good capacity to predict the scores.

When the data are complete, it is expected to approach 1 for people and 1 for items, as is the case here. We found a Pearson raw score-to-measure correlation = 1.00 and the item raw score-to-measure correlation = –.99, which aligns with what is desirable.

The Wright map (Fig. 2) shows the distribution of the respondents and items measured in DIGITS. The item difficulty level (–2.55 to 2.14 digits) on the left side of the graph and the respondents' ability level (–1.74 to 2.17) almost overlapped, with the average of the latter being slightly higher (.00 vs .21).

Finally, we checked the differential item functioning according to sex. In Figure 3, each point represents an item. Graphically, we generally did not observe sex differences except in items 16 and 20. Women have a higher performance in item 17 compared to men, quite the opposite of what

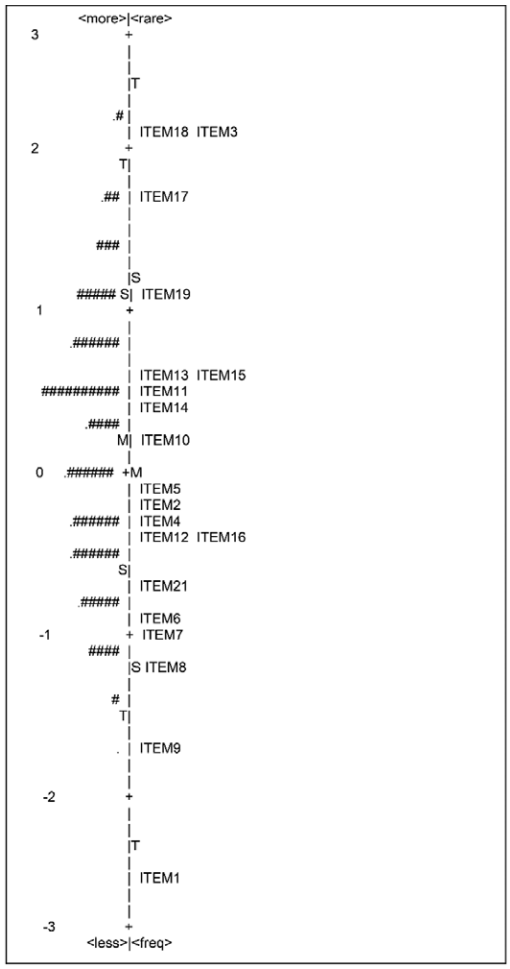

Figure 2. Wright Measure-Person map.

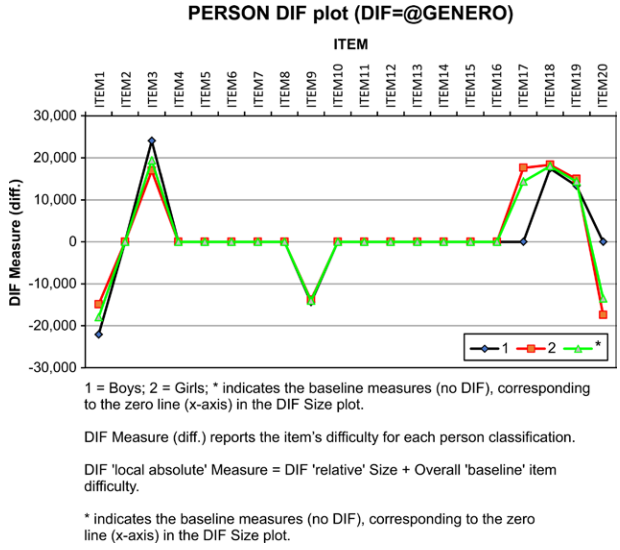

Figure 3. Cross plot of item difficulty measures by gender.

happens with item 20. Figure 3 shows the almost complete overlap of both distributions of difficulty according to gender. Focusing on the significance of the differences, with the Mantel–Haenszel test statistic, no significant differences by gender were associated with any items (the probability ranged from .0717–.9881).[32]

13. Discussion

The questionnaires used to evaluate performance in the CPR programs analyzed have not been validated, as mentioned above, which is the reason for this instrumental study.[13,15–22]

This study aimed to validate the PROCES questionnaire for the evaluation of CPR training programs for high school students, as it is the first to measure, in a structured manner and adjusted to the academic level of the students, the degree of knowledge acquired after the development of an educational intervention in CPR, since it fulfills all the requirements for evaluating the knowledge acquired by high school students in CPR techniques, As we have already indicated, it meets all the requirements for evaluating the knowledge acquired by high school students in CPR techniques, offering a real vision of the competence acquired in this area according to the academic levels so that its use should be emphasized due to its adaptability, as we have been able to confirm in our experience. The readability of the questionnaire is suitable for the level of potential users; therefore, the results are not biased by the greater or lesser command of the language of those evaluated, guaranteeing that the questionnaire only measures CPR knowledge.[26,27] In addition, the content validity index is appropriate, so the questionnaire measures what it purports to measure.[25]

