# Cross-Cultural Adaptation and Psychometric Evaluation of the Educational Content Validation Instrument in Health

Sergio Cazorla-Calderón, PhDc RN[1] ⓘ, José Manuel Romero-Sánchez, PhD RN[1,2] ⓘ, Elena Fernández-García, PhD RN[3] ⓘ, and Olga Paloma-Castro, PhD RN[1,2] ⓘ

## Abstract

Healthcare professionals often use multimedia patient education media, but not all have the same content quality. This study aimed to cross-culturally adapt the Educational Content Validation Instrument in Health to the Spanish setting and assess its psychometric properties. A methodological validation study was carried out between January and September 2020. Data collection took place from May to June 2020. A translation, back translation, committee review, and pre-testing was carried out. Subsequently, reliability (internal consistency), and validity (factorial and convergent) were assessed by requiring 210 Healthcare Professionals to complete the instrument based on video material. In addition, a refinement of the instrument was conducted based on the modification indexes. The instrument showed adequate internal consistency, although some redundancy in the items ($\alpha = .93$). Exploratory factor analysis suggested a unifactorial structure that explained an adequate variance (47.37%). Convergent validity was poor ($r = .11$; $P = .05$). After analysis, 6 items were deleted without impairing the validity results and eliminating redundancy. Therefore, a 12-item version of the instrument was created. It can help to assess more objectively the contents of the materials prescribed, facilitating the choice of those of higher quality and potentially improving their patients' health outcomes. Further studies are needed to confirm the previous results and reassess some of the shortcomings.

## Keywords

patient education, patient education handout, multimedia, instrument development, psychometrics, questionnaires, reliability and validity, Spain

---

**What Do We Already Know about This Topic?**
Although several instruments have been developed in the form of checklists to assess various aspects related to the content of PEM, none of these instruments found in a comprehensive literature search has been cross-culturally adapted and psychometrically tested in the Spanish setting.

**How Does Your Research Contribute to the Field?**
This study provides preliminary evidence on the reliability and validity of a 12-item cross-culturally adapted version of the Educational Content Validation Instrument in Health for health professionals to evaluate the content of video-based PEM for its use with patients in the Spanish context.

**What are Your Research's Implications Toward Theory, Practice, or Policy?**
The availability of a 12-item cross-culturally adapted of the Educational Content Validation Instrument in Health could be helpful for Spanish health professionals assisting them in selecting the PEM with the best possible content to offer their patients.

---

## Main Text

### Introduction

Patients receive large amounts of new information at a single consultation.[1] Healthcare professionals (HCPs) currently rely heavily on verbal communication and assume that patients have understood the information they provide fully.[2] However, they often use technical jargon and provide patients with very complex information in a short time, which makes it difficult to understand and memorize.[2] This situation leads to 40-80% of verbal information received after a formal consultation being immediately forgotten,[3] and almost half of the information patients recall is incorrect.[4] Given the difficulties in the spoken communication of health information, it is common to use patient education media (PEM) to complement it.

Video-based PEM provides visual and auditory information and is easier to absorb for patients than other patient educational methods[2] and thus could overcome barriers related to limited literacy and hearing impairment, respectively.[5] Thus, video-based PEM has become increasingly commonly used among healthcare practitioners.[6] However, it is logical to think that the content of all of them may not be of the same quality. PEM should meet relevant aspects of educational content, which are essential to fulfill the purposes of these materials and provide education as efficiently as possible.[7]

For this, several instruments have been developed in the form of checklists to assess various aspects related to the contents of PEM.[5,6,8,9] Among these tools, *The Patient Education Materials Assessment Tool (PEMAT)*[10] uses a systematic method to evaluate and compare the understandability and actionability of patient education materials, or the *Suitability Assessment of Materials (SAM)* instrument offers a systematic method to objectively assess the suitability of health information materials. Other instruments were developed to assets PEM in a specific format, such as the *Tool to Evaluate Materials Used in Patient Education (TEMPtEd),* established to help HCPs evaluate and select PEM in printed format but not in multimedia format. The Educational Content Validation Instrument in Health (ECVIH), a tool created by a group of Brazilian researchers to assist HCPs in assessing health education content in 3 content-related dimensions: objectives, structure and presentation and relevance.[7] It is also an instrument with few items and a simple administration and scoring system.

However, none of the instruments listed above or others found in the comprehensive literature search has been cross-culturally adapted and psychometrically tested in the Spanish setting.

Given the characteristics of brevity and ease of scoring and application of ECVIH, we believe it would be helpful for Spanish health professionals adapt the instrument that would assist them in selecting the PEM with the best possible content to offer their patients. Therefore, this study aimed to cross-culturally adapt the Educational Content Validation Instrument in Health to the Spanish setting and assess its psychometric properties.

