European Journal of Social Work

ISSN: 1369-1457 (Print) 1468-2664 (Online) Journal homepage: http://www.tandfonline.com/loi/cesw20

Content validity analysis of ISD-1: an instrument for social diagnosis in care homes for older persons

Silvia Patricia Cury, Andrés Arias Astray & José Luis Palacios Gómez

To cite this article: Silvia Patricia Cury, Andrés Arias Astray & José Luis Palacios Gómez (2017): Content validity analysis of ISD-1: an instrument for social diagnosis in care homes for older persons, European Journal of Social Work, DOI: 10.1080/13691457.2017.1364705

To link to this article: http://dx.doi.org/10.1080/13691457.2017.1364705

Published online: 17 Aug 2017.

EUROPEAN JOURNAL OF SOCIAL WORK, 2017
https://doi.org/10.1080/13691457.2017.1364705

Routledge
Taylor & Francis Group

Content validity analysis of ISD-1: an instrument for social diagnosis in care homes for older persons†

Análisis de la validez de contenido del IDIS.1: un instrumento para el diagnóstico social en residencias para personas mayores

Silvia Patricia Cury[a]*, Andrés Arias Astray[a] and José Luis Palacios Gómez[b]**

[a]Universidad Complutense de Madrid, Facultad de Trabajo Social, Madrid, Spain; [b]Rey Juan Carlos University, Public University in Móstoles, Madrid, Spain

ABSTRACT

This article presents the validity study for *ISD-1* (Instrument for Social Diagnosis), designed to be used in the specialist field of intervention in care homes for older persons. The study has focused on the evidence regarding the validity of its content. The definition of the operative area of *ISD-1* (social diagnosis in care homes), and its representativeness and relevance, are decisive aspects for its validity. Two validation procedures were used, with the participation of two independent groups of experts. Both procedures had the objective of obtaining a quantitative measure assessing the representation of the area and of the degree of association between the dimensions and the items of the instrument. We may conclude that there is a sufficient degree of evidence for the representativeness, relevance and usefulness of the content of *ISD-1*, meaning it may be considered a suitable instrument for the formulation of social diagnoses in care homes for older persons.

RESUMEN

Este artículo presenta el estudio de la validez del *IDIS.1* (Instrumento para el diagnóstico social), diseñado para para ser utilizado en el ámbito especializado de intervención de las residencias para personas mayores. El estudio se ha centrado en las evidencias de la validez de contenido. La definición operativa del dominio del *IDIS.1* (el diagnóstico social en residencias para personas mayores), y su representatividad y relevancia, resultan aspectos decisivos de su validez. Se emplearon dos procedimientos de validación, con la participación de dos grupos independientes de expertos. Ambos procedimientos tuvieron el objetivo de obtener una medida cuantitativa de la evaluación de la representación del dominio y del grado de vinculación entre las dimensiones y los ítems del instrumento. Se puede concluir que el *IDIS.1* cuenta con suficientes evidencias de representatividad, relevancia y utilidad de contenido, lo que permite considerarlo un instrumento válido para la formulación de los diagnósticos sociales en residencias para personas mayores.

KEYWORDS
Content validity; instrument; social diagnosis; care homes; older persons

PALABRAS CLAVE
instrumento; diagnóstico social; validez de contenido; residencias; personas mayores

CONTACT Silvia Patricia Cury silvia.cury@unir.net

†The first part of this work, which describes the ISD-1 design process and its results, has been published as an independent article in the European Journal of Social Work entitled 'Design of ISD-1: An instrument for social diagnosis in care homes for older persons'.

*Present address: Assistant Professor, Faculty of Legal and Social Sciences and Humanities, International University of La Rioja, UNIR, Calle Almansa 101, Edificio 1., 28040, Madrid, Spain.

**Present address: Ayuntamiento de Alcobendas, Patronato Sociocultural, Plaza Mayor 1, Madrid, 28100 Spain.

Introduction

This work presents the results of the validity study for *ISD-1* (Instrument for Social Diagnosis). *ISD-1* is aimed at enabling the correct formulation of such diagnosis in the specialist field of socio-healthcare intervention in care homes for older persons. Additionally, it seeks to unify and define the professional language used by social workers in this field of intervention.

The design of *ISD-1* was carried out in the 24 care homes in the Madrid Social Care Agency of the Community of Madrid (*Agencia Madrileña de Atención Social*; hereinafter *MSCA*), the institution that has sponsored this research. The instrument contains 627 items and is organised on the basis of a categorical system, including four fundamental dimensions of social diagnosis: the 'individual', 'social', 'family' and 'institutional' situations. These four dimensions are divided into 15 sub-dimensions, which contain the 83 categories for the formulation of the social diagnosis in care homes, according to the diagnostic group to which each category belongs. In order to correctly identify them, all the categories, sub-dimensions and dimensions of social diagnosis established in *ISD-1* are codified using an alphanumerical system. The 83 categories for the *ISD-1* social diagnosis are identified with a 'label' containing the name and definition of the category. A list of 544 diagnostic criteria has been designed to complete and specify this definition. In the cases where it has been deemed necessary, the minimum number of criteria required to formulate a particular diagnosis, the necessary type of information source and, in some cases, the period of time that must have passed since the appearance of specified signs or manifestations have also been established. *ISD-1* also has a 'Category Index' and a 'User Guide', which content describes the theoretical and technical basis for the classification system.

According to the nature and characteristics of *ISD-1* and its intended use in care homes, obtaining a sufficient degree of evidence of its validity as a social diagnosis instrument represents an indispensable methodological requirement, to be satisfied prior to its regular application in the environment of care homes for older persons.

Design of the *ISD-1* validity study

Validity is one of the most important psychometric concepts. Its definition has changed over time and continues to evolve, forming the object of intense ongoing debate among experts (Borsboom, 2009; Kane, 2013; Lissitz & Samuelsen, 2007; Markus & Borsboom, 2013). But its status as the most fundamental and important property in evaluating a test or measurement instrument has remained constant (Anastasi & Urbina, 1997; Cronbach, 1988; Kane, 2009). Therefore, the general aim of this work was to study the validity of *ISD-1* as an instrument for formulating social diagnoses in care homes for older persons.