PROCES was specially designed to be incorporated into the curricular material of children between 14 and 16 years of age in their educational centers so that pedagogical aspects were identified as a key factor for the success of the program. The formal structure of PROCES is based on universally accepted criteria in basic CPR following international guidelines (AHA)[9] and is pedagogically adapted to the target population for which the program is intended.[14]

In the preliminary aspects of IRT, the total explained variance was less than desired (≥40%),[28] but acceptable (28.9%). It should be noted that the explained variance depends on the scattering of people and items. For people with a wide range of abilities or items with a wide range of difficulties, this could indicate that the instrument is inappropriate. However, if the participants have a limited range of abilities (skills acquired after a brief workshop on CPR), and the items have a limited range of difficulty (valuing a small set of actions, with a small variability of difficulty, that constitutes CPR), the best sample may produce a modest, explained variance.[28] In addition, the items meet the requirement of creating a linear relationship concerning the latent variable, although 2 items (1 and 17) have somewhat low correlations (<.20); the questionnaire sufficiently meets the mandatory of unidimensionality and complies with the local independence requirement.[32]

According to the questionnaire's general INFIT and OUTFIT values, the model´s fit was almost perfect. Only item 14 has a lower ZSTD value. This suggested a possible random selection of responses. Rather than eliminating it, we studied their content and congruence with the item's construct and improves wording.

Wright's map indicates that there is an adequate fit between the difficulty of the items and the ability levels of the respondents, such that it is possible to correctly discriminate between the different levels without underestimating or overestimating anyone.

The advantages of using this questionnaire, in terms of its characteristics, are that it is easily understandable; the set of items is adjusted to the population examined, there being no items associated with total failure, it can discriminate different levels of ability in the respondents and its reliability is outstanding since it far exceeds the minimum necessary criteria.[21,33] Consequently, the questionnaire is valid and reliable and can be used for the purposes intended by Spanish speakers.

As a limitation, we could improve some of the worst-fitting items (items 1, 14, and 17) in future developments. We should also point out as a limitation the fact that the sample was not random.

14. Conclusion

The questionnaire is easily understandable and can discriminate between different ability levels without differential gender bias. It is reliable and outstanding, as it exceeds the minimum criteria. The readability of the questionnaire is suitable for the level of potential users. The version presented is valid and reliable and can be used for the designed purpose with guarantees.

Acknowledgements

The authors would like to thank the teachers of the Physical Education Department, the management team, and the School Council of IES José Zerpa, Santa Lucía de Tirajana, Gran Canaria, Canary Islands, Spain, for their collaboration. They also thank Wiley Editing Services for their assistance in preparing this article.

Author contributions

Conceptualization: Fernando Calvo-Francés, José Enrique Hernández-Rodríguez.
Formal analysis: Fernando Calvo-Francés, José Enrique Hernández-Rodríguez.
Investigation: Fernando Calvo-Francés, Lucía Cilleros-Pino, Paula Andrea Barraza-Illanes, Maximino Díaz-Hernández, José Enrique Hernández-Rodríguez.
Methodology: Fernando Calvo-Francés, José Enrique Hernández-Rodríguez.
Project administration: José Enrique Hernández-Rodríguez.
Resources: Maximino Díaz-Hernández, José Enrique Hernández-Rodríguez.
Validation: José Enrique Hernández-Rodríguez.
Visualization: Paula Andrea Barraza-Illanes, José Enrique Hernández-Rodríguez.
Writing – original draft: Fernando Calvo-Francés, Paula Andrea Barraza-Illanes, José Enrique Hernández-Rodríguez.
Writing – review & editing: Fernando Calvo-Francés, Lucía Cilleros-Pino, Paula Andrea Barraza-Illanes, Maximino Díaz-Hernández, José Enrique Hernández-Rodríguez.