## Methods

### Design

A methodological validation study was carried out.

### Cross-Cultural Adaptation

The guidelines proposed by Guillemin et al.[11] were used, slightly modified to adapt them to the purposes of the study to preserve equivalence in the cross-cultural adaptation of the instrument. The following phases were included:

(1) *Translation.* Three independent translations to Spanish were made of the original English instrument published by Leite et al.[7] Two translators were nurses and were warned of the objectives and relative concepts of the material to be translated. A third translator was a psychologist who was not warned to elicit unexpected meanings from the original instrument. All of them were bilingual, English, and Spanish; their mother tongue is the latter. The translators discussed the 3 versions and studied the discrepancies until a consensus version was reached.

(2) *Retro-translation.* The consensus version was retro-translated into English by a professional translator with a stable residence in Spain whose mother tongue is English and unfamiliar with the material.

(3) *Review.* A committee of experts was set up to compare the retro-translation with the original version of the instrument. This committee consisted of the previous translators together with the principal author of the original instrument. The decentering technique was used in which the items with problems were located, and the author provided a working version of them, maintaining the concept. On the other hand, given that the original tool was designed to assess written PEM and this study sought to validate it using the material in

[1]Research Group PAIDI CTS-1019 "MELES - Nursing Methods and Standard Languages", Universidad de Cádiz, Cádiz, Spain
[2]Faculty of Nursing and Physiotherapy, Nursing and Physiotherapy Department, Universidad de Cádiz, Cádiz, Spain.
[3]Faculty of Nursing, Physiotherapy and Podiatry,. Nursing Department, Universidad de Sevila, Sevilla, Spain.

**Corresponding Author:**
José Manuel Romero-Sánchez, Universidad de Cádiz. Facultad de Enfermería y Fisioterapia, Avda. Ana de Viya, 52; 11009 Cádiz (Spain).
Email: josemanuel.romero@uca.es

audio-visual format, it was agreed with the author to replace item 15, "The text size is adequate" with "The narrative/dialogues are heard clearly." Once the experts approved the back-translation, the original and the Spanish adaptation were given to 5 bilingual nurses to detect discrepancies.

## Participants

The sample comprised HCPs meeting the following inclusion criteria: (1) Hold a bachelor's degree in health science that qualifies them to prescribe PEM to patients and (2) practice or have practised the profession for which they are qualified in Spain. Visual and/or auditory impairments that prevent the viewing of multimedia material were only considered as an exclusion criterion.

This convenience sample was selected through non-probabilistic network sampling[12] to obtain subjects with diverse professional profiles and settings. The 10:1 ratio rule, ten participants per item, was adopted to determine a sufficient sample size.[13–15] Considering that the instrument to be psychometrically tested contains 18 items, the sample size would be set at 180 subjects. However, current trends recommend a sample size of at least 200 cases, even under optimal conditions of high communalities and well-defined factors.[16] Therefore, an attempt was made to maximize the sample size and thus obtain a higher number of subjects than initially proposed.

A panel of experts was also selected intentionally to enable the assessment of the convergent validity of criteria. It consisted of 9 people following the recommendation of *Polit and Tatano*.[17] To belong to the panel, they had to meet the eligibility criteria of the above sample and have proven experience, having developed at least 5 PEM in the last 2 years.

## Instruments

### The following instruments were used

(1) *Socio-demographic and professional data questionnaire.* Consisted of questions to assess the following variables: age (years), gender (male, female), healthcare professional profile (assistance, teaching, and research, management), educational profile (nurse, physician, psychologist), professional experience (years), academic degree (diploma, bachelor or graduate degree, official master's degree, doctorate).

(2) *Specific questions for expert's questionnaire.* Applied to experts only together with the other instruments. It consisted of questions to assess the following variables: format of the PE material created by them (written, multimedia), PEM created in the last 2 years in each format (number), and assessment of the contents of the patient education material shown assessed with a single question formulated as, "Rate the overall content of the material displayed to educate the patient on (subject of the video, in this case, Basic life support (BLS) from 0 to 10" and with a 10-point Likert-type measurement scale where 0 indicates that the overall content is "Very bad" and 10 "Very good."

(3) *Spanish version of the Educational Content Validation Instrument in Health (ECVIH-SV).* Developed for this study in the cross-cultural adaptation process described above. The original instrument was developed based on a narrative review of the literature and expert consensus.[7] It has 18 items allocated theoretically into 3 domains: (1) objectives (items 1-5), relating to purpose, goals, or target the use of educational materials; (2) structure and presentation (items 6-15), relating to general organization, structure, strategy, consistency, and sufficiency of the presentations; and (3) relevance (items 16-18), relating to the significance degree of the educational content and its ability to cause impact, motivation, and/or interest. The items were applied to the PEM to be assessed, using a three-point Likert response scale where 0 = disagree, 1 = partially agree, and 2 = totally agree. According to experts, the items obtained more than 80% agreement and good internal consistency in their domains in the content validation. ECVIH items are presented in Online Appendix 1.