The latest edition of the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) defines the concept of validity as the degree to which the evidence and theory support the interpretations of assessment instruments according to the use that will be made of them. In general, it should be noted that the current understanding of validity emphasises the interpretations that derive from the scores obtained. The impossibility of referring to the validity of an instrument in itself should also be emphasised, since this depends on the context of the evaluation and on the specific population under study. Moreover, the intended use of an instrument is taken into account when evaluating it, in such manner that the validity indices obtained for a specific application of an instrument cannot be generalised for other functions of the same instrument (Dings & Hershberger, 2002). Validity hence means the existence of proof and of evidence for each interpretation, and different degrees may be obtained relating to such evidence. As Messick (1980) would state, validity must be considered as a question of 'degree', rather than all or nothing.

Validity analysis is carried out by way of internal and external sources of evidence. The first group defines the validity of the content, the response process and the internal structure. The second group

encompasses evidence related with other variables and the social consequences of the use of a test. Determining the type of evidence that we must obtain for purposes of analysing the validity of *ISD-1* as an instrument for social diagnosis in care homes follows the proposal of Kane (2006). Kane maintains that validity involves two types of argument: the interpretative argument and validity *per se*. The interpretative argument consists of specifying the intepretations and use of the test (the assumptions that underpin the test, the components of the construct and the relationship of the construct with other variables). Meanwhile, the validity argument consists precisely of evaluating such interpretations and, on the basis thereof, determining the type of evidence that is necessary for their validation.

On the basis of all these recommendations, the validation process for *ISD-1* has been conducted with the objective of analysing its validity and usefulness in the institutional context of care homes for older persons, within the framework of a constantly evolving and continuing validation process, as stated by Cronbach (1971).

ISD-1 is not, strictly speaking, a test or a measurement instrument. Rather, it is an instrument for diagnosing the social situations of older persons in care homes. Its application does not, therefore, offer 'scores' that may be used in the statistical validation procedures necessary to analyse evidence based on its internal structure or its relation with other variables. In the case of evidence based on the relationship with other variables, the test scores are also used for purposes of comparison with the scores of certain external criteria. Though a study of this type could be undertaken to compare the diagnoses made using *ISD-1* and those formulated with another equivalent instrument (rather than comparing scores), at present there is no other specific instrument for social diagnosis in the context of interventions with older persons in care homes that offers comparable information. Nor do we have any other instrument that could be used to the same end without difficulties, meaning it is not feasible to perform an analysis of evidence based on the relationship with other variables.

Moreover, it was decided not to study evidence based on response processes or the social consequences of the use of *ISD-1*. In the former case, this kind of study would have meant conducting interviews with the *MSCA* social workers and gathering data during the course of the professional's intervention with the users. This circumstance would have interfered with the application of the instrument and with the spontaneous assessment to be undertaken by professionals with regard to *ISD-1*. As regards social consequences, while they are an important aspect of the use of an instrument, they are not in themselves true evidence of that instrument's validity (Borsboom, 2009; Borsboom, Mellenbergh, & van Heerden, 2004; Popham, 1997).

In light of all the foregoing, our study of the validity of *ISD-1* focused on the evidence for validity of its content. The need to analysis the content of a test as the first step in judging whether an instrument can be used for a specified aim had already appeared by 1954, as one of the APA recommendations for psychological tests and diagnostic techniques. The *Standards for educational and psychological testing* (AERA, APA, NCME, 2014, p. 14) highlight that 'important validity evidence can be obtained from an analysis of the relationship between the content of a test and the construct it is intended to measure'. Content does not merely encompass the group of themes, phrases and format of the items, but also includes the instructions for use and scoring of the instrument. Studying content-based evidence requires consideration of two fundamental factors (Sireci, 1998a, 2003): the definition and representation of the 'area'. The 'definition' of the area refers to the operating definition of its content, establishing the meaning and sense of the construct or area. The second aspect, the 'representation' of the area, incorporates the concepts of 'representativeness' and 'relevance'. Representativeness or coverage indicates whether the content of the test contemplates all the facets of the defined area and whether any aspects are under-represented. Relevance, meanwhile, refers to the degree to which the items composing the instrument measure the area and whether there is any content that is irrelevant in measuring it. In summary, the essential concept of content validity is that the items of an instrument must be relevant and representative of the construct for a particular evaluative purpose.

The two main procedures for evaluating content-based evidence belong to two groups: procedures based on expert opinions and statistical procedures. Assessment by judges or experts is the most commonly used procedure (Dings & Hershberger, 2002), but statistical procedures can also be used on the basis of subjects' responses to a particular test. Nonetheless, as Utkin (2006) maintains, expert opinions are an important part of the information available and have become the most widely used procedure to analyse the validity of content when experimental observations are limited. This procedure consists of certain judges evaluating the full content of the test or instrument and issuing their opinion as to the degree to which the content represents and defines the thematic area it purports to measure. In this case, a fairly low number of experts are used to issue a relatively high number of evaluations. It is essential to carefully select the expert group for this type of analysis, since the evidence with respect to the content is based on their opinions concerning the extent to which the area is well defined and the instrument adequately represents it (American Educational Research Association, & American Psychological Association and National Council on Measurement in Education, 1985). For this purpose, the procedure is based on the work of experts who issue opinions regarding the degree of matching between the established items and aims.

Once all the scores have been obtained from the different experts, the results are normally summarised using a statistical index. The congruence indices proposed by Rovinelli and Hambleton (1977) and Hambleton's relevance/representativeness index (1980, 1984) provide examples of the tools most frequently used to determine the evidence based on the test content. In the second case, that of statistical procedures (much less common in practice), some form of multivalent data reduction analysis technique is normally used: factor analysis (Dorans & Lawrence, 1987), multidimensional scaling and cluster analysis (Deville & Prometric, 1996; Sireci & Geisinger, 1992, 1995), or structural equation modelling (Dings & Hershberger, 2003) are examples of the first group of procedures, while studies of representativeness of area via the characteristic variance analyses (Green, 1983; Jarjoura & Brennan, 1982; Shavelson, Gao, & Baxter, 1995) offer examples of the generalizability theory approach.

Expert opinions were used to study the content-based evidence for *ISD-1*. Given the importance of assessing aspects relating to the potential under-representation of the area or construct or to the presence of factors that are irrelevant thereto in the design of *ISD-1*, two independent validation procedures were used, with the participation of two differentiated groups of experts.