References

[1] Roger VL, Go AS, Lloyd-Jones DM, et al. The American heart association statistics committee and stroke statistics subcommittee. Executive summary: heart disease and stroke statistics--2012 update: a report from the American heart association. Circulation. 2012;125:188–97.
[2] Yan S, Gan Y, Jiang N, et al. The global survival rate among adult out-of-hospital cardiac arrest patients who received cardiopulmonary resuscitation: a systematic review and meta-analysis. Crit Care. 2020;24:61.
[3] Prado F. Incidence of resuscitation maneuvers by a bystander on the prognosis of patients suffering cardiac arrest on public roads and premises in Madrid. [Incidencia de las maniobras de reanimación por un testigo en el pronóstico de los pacientes que sufren un paro cardíaco en vías y locales públicos de Madrid]. PhD thesis [Tesis doctoral]. Madrid: Universidad Complutense; 2017. Available at: https://eprints.ucm.es/41539/ [access date September 23, 2019].
[4] Bhanji F, Donoghue AJ, Wolff MS, et al. Part 14: education: 2015 American heart association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation. 2015;132(18 Suppl 2):561–73.
[5] Soar J, Maconochie I, Wyckoff MH, et al. International consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. Resuscitation. 2019;145:95–150.
[6] García GR, Cerdà M. Teaching cardiopulmonary resuscitation to the population: One of the pillars to improve the survival of patients in cardiac arrest. [Enseñanza de la reanimación cardiopulmonar a la población: Uno de los pilares para mejorar la supervivencia de los pacientes en paro cardíaco]. Med Clín (Barc). 2005;124:13–5.
[7] Yin G, Chen L, Wu Y, et al. The implementation of a community-centered first aid education program for older adults – community health workers perceived barriers. BMC Health Serv Res. 2023;23:1–9.
[8] Perkins GD, Graesner JT, Semeraro F, et al. European resuscitation council guidelines 2021: executive summary. Resuscitation. 2021;9–13.
[9] Cheng A, Bhanji F, Bigham BL, et al. Part 6: resuscitation education science. 2020 American heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation. 2021;142:S551–79.
[10] Jones I, Whitfield R, Colquhoun M, et al. At what age can schoolchildren provide effective chest compressions? An observational study from the Heartstart UK schools training program. BMJ. 2007;334:1201.
[11] Rajacich D, Khasawneh J, Cameron S, et al. Development of human resources in nursing: a collaborative initiative in CPR. J Contin Educ Nurs. 2001;32:27–30.
[12] Xanthos T, Ekmektzoglou KA, Bassiakou E, et al. Nurses are more efficient than doctors in teaching basic life support and automated external defibrillator in nurses. Nurse Educ Today. 2019;29:224–31.
[13] Martínez-Isasi S, García-Suárez M, De La Peña Rodríguez MA, et al. Basic life support training program in schools by school nurses: how long and how often to train? Medicine (Baltim). 2021;100:e24819.
[14] Miró O, Díaz N, Escalada X, et al. Review of the initiatives carried out in Spain to implement the teaching of basic cardiopulmonary resuscitation in schools. [Revisión de las iniciativas llevadas a cabo en España para implementar la enseñanza de la reanimación cardiopulmonar básica en las escuelas]. An Sist Sanit Navar. 2012;35:477–86. Available at: https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1137-66272012000300014 [access date September 23, 2019].
[15] Miró O, Escalada X, Jiménez-Fábrega X, et al. Cardiopulmonary resuscitation program for secondary schools (PROCES): analysis of the pilot study. [Programa de Reanimación Cardiopulmonar orientado a centros de Enseñanza Secundaria (PROCES): análisis del estudio piloto]. Med Clin (Barc). 2005;124:4–9. Available at: https://www.elsevier.es/es-revista-medicina-clinica-2-articulo-programa-reanimacion-cardiopulmonar-orientado-centros-13070434 [access date September 23, 2019].
[16] López MC, Freire M, Rasines R, et al. CPR in the classroom: basic cardiopulmonary resuscitation teaching program for high school students. [RCP na aula: programa de enseñanza de la reanimación cardiopulmonar básica para estudiantes de secundaria]. Emergencias. 2012;24:76. Available at: http://emergencias.portalsemes.org/descargar/rcp-na-aula-programa-de-ensenanza-de-la-reanimacion-cardiopulmonar-basica-para-estudiantes-de-secundaria/ [access date September 23, 2019].
[17] SAMUR-Civil Protection of the Madrid City Council. Citizen training. Alert, Assess and Act program. [SAMUR-Protección Civil del Ayuntamiento de Madriid.Formación ciudadana. Programa alertante. Alerta, Valora y Actúa]. V1.2017-2018. Available at: https://www.bandomovil.com/userFiles/4S/4SMmU230920140755_Programa-Alertante.pdf [access date September 23, 2019].
[18] Díaz MA, Fernández A, Díaz A, et al. Influence of sociodemographic factors on the acquisition of basic life support skills. [Influencia de los factores sociodemográficos en la adquisición de habilidades en soporte vital básico]. Emergencias. 2014;26:202–5. Available at: http://emergencias.portalsemes.org/descargar/influencia-de-los-factores-sociodemograficos-en-la-adquisicion-de-habilidades-en-soporte-vital-basico/force_download/ [access date September 23, 2019].
[19] Cheng YH, Yeung CY, Sharma A, et al. Non-resuscitative first aid training and assessment for junior secondary school students: a pre-post study. Medicine (Baltim). 2021;100:e27051.
[20] Fonseca Del PFJ, Valle AJ, Canales VNB, et al. Basic life support knowledge of secondary school students in cardiopulmonary resuscitation training using a song. Int J Med Educ. 2016;7:237–41.