## Data Collection

The authors designed 2 non-validated anonymous electronic surveys to collect the data specified in the previous section, one for the experts' sample and the other for the general sample.

Participants of both experts and general samples were contacted by mail, giving information about the purpose of the study, the identity of the investigators, and the time investment for the survey. To fill in the survey, access to the website was through a link to the survey, password-protected to prevent malicious access. Password was sent together with the invitation email. The electronic survey system prevented multiple entries from the same individual through email addresses. The anonymity of the responses was ensured. No advertising was used.

Both surveys were designed using the commercial web survey provider www.surveymonkey.com. The account was only accessible to the leading investigator for data protection purposes. The web survey was tested before the start of the study by colleagues. Questions asked to all participants were mandatory (an automatic option in the survey program). Respondents could review and change existing answers until the survey was finished through a "back button."

PEM to be assessed is needed to apply the instrument. In this case, 3 PEM were included in the surveys, precisely 3 different videos on the same topic, "how to perform Basic Life Support with an automatic defibrillator," aimed at the

general public. These videos were pre-selected by the research team based on their opinion and regarding their quality, looking for differences between them. The authors of the videos that were not classified as free use were asked for authorization.

Each participant in the general sample was required to use the ECVIH instrument to assess the content of only one of these videos assigned randomly by the Web application. Video 1 was assessed by 83 participants (39.5%), video 2 by 70 (33.3%), and video 3 by 57 (27.1%). In contrast, experts assessed all 3 videos since their assessment was necessary to assess convergent validity. The videos were shown to experts in random order to prevent the learning effect in the application of the instrument from constituting a bias in the scores.

Data collection was conducted from May to June 2020.

## Ethical Considerations

This study is part of a larger project approved by the Andalusian Coordinating Committee on Ethics in Biomedical Research (PI-0019-2017). Participation was voluntary, and the anonymity of the participants was guaranteed, so the web application did not record any data that could be used to identify them. The participants were informed of the purpose of the study and the processing of the data. They also signed a written consent form before filling the survey. Written permission was obtained from the authors of the instrument for its cross-cultural adaptation and validation.

## Data Analysis

The normality of the variables was checked using the Kolmogorov-Smirnov-Lilliefors test.[18] Data relating to the distribution scores and the sample characteristics were summarized using the descriptive statistics, the mean and standard deviation for the quantitative variables, while frequency and percentage were used for the categorical ones.

The reliability assessment included tests of internal consistency: (a) Cronbach's alpha ($\alpha$) for the complete instrument and when each item was removed. A value of $\alpha$ of .70-.90 was considered appropriate and that there were no substantial changes in value when the item was excluded; (b) item-total correlation corrected using the Pearson correlation coefficient. An item-total correlation of $\geq$.40 was considered appropriate.[19]

The convergent validity of the instrument was assessed using the assessment of the overall contents of the PEM by the experts. Given that no instrument validated in Spain that measures the same construct was found in the literature search, an ad hoc question was used to assess the content in general terms. The total scores on the instrument were correlated with the averages obtained in the previous question in each of the 3 videos shown to the experts through Pearson's test. A correlation coefficient ranging from .30-.40 indicates a high convergent validity.[13]

The construct validity was assessed in terms of factorial validity through exploratory factor analysis (EFA) using principal component analysis (PCA). An EFA was carried out using the main components analysis method. As well as conceptual interpretability, the following criteria were used to identify the optimal number of factors: (a) Scree plot[20] determined the point where the slope of the curve is leveling off;[21] (b) explained variance of at least 40%[22] and that each factor contains at least 3 items.[23] The factorial matrix was examined, extracting those items that load unambiguously on the extracted factors, considering those with factor loading >.50 on exactly 1 factor and <.40 on all other factors.[24] Data were checked beforehand for suitability for factor analysis by examining Kaiser–Meyer–Olkin (KMO) indices test of sampling adequacy and Bartlett's test of sphericity (BTS). A value of KMO <.50 and a BTS significance level of $P < .05$ was being considered appropriate.[25]