Given the complexity of design of this study and the participation of different groups of experts, in Figure 1 we show a descriptive summary of the different phases and group processes used for the design and subsequent validation of *ISD-1*, as well as of the analysis units and of the composition of the three expert groups that participated in these processes. This description includes the procedures and participating groups for both the design of *ISD-1* and the subsequent validity study, since we consider it appropriate to offer the reader a full overview of our research and of the importance we have attributed to the different methodological procedures as regards the content of the instrument since the beginning of this study (Grant & Davis, 1997; Kane, 2013).

Research hypothesis

All the procedures described in this work were conducted for purposes of obtaining information that would permit an empirical analysis of whether the hypotheses as to the validity of *ISD-1* were satisfied. These hypotheses are as described below:

1. The results of the evaluation of the instrument, used systematically in all the MSCA care homes for a three-month period, show a high level of consistency among social workers with respect to its content and use for the formulation of social diagnoses, which is expressed through the following indicators:

1.1. At least 75% of social workers affirm that the four social diagnosis dimensions of *ISD-1* (individual, social, family and institutional situations) correspond to the main areas of social diagnosis in care homes.

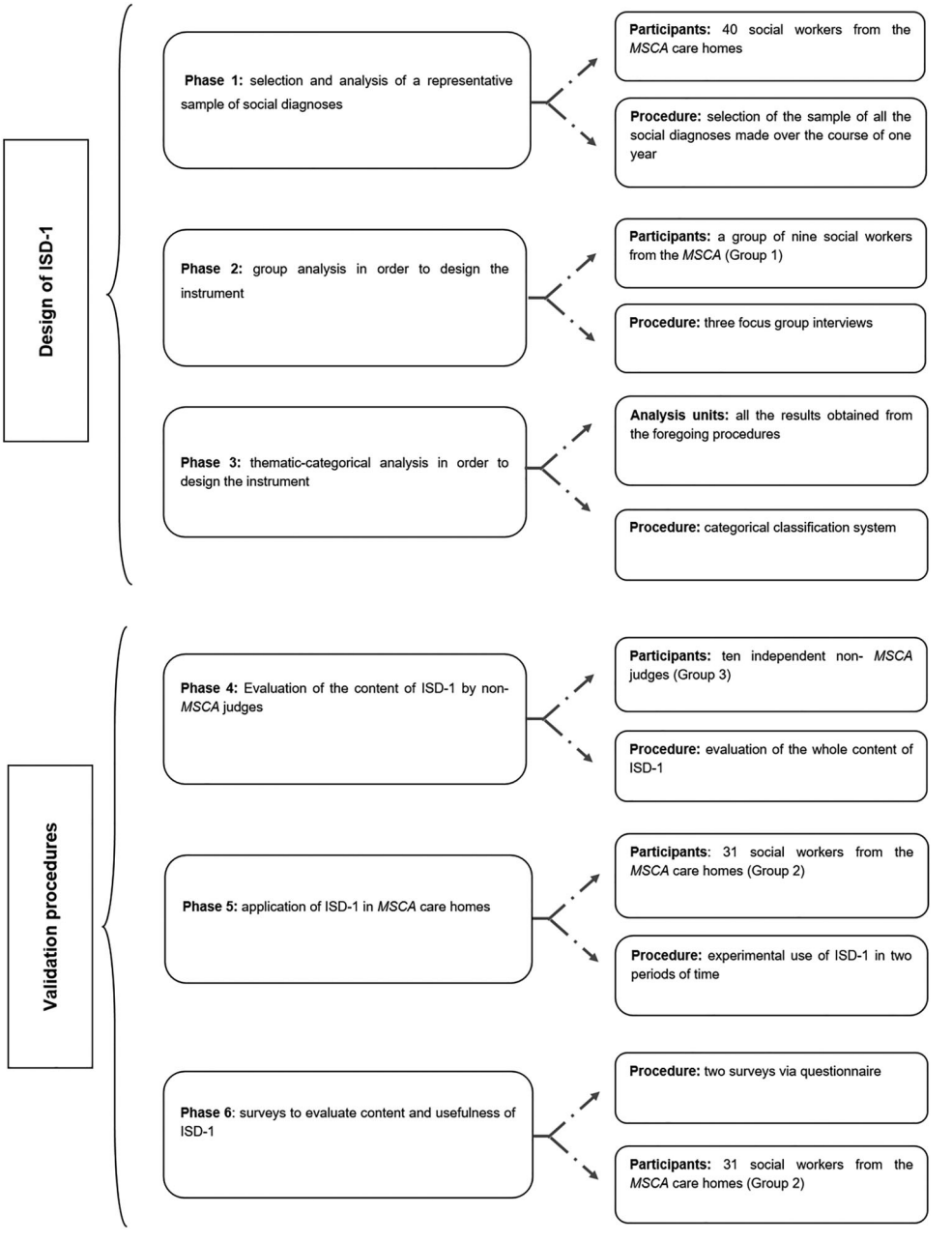

Figure 1. Methodological phases, procedures, analysis units and participants.

1.2. At least 75% of social workers affirm that *ISD-1* contemplates the main social situations diagnosed in care homes.

1.3. At least 75% of social workers affirm that the *ISD-1* user guide is useful when applying the instrument.

1.4. At least 75% of social workers affirm that the terminology used in *ISD-1* for the formulation of social diagnoses in care homes is appropriate.

1.5. At least 75% of social workers affirm that the definitions set out in *ISD-1* to describe each of the social diagnosis categories are correct.

1.6. At least 75% of social workers affirm that the diagnostic criteria set out in *ISD-1* are appropriate in more precisely defining the different social diagnosis categories.

1.7. At least 75% of social workers consider that a sufficient quantity of diagnostic criteria is set out in *ISD-1* to carry out social diagnosis.

1.8. At least 75% of social workers affirm that *ISD-1* makes it easier to carry out social diagnosis.

1.9. At least 75% of social workers affirm that the different categories in *ISD-1* are appropriate for carrying out a complete social diagnosis.

1.10. The mean and median that are calculated from the results of *MSCA* social workers' evaluations with respect to the content of *ISD-1* exceed 6 points, based on a scale ranging from 1 to 7 points where 7 is the value that expresses the highest degree of agreement.