[21] García del ÁJJ, López E, Escamilla R, et al. Training of school children in basic life support by their teachers. [Formación de escolares en soporte vital básico por sus propios profesores]. Emergencias. 2019;31:185–8. Available at: http://emergencias.portalsemes.org/descargar/formacin-de-escolares-en-soporte-vital-bsico-por-sus-propios-profesores/force_download/ [access date September 23, 2019].
[22] Tobase L, Peres HHC, Tomazini EAS, et al. Basic life support: evaluation of learning using simulation and immediate feedback devices. Rev Lat Am Enfermagem. 2017;25:e2942.
[23] Carretero-Dios H, Pérez C. Normas para el desarrollo y revisión de estudios instrumentales: consideraciones sobre la selección de tests en la investigación psicológica. Int J Clin Health Psychol. 2007;7:863–82.
[24] Montero I, León O. A guide for naming research studies in psychology. Int J Clin Health Psychol. 2007;7:847–62.
[25] Hidalgo-Montesinos MD, French BF. A didactic introduction to item response theory for understanding scale construction. [Una introducción didáctica a la Teoría de Respuesta al Ítem para comprender la construcción de escalas]. Revista de Psicología Clínica con Niños y Adolescentes. 2016;3:13–21. Available at: https://www.revistapcna.com/sites/default/files/16-11.pdf [access date November 20, 2019].
[26] Templin J, Henson RA. Measurement invariance within and between persons: a distinct problem in testing the equivalence of mental constructs. Multivariate Behav Res. 2019;54:327–41.
[27] Rijmen F, Tuerlinckx F, De Boeck P, et al. (2020). Investigating the influence of time-varying item features on emotion measurement using dynamic item response theory models. Emotion. 2020;20:987–1002.
[28] Nunnally JC. Psychometric theory [Teoría psicométrica]. (2ª ed). Nueva York: McGraw-Hill; 1978.
[29] Tristán-López A. Modification of Lawshe's model for the content validity of an objective instrument. [Modificación al modelo de Lawshe para el dictamen de validez de contenido de un instrumento objetivo]. Avances en Medición. 2008;6:37–48. Available at: https://studylib.es/doc/6257335/modificaci%C3%B3n-al-modelo-de-lawshe-para-el-dictamen-cuantit [access date October 23, 2019].
[30] Barrio IM. Readability and Health. Legibility measurement methods and their application to the design of health education brochures. [Legibilidad y Salud. Los métodos de medición de la legibilidad y su aplicación al diseño de folletos educativos sobre salud]. PhD thesis [Tesis doctoral]. Servicio de Publicaciones de la Universidad Autónoma de Madrid; 2007. Available at: https://repositorio.uam.es/handle/10486/2488 [access date October 23, 2019].
[31] Muñoz A, Muñoz J. Readable. Text readability analyzer [Legible. Analizador de legibilidad de texto]. 2020. Available at: https://legible.es/ [access date November 20, 2019].
[32] Linacre JM, Wright BD. WINSTEPS: Multiple-Choice, Rating Scale, and Partial Credit Rasch Analysis. Chicago, IL: MESA Press; 2009.
[33] Linacre JM. Winsteps® (Version 4.4.4) [Computer Software]. Beaverton, OR: Winsteps.com; 2019. Available at: https://www.winsteps.com [access date November 20, 2019].
[34] Arias B, Arias V, Gómez LE. Calibration of the conners hyperactivity index using the rasch model. [Calibración del índice de Hiperactividad de Conners mediante el modelo de Rasch]. Universitas Psychologica. 2013;12:957–70.
[35] Junco JI, Prieto G. Analysis of the neuropsychological test addenbrooke's cognitive examination using the rasch model. [Análisis del test neuropsicológico Addenbrooke's Cognitive Examination mediante el Modelo de Rasch]. Revista de Psicología. 2014;23:40–52.
[36] Prieto G, Delgado AR. Analysis of a test using the Rasch model [Análisis de un test mediante el modelo de Rasch]. Psicothema. 2003;15:94–100. Available at: http://www.psicothema.com/pdf/1029.pdf [access date November 20, 2019].