Once the optimal factor solution according to the EFA was obtained, a confirmatory factor analysis (CFA) was carried out to determine whether the data fit the proposed model. The following adjustment indices were used: (a) Significance of chi-square ($\chi2$), being desirable a $\chi2$ not significant for the model;[26] (b) comparative fit index (CFI) and Tucker–Lewis Index (TLI) for which values greater than .90 indicate a good fit, while values greater than .95 indicate an excellent fit[27,28] and (c) root mean square error of approximation (RMSEA) for which values below .08 indicate an acceptable model fit and .06 an excellent fit, such as at its upper limit of the 90% confidence interval.[29] If the model did not fit the data obtained in the sample, refining the instrument by studying the discrepancies between the model and the data through the modification indexes (MIs) was considered.[30] The MIs indicate the amount of decrease in the $\chi2$ of the model with a degree of freedom. A high MI value indicates that the relevant fixed-parameter should be released to improve the model fit.[26] To free it, it was decided to eliminate the specific elements with higher MIs, which meant an improvement in the model fit, with a decrease of $\chi2$ and a substantial improvement of the other adjustment indices.[31] The elimination was carried out one by one, starting with the one with the highest MI and so on successively since changing a single parameter in a model could affect other parts of the solution.[32] After each elimination, the CFA was repeated on the remaining elements, assessing the specific adjustment rates. These steps were repeated progressively until the indices showed a good model fit according to the criteria specified in the previous paragraph.

Once the items that comprised the definitive model were retained, the reliability and convergent validity tests were repeated to check if they also obtained similar or better psychometric properties in addition to their dimensionality.

The results were considered statistically significant if the P-values were <.05. Data were analyzed using IBM SPSS Statistics for Windows, version 25, and IBM AMOS, version 23 (IBM Corp, Armonk, NY, USA).

## Results

### Cross-Culturally Adapted Instrument

The translations made independently by the translators obtained complete coincidence in 3 items, partial coincidence in 4, and discrepancy in eleven. Partial coincidences were solved by choosing the translation of the 3 that best represented the item, with each voting for the 1 that seemed correct. Translations of dissenting items were resolved by joint discussion, obtaining a consensus translation for each item. The retro-translation was revised by the translators and one of the authors of the original instrument. She only adjusted 3 items. The translations of these were modified concerning their contributions and sent back for comparison, after which the translation was approved. The bilingual nurses who reviewed the Spanish and the original instruments endorsed the equivalence of both versions. The Spanish version of the ECVIH is given in Online Appendix 1.

### Description of the Participants

Of the 600 questionnaires sent out, 222 were returned correctly filled in, implying a response rate of 37%. Of these, 12 were eliminated as they did not meet the inclusion criteria (8 were non-university senior healthcare technicians, and 4 were non-graduate nursing students). Therefore, the final sample consisted of 210 participants, most of whom were women (55.5%) with an average age of 40.8 years (SD = 10.41). Most of the sample were nurses (79%), working mainly in the healthcare field (80%) with an accumulated professional experience of 16.53 years (SD = 9.95). Concerning training, 52.6% of the participants had some postgraduate university education. The participants' demographic and professional properties are shown in detail in Table 1.

Regarding the experts, 8 were nurses (88.9%) and 1 was a physician. 77.8% had postgraduate training. Most of them had healthcare functions (88.9%) and an accumulated professional experience of 18 years (SD = 6.46). Most of them had experience in creating patient education materials in both written and multimedia formats (77.8%), the rest in written format only, with an average of 19.44 materials in written format (SD = 19.70) and 22.00 in multimedia format (SD = 29.26).

### Score Distribution

The average scale score for the sample was 29.88 (SD = 2.77) out of a maximum of 36 points. The items had average scores in the range of 1.33-1.87. The means and standard deviations of each item are shown in Table 2.

### Reliability

*Internal consistency.* The α value for the instrument was .93. No item significantly increased or decreased the α when it was removed. All items reached good item-total correlations

**Table 1.** Participants' Demographic and Professional Characteristics.

| | N | % | M | SD |
|---|---|---|---|---|
| **Sex** | | | | |
| Women | 116 | 55.5% | | |
| Men | 91 | 43.5% | | |
| N/A | 2 | 1% | | |
| Age (years) | | | 40.8 | 10.41 |
| **Field of work** | | | | |
| Health care | 168 | 80% | | |
| Managing | 13 | 6.2% | | |
| Professor and/or researcher | 29 | 13.8% | | |
| **Occupation** | | | | |
| Nurse | 166 | 79% | | |
| Physician | 37 | 17.6% | | |
| Psychologist | 5 | 2.4% | | |
| Other | 2 | 1% | | |
| Professional experience *(years)* | | | 16.53 | 9.95 |
| **Academic degree** | | | | |
| Bachelor | 99 | 47.4% | | |
| Master | 87 | 41.6% | | |
| PhD | 23 | 11% | | |

M, mean; SD, standard deviation.

(.41-.76), except item 14, which obtained a low correlation (.28). These results are shown in detail in Table 2.

### Validity

*Convergent validity.* A poor, though significant, correlation (r = .11; P = .05) was found between the total instrument scores and the average of the scores given by the experts for each video.