2. The results of evaluation of the content of ISD-1 by a group of non-MSCA social workers show a high level of inter-judge agreement, which is expressed through the following indicators:

2.1. The calculated mean of the results of evaluation of the content of *ISD-1* by a group of non-*MSCA* social workers is at least 4 points, based on a scale ranging from 1 to 5 points where 5 is the value that expresses the highest possible degree of agreement.

2.2. The calculated median of the results of evaluation of the content of *ISD-1* by a group of non-*MSCA* social workers is at least 4 points, based on a scale ranging from 1 to 5 points where 5 is the value that expresses the highest possible degree of agreement.

Method

Evaluation of the content of ISD-1 *by non-*MSCA *judges*

Participants

The first evaluation of the content of *ISD-1* was carried out by 10 independent, non-*MSCA* judges. This group of judges comprised 70% women, with an average age of 41.2 years. The judges were social workers with an average of 18.2 years' professional practice experience, and an average of 11 years spent caring for older persons in residential environments.

Validation technique

The technique used was the evaluation of the whole content of ISD-1 by the participating judges, who stated their opinion with respect to the degree to which the ISD-1 items correspond to the different social diagnosis dimensions, sub-dimensions and categories and, therefore, to the conceptual area that the instrument is intended to measure. An item-objective congruence task was carried out for this purpose, and the 'relevance-representation indices' proposed by Hambleton (1980, 1984) were calculated. Each of the 10 judges expressed their opinion regarding each of the 627 *ISD-1* items, using the following 5-point Likert scale: 1: strongly disagree/ 2: disagree/ 3: neither agree nor disagree/ 4: agree/ 5: strongly agree.

On the basis of all the scores given by the 10 judges, the mean and the median were calculated and used as 'relevance indices' for the item. The average of the relevance indices for the 627 items was used as a 'representation index' for the entire content of *ISD-1*. The representation indices obtained as a result of this procedure reflected the level of inter-judge agreement with respect to the content of *ISD-1*. It was established that the mean and the median calculated for the result of all the assessed areas must exceed a score of 4 points in order for the level of agreement to be considered acceptable.

Instrument

A specific instrument was designed to evaluate the content of *ISD-1*, which included the complete content of the instrument and an introduction describing the structure of *ISD-1* and the theoretical and technical foundations for the instrument, together with the definition of the concept of 'social diagnosis'.[1]

Application of ISD-1 *in* MSCA *care homes*

ISD-1 was experimentally applied in 17 of the 24 *MSCA* care homes in order for the *MSCA* social workers to be able to assess the content and practical utility of *ISD-1* as an instrument for the formulation of social diagnosis in care homes for older persons.

Participants

Of the practising social workers in 17 of the 24 *MSCA* care homes, 31 out of 40 participated. The population of these social workers presented a profile with a clear majority of women, who represented 96.47%, with an average age of 47 years, an average of 21.65 years in professional social work practice, and an average of 16 years spent caring for older persons in residential environments. All of the participants held diplomas in social work, while six social workers also had other related qualifications.

Validation technique

The participating social workers applied *ISD-1* to formulate all the social diagnoses in their centres for two three-month periods, with an interim rest period of three months. The differentiation of the two application periods was due to the need to avoid the effects of the novelty of applying the instrument during the first period and the possible sources of bias arising out of the fatigue that may result from the task entrusted to them (Selltiz, Wrightsman, & Cook, 1980).

Instrument

The instrument used was the complete content of *ISD-1*.

Surveys to evaluate the content and usefulness of ISD-1

Participants

The participants were the 31 *MSCA* social workers who applied *ISD-1* when formulating social diagnoses during the two established three-month periods (Group 2).

Validation technique

At the end of each *ISD-1* application period, two self-administered surveys were conducted using an online questionnaire sent to participating social workers, in order to know their opinions regarding the content and usefulness of *ISD-1* for the formulation of social diagnoses in care homes. To safeguard the accessibility of information, it was located using a PHP/MySql hosting service, which is developed using responsive technology in order to secure its visualisation. The use of this online platform also permits anonymous responses, thus guaranteeing the protection of respondents' personal data.

Instrument

To carry out the two surveys, it was decided to design a questionnaire, the specific nature and brevity of which would make it easier for the professionals to complete. Nine closed questions were established with regard to each of the fundamental characteristics of *ISD-1* and concerning the suitability of the application for the formulation of social diagnoses in care homes. The content of the nine questions corresponded exactly to the formulation of the nine indicators relating to the first research hypothesis (see Figure 2).

The same questionnaire was used to conduct the surveys at the end of each of the two periods of application of *ISD-1*. As such, the results of the first and second surveys could be compared. We were hence able to evaluate whether the results of the first survey remained unaltered after the three-month rest period and the second application period lasting a further three months, as well as whether the results of the second survey showed significant differences in comparison with the results of the first survey.

Questionnaire
1. The 4 ISD-1 diagnostic dimensions correspond to the main areas of social diagnosis in care homes
2. ISD-1 covers the main situations subject to social diagnosis in care homes
3. The User Guide is useful as an introductory guide to ISD-1
4. The terminology used in ISD-1 is suitable for formulating social diagnoses in care homes
5. The definitions established in ISD-1 to describe the categories of social diagnosis in care homes are correct
6. The diagnostic criteria established in ISD-1 are suitable for defining the social diagnosis categories
7. ISD-1 contains a sufficient number of diagnostic criteria to perform the social diagnosis
8. ISD-1 makes it easier to draft social diagnoses
9. The categories established in ISD-1 are suitable for performing a complete social diagnosis

Figure 2. Survey questionnaire for social workers.

As the brevity of the questionnaire permitted it, and given the importance of obtaining data as to the different levels of agreement between the social workers' evaluations, it was decided to use the following Likert scale (from 1 to 7) to evaluate each of the nine items: 1. totally disagree/ 2. disagree to an extent/ 3. disagree more than agree/ 4. neither agree nor disagree/ 5. agree more than disagree/ 6. agree to an extent/ 7. totally agree. For the overall calculation of the percentage of social workers who expressed their agreement with the nine items, all the scores between values 5 and 7 on the scale (5: 'Agree more than disagree', 6: 'Agree to an extent' and 7: 'Totally agree') were incorporated.

Results

Results of evaluation of the content of ISD-1 *by non-*MSCA *judges*

Table 1 shows the relevance and representation indices obtained by calculating the scores given by the 10 judges for the group of 627 items making up *ISD-1*. The previously established value for the purpose of considering the content of *ISD-1* to be representative was 4 points for the mean and for the median. The final result of the evaluation, calculated from the scores of the 10 judges for the entire content of *ISD-1*, showed a mean of 4.46 points and a median of 5 points. These results indicate a high level of inter-judge agreement with respect to the content of *ISD-1*.

Results of the two ISD-1 *evaluation surveys*

The results of the two surveys conducted at the end of the first and second *ISD-1* application periods in the *MSCA* care homes are presented below.

Results of the first evaluation survey

In the first survey conducted, the percentage of the 31 participating social workers who expressed their agreement by giving scores of 5, 6 or 7 points (5: 'Agree more than disagree', 6: 'Agree to an extent' and 7: 'Totally agree') to the nine items on the questionnaire exceeded the established cut-

Table 1. Relevance and representation indices for the group of items.

Judges	Relevance items
1	4.71
2	4.74
3	4.03
4	4.71
5	4.28
6	3.46
7	4.30
8	4.96
9	4.85
10	4.81
Mean	4.46
Median	5.00

Table 2. Agreement percentage for social workers in the first survey.

Items	Scores			Total
	7	6	5	
1. The four ISD-1 diagnostic dimensions correspond to the main areas of social diagnosis in care homes	61.29	22.58	16.13	100.00
2. ISD-1 covers the main situations subject to social diagnosis in care homes	41.94	38.71	12.90	93.55
3. The user guide is useful as an introductory guide to ISD-1	61.29	22.58	16.13	100.00
4. The terminology used in ISD-1 is suitable for formulating social diagnoses in care homes	25.81	51.61	22.58	100.00
5. The definitions established in ISD-1 to describe the categories of social diagnosis in care homes are correct	38.71	22.58	29.03	90.32
6. The diagnostic criteria established in ISD-1 are suitable for defining the social diagnosis categories	35.48	38.71	22.58	96.77
7. ISD-1 contains a sufficient number of diagnostic criteria to perform the social diagnosis	22.58	51.61	12.90	87.10
8. ISD-1 makes it easier to draft social diagnoses	64.52	12.90	19.35	96.77
9. The categories established in ISD-1 are suitable for performing a complete social diagnosis	41.94	35.48	12.90	90.32

off point (75%) in all cases. An agreement percentage equal to or greater than 90% was obtained for all items, with the exception of item number 7, which obtained an 87.10% agreement percentage (see Table 2).

The first analysis of these results shows a very high percentage of social workers expressing their agreement with the content and usefulness of *ISD-1*. On the basis of these results, a more detailed analysis was carried out of the different levels of agreement expressed by the social workers in this first survey. In order to do so, the agreement percentage was also calculated excluding the responses that gave these items a score of 5 ('agree more than disagree'), that is, a somewhat lower level of agreement. The objective of this analysis was to observe whether, according to this sample of evaluators, these items reached the established cut-off point (75%). Otherwise, together with the results of the second survey, this analysis would permit the detection of elements that could be modified and perfected in future research and/or future versions of *ISD-1*. The results of this procedure showed that, even submitting the analysis to this stricter criterion, 6 of the 9 items exceeded the established cut-off point of 75% and only 3 items obtained a slightly lower percentage: 70.97% for item 5 and 74.19% for items 6 and 7 (see Table 3).

The mean and the median, as representation indices for the content of *ISD-1* and which were calculated by incorporating all the scores given by the 31 social workers for all the items, were equal to or greater than the established cut-off point of 6 points on the 1–7 scale used (see Table 4).

It may therefore be considered that, in this survey, the evaluators expressed a high level of agreement with the content of *ISD-1*, which was considered representative, relevant and useful as an instrument for the formulation of social diagnosis for older persons in care homes.

Table 3. Agreement percentage of social workers in the first survey, excluding scores with a value of 5 ('agree more than disagree').

Items	Agreement percentage
1. The four ISD-1 diagnostic dimensions correspond to the main areas of social diagnosis in care homes	83.87
2. ISD-1 covers the main situations subject to social diagnosis in care homes	80.65
3. The user guide is useful as an introductory guide to ISD-1	83.87
4. The terminology used in ISD-1 is suitable for formulating social diagnoses in care homes	77.42
5. The definitions established in ISD-1 to describe the categories of social diagnosis in care homes are correct	70.97
6. The diagnostic criteria established in ISD-1 are suitable for defining the social diagnosis categories	74.19
7. ISD-1 contains a sufficient number of diagnostic criteria to perform the social diagnosis	74.19
8. ISD-1 makes it easier to draft social diagnoses	77.42
9. The categories established in ISD-1 are suitable for performing a complete social diagnosis	77.42

Table 4. Scores given by social workers for the nine items in the first survey.

Items	Mean	Median
1. The four ISD-1 diagnostic dimensions correspond to the main areas of social diagnosis in care homes	6.45	7.00
2. ISD-1 covers the main situations subject to social diagnosis in care homes	6.10	6.00
3. The user guide is useful as an introductory guide to ISD-1	6.45	7.00
4. The terminology used in ISD-1 is suitable for formulating social diagnoses in care homes	6.00	6.00
5. The definitions established in ISD-1 to describe the categories of social diagnosis in care homes are correct	6.00	6.00
6. The diagnostic criteria established in ISD-1 are suitable for defining the social diagnosis categories	6.06	6.00
7. ISD-1 contains a sufficient number of diagnostic criteria to perform the social diagnosis	6.00	6.00
8. ISD-1 makes it easier to draft social diagnoses	6.39	7.00
9. The categories established in ISD-1 are suitable for performing a complete social diagnosis	6.03	6.00

Results of the second survey

The results of the second survey also showed a high level of homogeneity, with the percentage of social workers who expressed their agreement being equal to or greater than 80% for all the items of the questionnaire (see Table 5). The frequency distributions that we see in the table show that the percentage of social workers who gave scores of 5 was very low, meaning that even when these scores were excluded, eight of the nine items obtained a percentage exceeding 80% and only item 6 obtained a slightly lower percentage (74.19%).