*Factorial construct validity.* The Bartlett sphericity test ($\chi^2$ = 2125.66; df = 153; P < .001) reached statistical significance, and the KMO value was adequate (KMO = .92). The preliminary PCA revealed the presence of 3 components with eigenvalues greater than 1, which explained the 47.37%, 7.30%, and 5.99% of the variance, respectively. However, an inspection of the scree plot indicated a significant slope after the first component, followed by a clear plateau. The factorial matrix was examined, determining that all but 3 items had high factorial loadings in the first component (.45-.86) and low ones in the rest (<.40). Items 16 and 18 load above .40 on both the first and second components, although the loadings on the first were significantly higher than on the second (item 16: .69 vs .44; item 18: .71 vs .41). Item 14 loads high on the second (.58) and third component (.53). Considering the scree plot, the high factor loading of almost all items on the first component, the acceptable variance explained by it, and that the other 2 components did not satisfy the condition, at least 3 items load the highest on that factor. It was decided to carry out a new PCA by forcing the

**Table 2.** Item Scores, Reliability Results, and Factor Loadings of ECVIH.

| Items | Item Score | | Total Score if Item Deleted, M | Corrected Item-Total Correlation | Cronbach's Alpha if Item Deleted | Factor Loadings |
|---|---|---|---|---|---|---|
| | M | SD | | | | |
| 1. Contemplates the proposed theme | 1.84 | 0.41 | 28.05 | 0.54 | 0.93 | 0.59 |
| 2. Suits the teaching-learning process | 1.59 | 0.57 | 28.30 | 0.70 | 0.93 | 0.74 |
| 3. Clarifies doubts on the addressed theme | 1.42 | 0.65 | 28.47 | 0.61 | 0.93 | 0.66 |
| 4. Provides reflection on the theme | 1.56 | 0.65 | 28.32 | 0.67 | 0.93 | 0.72 |
| 5. Encourages behavior change | 1.57 | 0.64 | 28.31 | 0.69 | 0.93 | 0.74 |
| 6. Language appropriate to the target audience | 1.66 | 0.58 | 28.23 | 0.74 | 0.92 | 0.78 |
| 7. Language appropriate to the educational material | 1.70 | 0.56 | 28.19 | 0.71 | 0.93 | 0.76 |
| 8. Interactive language, enabling active involvement in the educational process | 1.33 | 0.74 | 28.55 | 0.60 | 0.93 | 0.64 |
| 9. Correct information | 1.59 | 0.61 | 28.30 | 0.63 | 0.93 | 0.68 |
| 10. Objective information | 1.72 | 0.54 | 28.17 | 0.69 | 0.93 | 0.74 |
| 11. Enlightening information | 1.52 | 0.63 | 28.36 | 0.74 | 0.92 | 0.78 |
| 12. Necessary information | 1.87 | 0.39 | 28.02 | 0.61 | 0.93 | 0.67 |
| 13. Logical sequence of ideas | 1.71 | 0.55 | 28.17 | 0.69 | 0.93 | 0.73 |
| 14. Current theme | 1.86 | 0.40 | 28.02 | 0.28 | 0.93 | 0.32 |
| 15. Appropriate text size | 1.78 | 0.46 | 28.11 | 0.41 | 0.93 | 0.46 |
| 16. Encourages learning | 1.70 | 0.50 | 28.18 | 0.65 | 0.93 | 0.69 |
| 17. Contributes to knowledge in the area | 1.77 | 0.46 | 28.11 | 0.76 | 0.93 | 0.81 |
| 18. Arouses interest in the theme | 1.70 | 0.52 | 28.19 | 0.67 | 0.93 | 0.71 |

M, mean; SD, standard deviation.

extraction of a single component that explained a variance of 47.37%, with all items with high loadings on it (.46-.81) except item 14 that has a slightly lower loading on the component (.32) (Table 2). Therefore, the results of this second analysis support the one-dimensionality of the instrument.

The CFA carried out, assuming a one-factor model as extracted from the EFA, determined that this model did not fit well with the data obtained with poor adjustment indices: $X^2 = 474.89$ ($P < .0001$); $CFI = .83$; $TLI = .81$; $RMSEA = .11$ (90% $CI$:0.10-.12).

### Instrument Refinement

Based on these results, it was decided to refine the tool by re-estimating and reassessing it multiple times to adjust the data to the chosen model by studying the MIs changes by pairs of items. Items 6, 18, 5, 3, 10, and 14 were progressively eliminated until an appropriate fit of the data to the model was achieved, obtaining excellent fit indices with the remaining 12 items except for the $\chi^2$ ($\chi^2 = 101.21$ ($P < .0001$); $CFI = .96$; $TLI = .95$; $RMSEA = .06$ (90% $CI$:0.04-.08). Figure 1 shows the selected final model with its factor loadings. As can be seen, all the items have loadings higher than .50 except item 15, which loadings above .40. A 12-items ECVIH was obtained from this refinement process (Online Appendix 2).