The mean and median were also calculated for all the scores given by the social workers for the nine questionnaire items, for purposes of obtaining the two representation indices for the content of *ISD-1*. In both cases, the representation indices were equal to or greater than the cut-off point, which was established as 6 points (see Table 6).

The results obtained for the procedures carried out appear to corroborate all the hypotheses that guided this research. We may therefore conclude that there is a sufficient degree of evidence for the

Table 5. Agreement percentage of social workers in the second survey.

Items	Scores			Total
	7	6	5	
1. The 4 ISD-1 diagnostic dimensions correspond to the main areas of social diagnosis in care homes	48.39	32.26	3.23	83.88
2. ISD-1 covers the main situations subject to social diagnosis in care homes	31.03	51.61	3.23	83.87
3. The user guide is useful as an introductory guide to ISD-1	54.84	28.21	3.23	83.88
4. The terminology used in ISD-1 is suitable for formulating social diagnoses in care homes	38.71	41.94	3.23	83.88
5. The definitions established in ISD-1 to describe the categories of social diagnosis in care homes are correct	41.94	38.71	3.23	83.88
6. The diagnostic criteria established in ISD-1 are suitable for defining the social diagnosis categories	29.03	45.16	9.68	83.87
7. ISD-1 contains a sufficient number of diagnostic criteria to perform the social diagnosis	23.86	58.06	3.23	83.88
8. ISD-1 makes it easier to draft social diagnoses	51.61	31.13	3.23	83.87
9. The categories established in ISD-1 are suitable for performing a complete social diagnosis	45.16	32.26	6.45	83.87

Table 6. Scores given by social workers for the nine items in the second survey.

Items	Mean	Median
1. The four ISD-1 diagnostic dimensions correspond to the main areas of social diagnosis in care homes	6.00	7.00
2. ISD-1 covers the main situations subject to social diagnosis in care homes	6.00	6.00
3. The user guide is useful as an introductory guide to ISD-1	7.00	7.00
4. The terminology used in ISD-1 is suitable for formulating social diagnoses in care homes	6.00	6.00
5. The definitions established in ISD-1 to describe the categories of social diagnosis in care homes are correct	6.00	7.00
6. The diagnostic criteria established in ISD-1 are suitable for defining the social diagnosis categories	6.00	6.00
7. ISD-1 contains a sufficient number of diagnostic criteria to perform the social diagnosis	6.00	6.00
8. ISD-1 makes it easier to draft social diagnoses	7.00	7.00
9. The categories established in ISD-1 are suitable for performing a complete social diagnosis	6.00	7.00

representativeness, relevance and usefulness of the content of *ISD-1* for it to be considered a valid and potentially suitable instrument for the purpose for which it was designed, in the institutional context in which it is intended to be used.

Discussion and conclusion

ISD-1 is a specific instrument for the formulation of social diagnoses in care homes for older persons and the validation process has been conducted with the objective of analysing its validity and usefulness. On the basis of the intended purpose of *ISD-1*, to be used as a social diagnosis instrument, the operating area definition (social diagnosis in care homes for older persons) and its representativeness and relevance are decisive aspects for its validity (Kane, 2006). The definition of area means that *ISD-1* must have been identified and defined the significance and meaning of social diagnosis in care homes with the greatest possible precision. Its representativeness and relevance refer to the instrument needing to cover all the various facets of social diagnosis in this field of intervention. Its measurement ensures that none of its fundamental aspects are under-represented. Taking into account the complexity of the construct that *ISD-1* seeks to define, as well as the difficulty involved in defining a thematic area so broad as social diagnosis in care homes, a content-based study of evidence is fundamental in testing the validity of the instrument.

Among the possible limitations of content-based validity evidence, it should be recalled that authors such as Nunally (1978) argue that the validity of content as an appropriate sample of ideas and formulation of items depends inexorably on rational judgments. Certain authors also state that content validity, understood as the relevance and representativeness of items as an appropriate sample of a previously defined area, is a property of the instrument and not of the inferences that we make on the basis of its use (Guion, 1977; Messick, 1980, 1981). Moreover, Guion (1977) warns that, since content-based evidence is frequently analysed by the same researchers who have developed the test or instrument, these analyses present a significant risk of potential bias. Nonetheless, many authors consider that although content-based evidence may always be questioned, it is useful, particularly if the analysed area has been carefully specified and systematically sampled and the instrument or test has been properly evaluated (Cronbach, 1971; Cureton, 1951; Ebel, 1961; Flockton & Crooks, 2002; Kane, 1982; Kane, Crooks, & Cohen, 1999; Lissitz & Samuelsen, 2007; Sireci, 1998b).

In any case, the content-based validity study of *ISD-1* owes to the view expressed by the American Educational Research Association et al. (2014) with respect to the importance of the content and the need to assess the relationship between content and the construct or 'area' that one is seeking to measure. Following the views expressed by Yallow and Popham (1983), we consider that if the content is relevant and representative and these characteristics are well explained, this constitutes an implicit inference meaning that, though the content may be a property of the instrument, this property makes it possible to interpret the results of that instrument's use. At the same time, we have taken into account the contribution of Ebel (1983), who argues that an instrument is valid if its items properly reflect what it seeks to confirm and that validity is always ultimately established through rational judgments. Ebel also notes that there is an exaggerated emphasis on empirical

verification for validity studies, and that the fundamental issue is the intrinsic logical validity of an instrument. Though the debate continues with respect to content validity, all authors agree on its guiding function in instrument design as well as its importance in interpreting a construct and the inferences that may be drawn (Kane, 2009).

The specific objective of this *ISD-1* validation study consisted of verifying the validity of the instrument using content-based evidence, taking as a starting point the fundamental fact that, after its content was analysed by judges with expertise in the field, there was a sufficient degree for evidence for the validity of the instrument to consider it potentially suitable for use in care homes for older persons.

On the basis of the main limitations set forth in the literature with respect to the difficulty of defining universally accepted and shared concepts and of obtaining clear criteria for ascertaining the extent to which there is sufficient content validity evidence, this study recognised the need to engage two independent groups of judges who conducted two different procedures.

Moreover, the fact the correct and complete evaluation of *ISD-1* required its use by the social workers for the formulation of social diagnoses in care homes for older persons was taken into account, so that its validity had to be ensured in this specific context for the use for which it was designed (AERA, APA, NCME, 2014). As Kane (2009) affirms, the adoption of a single analysis procedure, limited to the representativeness of the instrument content, would have been insufficient as evidence of the validity of *ISD-1*. Hence, once assessments had been obtained from the 10 non-*MSCA* judges as regards the relevance and representativeness of each one of the 627 items that make up *ISD-1*, the instrument was experimentally applied for the formulation of all the social diagnoses in the 17 participating centres. In this manner, the *MSCA* social workers could evaluate the usefulness of *ISD-1* and the relevance and representativeness of its content for formulating social diagnoses on the basis of the practical experience accrued by applying it in the centres.

It is also noteworthy that the results of this procedure require a guarantee of adequate inter-judge reliability, in the sense that their assessments are consistent. The main limitation of these procedures is that by informing the experts about what the test is supposed to measure, we are restricting their evaluations to the proposed dimensions and aims and, as such, influencing their perceptions as to what the test measures. As the experts know the aims of the test, there may be a certain level of sensitivity toward the researcher's expectations, giving rise to a potential demand bias that may affect the experts' opinions. However, in this study the 10 non-*MSCA* judges did not know each other, and nor had they had any previous contact with the researchers or with the research project. Therefore, the potential bias that could have arisen from the judges having greater involvement in the research and/or in the institution that sponsored the study was limited.

For the second procedure conducted to analysis the evidence for the validity of *ISD-1*, it was applied by 31 *MSCA* social workers in the care homes. It should be recalled that, as shown in Figure 1, this group did not participate in the previous group procedures for the *ISD-1* design. The social workers who participated in the *ISD-1* validity study only had contact with the researcher during the personal interview carried out in each care home during the first phase of the study, and had contributed the requested sample of social diagnoses for analysis. In this manner, we limited the possible sources of bias deriving from professionals' expectations with regard to the results of this study and from their relationship with the researcher (Landsberger, 1958; Rosenthal, 1966). It is also worth highlighting that the potential bias emerging from the social workers' prior knowledge of the aim of this study simultaneously represents a guarantee of the reliability of their evaluations. The social workers had been informed of the possible interest of *MSCA* in the future use of *ISD-1* as an instrument for regular use in care homes to formulate all social diagnoses. If we take into account that the formulation of a social diagnosis constitutes the fundamental basis from which social workers plan the design of interventions in care homes, and is therefore one of the most fundamental tasks that they carry out as professionals, we may consider that there are sufficient guarantees of reliability with respect to their evaluation of the content and usefulness of *ISD-1*. The social workers knew the decisive importance of their critical opinions in this respect, and were

aware that the regular implementation of the instrument in their centres could depend on their views.

The design of this procedure allowed social workers to apply *ISD-1* for a sufficient period of time to be able to subsequently evaluate its content and usefulness. The differentiation of the two application periods owed first to the need to avoid the novel effect of application of the instrument during the first period and, second, to the wish to alleviate social workers' fatigue due to having to familiarise themselves with a new instrument and its application to the formulation of social diagnoses. In this manner, the main sources of bias deriving from the reactive effects of the techniques used in the study were limited (Selltiz et al., 1980).

In conclusion, it may be noted that the results obtained have confirmed all the hypotheses that guided this research. These results have shown that a broad sample of qualified judges who participated in the *ISD-1* validity analysis procedures express a high degree of agreement with respect to the representation, relevance and usefulness of this instrument for the formulation of social diagnoses in care homes.

The *ISD-1* validation study sought to verify the validity of the instrument for the proposed purposes. Based on the procedures used for its analysis, on the qualification of the participating evaluators and on the results obtained, we may conclude that there is a sufficient degree of evidence for the representativeness, relevance and usefulness of the content of *ISD-1* to enable it to be considered a useful instrument for application in care homes with older persons. It is therefore to be hoped that this line of research will be continued, with a more prolonged and regular application in care homes and the completion of new *ISD-1* validity studies, broadening the approach used in this work and obtaining more evidence as to its validity and improvements, extensions and/or adaptation of the instrument's content.[2]

Notes

1. Definition of the concept of 'social diagnosis' (Blinded for review, 2016):

   > Social diagnosis is the professional judgment made by a social worker as a result of the study and interpretation of a given social situation, and which constitutes the basis for social intervention in that situation. The correct formulation of the social diagnosis is the responsibility and competence of the social worker, and must take into account the difficulties and strengths of the person and of their individual, family, social and institutional situation.

2. Note: Micro-data corresponding to this article are available in:
   https://goo.gl/DBeobq (Results of evaluation of the content of ISD-1 by non-*MSCA* judges)
   https://goo.gl/ebjCSp (Results of the first IDS-1 evaluation survey)
   https://goo.gl/jtKq6n (Results of the second IDS-1 evaluation survey).

Acknowledgements

The authors particularly wish to thank all the social workers who have participated in this study, since their dedication and involvement have made its completion possible.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors acknowledge the Autonomous Agency Madrid Social Care, which has sponsored and at all times supported the completion of this study, and Universidad Complutense, Madrid, which has financed this research by providing Silvia Patricia Cury with a four-year pre-PhD grant.

Notes on contributors

Silvia Patrica Cury has a PhD in Social Work, a University Master's in Community Social Work and Management and Evaluation of Social Services and a Diploma in Social Work from Universidad Complutense, Madrid (UCM), with an Outstanding PhD Award (*Premio Extraordinario de Doctorado*) in Social Work. She is currently an associate professor at the International University of La Rioja for the degree in Social Work and the master's in Social Intervention in Knowledge Societies, coordinator for the master's in Socio-Healthcare Management and Intervention, and associate on the UCM PhD programme in Social Work. She has published various monograph chapters and articles in scientific social work journals.

Andrés Arias Astray is a Professor of Social Work at Complutense University (Madrid). His research interest revolves around the Improvement of social and personal well-being through the critical analysis of social work practices and theories and the development of new professional tools and instruments. He is also very interested in the research and professional practice of social work with groups. He is the Academic Coordinator of the Social Work PhD programme at Complutense and a member of the Executive Committee of the European Association of Schools of Social Work.