Once the instrument had been refined, the rest of the psychometric properties were reassessed considering only the 12 items retained, with the following results. Regarding reliability in terms of internal consistency, the α for the 12-items instrument was .89. No item significantly increased or decreased the α when it was removed (.88-.90). Regarding the convergent validity, a poor correlation was again found between the total score of the 12-items instrument and the average of the scores given by the experts for each video, although more significant than for the 18-items ECVIH (r = .14; $P = .04$).

## Discussion

This study aimed to cross-culturally adapt and validate the Educational Content Validation Instrument in Health (EC-VIH). It is the first assessment of the psychometric properties of ECVIH on a sample of HCPs. The instrument, developed for the Brazilian context, was only previously assessed for content validity by experts in the same study that developed it.[7] This study, therefore, provides evidence of the internal properties of ECVIH when applied by HCPs on PEM, endorsing its use with a more excellent scientific basis.

The results endorse the reliability of the 18-items instrument in terms of internal consistency with an adequate psychometric performance of most items; however, the α value above .90 for the entire instrument exceeds the values
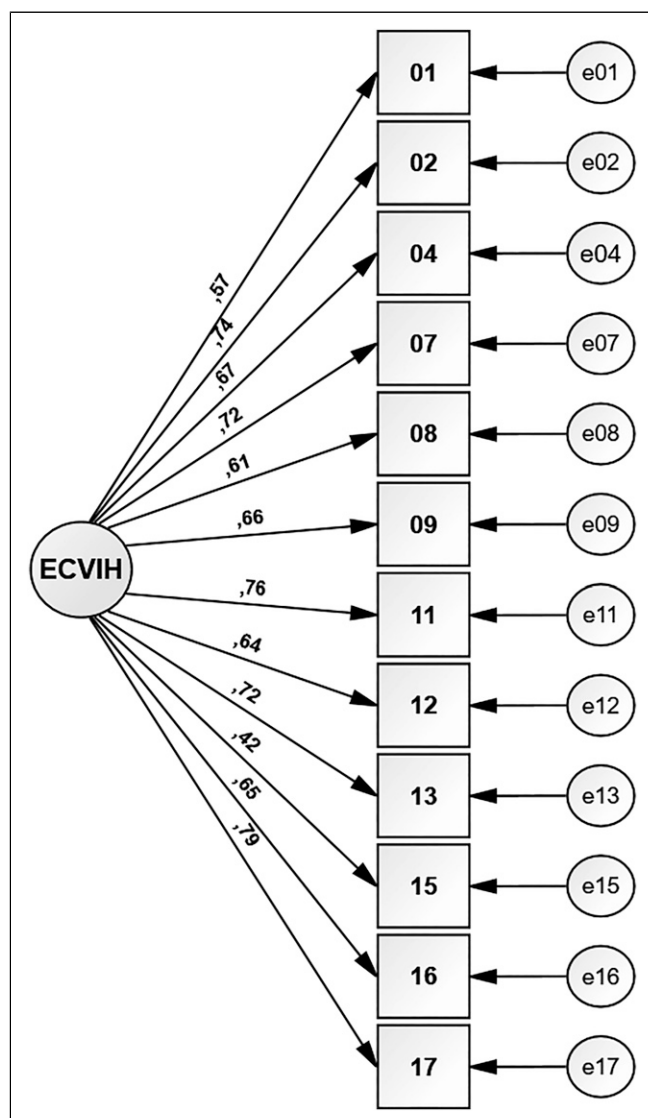
**Figure 1.** Items structure.

set as acceptable. Extremely high alpha values suggest redundancy between some of the items;[33] some items measure the same aspect of the construct.[34] In this case, it is usual to eliminate the redundancy by eliminating item.[17] In this study, items were eliminated based on the factorability of the instrument. However, even though the refinement was carried out based on a validity property, the re-assessment undertaken of the reliability of the 12-item ECVIH, all items retained or improved performance, and the value of alpha for the entire instrument decreased to optimal values, thus eliminating redundancy.

Although the items were allocated into 3 dimensions theoretically to structure the instrument in its creation,[7] at the empirical level, the scale seems to respond to a one-factor structure suggested by the EFA results, which accumulated adequate variance and factor loadings on the appropriate factor except for 1 item in the 18-items original instrument.

No good fit was obtained when the one-factor structure was tested through a CFA on the data for the 18 items. In the context of structural equation models, MIs are a measure often used to reduce discrepancies between model and data,[30] so it was decided to use them to refine the instrument. Based on these, a model 1 showing an excellent fit of all the calculated indices was obtained, except for the χ2, which is a parameter that is extremely sensitive to the sample size.[35] A reduced form was obtained from the process with a clear structure consisting of 1 factor and better reliability of the 12 items retained. It also reduced the instrument application times.