José Luis Palacios Gómez has a Degree and PhD in Political Sciences and Sociology from UCM, Madrid, as well as a PhD from Universidad Autónoma, Madrid on its Methodology of Behavioural and Health Sciences programme. He carries on his professional activities as a sociologist on the special management scale, higher technical class, of local government administration. He has worked as an associate professor at various universities, teaching the subjects of Social Research Methods and Techniques and Statistics, among others. He has published seven books and numerous articles in scientific journals of sociology, economics and business administration.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anastasi, A., & Urbina, S. (1997). *Psychological testing*. New York, NY: Macmillan.

Borsboom, D. (2009). *Measuring the mind. Conceptual issues in contemporary psychometrics*. New York, NY: Cambridge University Press.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*, 1061–1071. doi:10.1037/0033-295X.111.4.1061

Cronbach, L. (1971). Test validation. In R. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 335–355). Washington, DC: American Council of Education.

Cronbach, L. (1988). Five perspectives on the validity argument. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 3–17). Hillsdale, NJ: Erlbaum.

Cureton, E. (1951). Validity. In E. Lindquist (Ed.), *Educational measurement* (pp. 621–694). Washington, DC: American Council on Education.

Deville, C. W., & Prometric, S. (1996). An empirical link of content and construct validity evidence. *Applied Psychological Measurement*, *20*, 127–139. doi:10.1177/014662169602000202

Dings, C., & Hershberger, S. C. (2002). Assessing content validity and content equivalence using structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 283–297. doi:10.1207/S15328007SEM0902_7

Dorans, N. J., & Lawrence, I. M. (1987). *The internal construct validity of the Scholastic Aptitude Test*. Research Report. Princeton, NJ: Educational Testing Service.

Ebel, R. (1961). Must all tests be valid? *American Psychologist*, *16*, 640–647.

Ebel, R. L. (1983). The practical validation of tests of ability. *Educational Measurement: Issues and Practice*, *2*(2), 7–10.

Flockton, L., & Crooks, T. (2002). *Social studies assessment results 2001*. Dunedin: Educational Assessment Research Unit, University of Otago.

Grant, J., & Davis, L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health*, *20*, 269–274. doi:10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.3.CO;2-3

Green, B. F. (1983). Identifiability of spurious factors using linear factor analysis with binary items. *Applied Psychological Measurement*, *7*, 139–147. doi:10.1177/014662168300700202

Guion, R. M. (1977). Content validity: Three years of talk-what's the action. *Public Personnel Management*, *6*, 407–414.

Hambleton, R. (1980). Test score validity and standard-setting methods. In R. A. Berk (Ed.), *Criterion-referenced measurement: The state of art* (pp. 80–123). Baltimore, MD: John Hopkins University Press.

Hambleton, R. (1984). Validating the test score. In R. A. Berk (Ed.), *A guide to criterion referenced tests construction* (pp. 199–230). Baltimore, MD: John Hopkins University Press.
Jarjoura, D., & Brennan, R. (1982). A variance components model for measurement procedures associated with a table of specifications. *Applied Psychological Measurement*, *6*, 161–171. doi:10.1177/014662168200600202
Kane, M. (1982). A sampling model for validity. *Applied Psychological Measurement*, *6*, 125–160. doi:10.1177/014662168200600201
Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4nd ed., pp. 17–64). Westport, CT: National Council on Measurement in Education and American Council on Education-Praeger Series on Higher Education.
Kane, M. (2009). Validating the interpretations and uses of test scores. In R. Lissitz (Ed.), *The concept of validity* (pp. 39–64). Charlotte, NC: Information Age.
Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1–73. doi:10.2307/23353796
Kane, M., Crooks, T., & Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, *18*, 5–17. doi:10.1111/j.1745-3992.1999.tb00010.x
Landsberger, H. (1958). *Hawthorne revisited: Management and the worker, its critics, and developments in human relations in industry*. New York, NY: Ithaca.
Lissitz, R., & Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding. *Educational Researcher*, *36*, 635–694. doi:10.3102/0013189X07311286
Markus, K., & Borsboom, D. (2013). *Frontiers of test validity theory*. New York, NY: Routledge.
Messick, S. (1980). Test validity and ethics of assessment. *American Psychologist*, *35*, 1012–1027. doi:10.1037//0003-066X.35.11.1012
Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Researcher*, *10*, 9–20. doi:10.3102/0013189X010009009
Nunally, J. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.
Popham, W. (1997). Consequential validity: Right concern-wrong concept. *Educational Measurement: Issues and Practice*, *16*, 9–13. doi:10.1111/j.1745-3992.1997.tb00586.x
Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York, NY: Appleton-Century-Crofts.
Rovinelli, R. J., & Hambleton, R. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, *2*, 49–60.
Selltiz, C., Wrightsman, S., & Cook, S. (1980). *Métodos de investigación en las relaciones sociales*. Madrid: Rialp.
Shavelson, R., Gao, X., & Baxter, G. (1995). On the content validity of performance. In M. Bierembaum & F. Douchy (Eds.), *Alternatives in assessment of achievements, learning process, and prior knowledge* (pp. 131–141). Boston, MA: Kluwer Academic.
Sireci, S. (1998a). Gathering and analyzing content validity data. *Educational Assessment*, *5*, 299–321. doi:10.1207/s15326977ea0504_2
Sireci, S. (1998b). The construct of content validity. *Social Indicators Research*, *45*, 83–117. doi:10.1023/A:1006985528729
Sireci, S. (2003). Content validity. In R. Fernández (Ed.), *Encyclopedia on psychological assessment* (pp. 1075–1077). London: Sage.
Sireci, S., & Geisinger, K. (1995). Using subject-matter experts to assess content representation: An MDS analysis. *Applied Psychological Measurement*, *19*, 241–255. doi:10.1177/014662169501900303
Sireci, S., & Geisinger, K. F. (1992). Analyzing test content using cluster analysis and multidimensional scaling. *Applied Psychological Measurement*, *16*, 17–31. doi:10.1177/014662169201600102
Utkin, L. V. (2006). A method for processing the unreliable expert judgments about parameters of probability distributions. *European Journal of Operational Research*, *175*, 385–398. doi:10.1016/j.ejor.2005.04.041
Yallow, E. S., & Popham, W. J. (1983). Content validity at the crossroads. *Educational Researcher*, *12*(8), 10–21. doi:10.3102/0013189X012008010