None of the versions, the original 18 or the refined 12-items version, gave good results for convergent validity. Among the reasons behind these negative findings is that the experts' opinion was collected with a single question developed ad hoc, which may not capture the entire construct.

There are various potential limitations of the study that need to be discussed. The intervention of a single translator in the back-translation process of the instrument made it impossible to implement a double-check mechanism. The use of a convenience sample limits the generalization of the results. Not all aspects of reliability or validity were assessed. Sensitivity to change was also not evaluated, a property that must be known if the instrument is to be used as a tool by the creators of health education material to assess improvements in that material's content.

Moreover, in this study, the CFA was carried out on the same data as the EFA, as was the determination of properties of the refined version of 12 items. For all the above reasons, it would be necessary to carry out future studies with more significant and representative samples, if possible, to mitigate these deficient aspects. Similarly, it would be necessary to determine whether the source version and other cross-cultural adaptations to other international contexts provide similar psychometric results. Finally, the educational material on which the validation was carried out was only of a multimedia type, and it would be necessary to assess the tool's performance on the material in written format if it is to be used on it.

## Conclusion

The results of this study provide preliminary evidence on the reliability and validity of ECVIH for health professionals to evaluate the content of video-based PEM for its use with patients in the Spanish context. A refined version of 12 items is proposed, which, in addition to improving the factorability of the original instrument, allows it to be completed in less time, improving or maintaining the psychometric properties of the original 18-items version. The availability of an instrument makes it easier for HCPs to assess the contents of the educational material they prescribe to patients with greater objectivity, facilitating the choice of the highest quality ones and potentially improving their patients' health outcomes. It

also allows the creators of PEM to assess the content of the material, allowing for improvement in those areas where HCPs perceive deficiencies.

## Author Contributions

Sergio Cazorla-Calderon: Conceptualization, investigation, resources, writing—original draft.
Jose Manuel Romero-Sanchez: Methodology, formal analysis, data curation, writing—original draft.
Olga Paloma-Castro: Conceptualization, resources, writing—review and editing, project administration, funding acquisition.
Elena Fernandez-Garcia: Investigation, writing—review and editing.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## ORCID iDs

Sergio Cazorla-Calderón  https://orcid.org/0000-0003-4965-6011
José Manuel Romero-Sánchez  https://orcid.org/0000-0001-8227-9161
Elena Fernández-García  https://orcid.org/0000-0002-7922-2663

## Supplemental Material

Supplemental material for this article is available online.

## References

1. Selic P, Svab I, Repolusk M, Gucek NK. What factors affect patients' recall of general practitioners' advice? *BMC Fam Pract*. 2011;12(1):141. doi:10.1186/1471-2296-12-141.

2. Abu Abed M, Himmel W, Vormfelde S, Koschack J. Video-assisted patient education to modify behavior: A systematic review. *Patient Educ Counsel*. 2014;97(1):16-22. doi:10.1016/j.pec.2014.06.015

3. Kessels RPC. Patients' memory for medical information. *J R Soc Med*. 2003;96(5):219-222. doi:10.1258/jrsm.96.5.219.

4. Jansen J, van Weert J, van der Meulen N, van Dulmen S, Heeren T, Bensing J. Recall in older cancer patients: Measuring memory for medical information. *Gerontol*. 2008;48(2):149-157. doi:10.1093/geront/48.2.149.

5. Farwana R, Sheriff A, Manzar H, Farwana M, Yusuf A, Sheriff I. Watch this space: A systematic review of the use of video-based media as a patient education tool in ophthalmology. *Eye*. 2020;34(9):1563-1569. doi:10.1038/s41433-020-0798-z.

6. Klein-Fedyshin M, Burda ML, Epstein BA, Lawrence B. Collaborating to enhance patient education and recovery. *J Med Libr Assoc: JMLA*. 2005;93(4):440-445.

7. Leite SDS, Áfio ACE, Carvalho LVD, Silva JMD, Almeida PCD, Pagliuca LMF. Construction and validation of an educational content validation instrument in health. *Rev Bras Enferm*. 2018;71(suppl 4):1635-1641. doi:10.1590/0034-7167-2017-0648.

8. Ferguson LA. Implementing a video education program to improve health literacy. *J Nurse Pract*. 2012;8(8):e17-e22. doi:10.1016/j.nurpra.2012.07.025.

9. Freda MC. Issues in patient education. *J Midwifery Wom Health*. 2004;49(3):203-209. doi:10.1016/j.jmwh.2004.01.003.

10. Shoemaker JS, Wolf MS, and Brach C. *The Patient Education Materials Assessment Tool (PEMAT) and User's Guide*. Rockville: Agency for Healthcare Research and Quality U.S.: 1-5. https://www.ahrq.gov/professionals/prevention-chronic-care/improve/self-mgmt/pemat/index.html (2017).

11. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417-1432. doi:10.1016/0895-4356(93)90142-n.

12. Clayton LH. TEMPtEd: Development and psychometric properties of a tool to evaluate material used in patient education. *J Adv Nurs*. 2009;65(10):2229-2238. doi:10.1111/j.1365-2648.2009.05049.x.

13. Nunnally JC, Bernstein IH. *Psychometric Theory*. New York: McGraw-Hill; 1994.

14. Thorndike RL. *Applied Psychometrics*. Boston: Houghton-Mifflin; 1982.

15. MacCallum R, Widaman K, Zhang S, Hong S. Sample size in factor analysis. *Psychological Methods - Psychol Methods*. 1999;4:84-99. doi:10.1037/1082-989X.4.1.84.

16. Ferrando PJ, Lorenzo-Seva U. El análisis factorial exploratorio de los ítems: algunas consideraciones adicionales. *An Psicolog* 2014;30(3):1170-1175. doi:10.6018/analesps.30.3.199991.

17. Polit DF, and Beck CT. *Nursing Research: Principles and Methods*. Philadelphia: Lippincott Williams & Wilkins. https://books.google.es/books?id=5g6VttYWnjUC (2004).

18. Lilliefors HW. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J Am Stat Assoc*. 1967; 62(318):399-402. doi:10.1080/01621459.1967.10482916.

19. Norman GR, and Streiner DL. *Biostatistics: The Bare Essentials*. Hamilton: B.C. Decker. https://books.google.es/books?id=8rkqWafdpuoC (2008).

20. Cattel RB. The Scree Test For The Number of Factors. *Multivariate Behavioral Research* 1996;1(2):245-276. doi:10.1207/s15327906mbr0102_10. In press.

21. Kim J, Mueller CW. *Factor Analysis: Statistical Methods and Practical Issues*. Thousand Oaks, CA, USA: Sage Publications; 1978.

22. Manzar D, Zannat W, Mohammed H, et al. Dimensionality of the Pittsburgh Sleep Quality Index in young collegiate adults. *SpringerPlus*. 2016;5. doi:10.1186/s40064-016-3234-x.

23. Netemeyer R, Bearden W, Sharma S. *Scaling Procedures. Issues and Applications*. Sage; 2003. doi:10.4135/9781412985772.

24. Farin E, Schmidt E, Gramm L. Patient communication competence: Development of a German questionnaire and correlates of competent patient behavior. *Patient Educ Counsel*. 2014;94(3):342-350. doi:10.1016/j.pec.2013.11.005.

25. Denise F. Polit. *Statistics and Data Analysis for Nursing Research*. 2nd ed.. Upper Saddle River: Pearson; 2010.

26. Wang J, Hefetz A, Liberman G. La aplicación del modelo de ecuación estructural en las investigaciones educativas. *Cultura y Educacion*. 2017;29(3):563-618. doi:10.1080/11356405.2017.1367907.

27. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107(2):238-246. doi:10.1037/0033-2909.107.2.238.

28. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973;38(1):1-10. doi:10.1007/BF02291170.

29. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model: A Multidisciplinary Journal*. 1999; 6(1):1-55. doi:10.1080/10705519909540118.

30. Sörbom D. Model modification. *Psychometrika*. 1989;54(3): 371-384. doi:10.1007/BF02294623.

31. Jöreskog KG, Olsson UH, Wallentin F Y. *Multivariate Analysis with LISREL*. Cham: Springer International Publishing; 2016. doi:10.1007/978-3-319-33153-9.

32. MacCallum RC, Roznowski M, Necowitz LB. Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychol Bull*. 1992;111(3):490-504. doi:10.1037/0033-2909.111.3.490.

33. DeVellis RF. *Scale Development: Theory and Applications*. SAGE Publications; 2016. https://books.google.es/books?id=48ACCwAAQBAJ.

34. Jaju A, Crask MR, Menon A, Sharma A. The Perfect Design: Optimization Between Reliability, Validity, Redundancy in Scale Items and Response Rates. In: *Winter Educators Conference*. Vol. 10. American Marketing Association; 1999: 127-131.

35. Wang J, Hefetz A, Liberman G. Applying structural equation modelling in educational research. *Cultura y Educacion*. 2017; 29:563-618. doi:10.1080/11356405.2017.1367907.

36. Helitzer D, Hollis C, Cotner J, Oestreicher N. Health literacy demands of written health information materials: An assessment of cervical cancer prevention materials. *Cancer Control*. 2009;16(1):70-78. doi:10.1177/107327480901600111.