Psychometric Evaluation of the Swedish Adaptation of the Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals Revised (IAPCC-R)

Helen Olt, Maria Jirwe, Petter Gustavsson and Azita Emami

J Transcult Nurs 2010; 21; 55 originally published online Oct 12, 2009;
DOI: 10.1177/1043659609349064

Journal of Transcultural Nursing
21(1) 55–64

DOI: 10.1177/1043659609349064
http://tcn.sagepub.com
SAGE

Psychometric Evaluation of the Swedish Adaptation of the Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals–Revised (IAPCC-R)

Helen Olt, MSc, RN, RNT,[1] Maria Jirwe, PhD, RN,[1,2] Petter Gustavsson, PhD,[1] and Azita Emami, PhD, RN, RNT[1,2]

Abstract

The purpose of this study was to describe the translation, adaption, and psychometric evaluation process in relation to validity and reliability of the Swedish version of the instrument, Inventory for Assessing The Process of Cultural Competence Among Healthcare Professionals–Revised (IAPCC-R) following the translation, adaptation, and psychometric evaluation process. Validity tests were conducted on the response processes ($N = 15$), the content ($N = 7$), and the internal structure of the instrument ($N = 334$). Reliability ($\alpha = .65$ for the total scale varying between −.01 and .65 for the different subscales) was evaluated in terms of internal consistency. Results indicated weak validity and reliability though it is difficult to conclude whether this is related to adaptation issues or the original construction. The testing of the response process identified problems in relation to respondents' conceptualization of cultural competence. The test of the content identified a weak correspondence between the items and the underlying model. In addition, a confirmatory factor analysis did not confirm the proposed structure of the instrument. This study concludes that this instrument is not valid and reliable for use with a Swedish population of practicing nurses or nursing students.

Keywords

instrument evaluation, cultural competence, psychometrics, reliability, validity, Swedish nurses, IAPCC-R

Numerous theories, models, and instruments have been developed to describe, evaluate, and measure cultural competence (Campinha-Bacote, 2003a; Kim-Godwin, Clarke, & Barton, 2001; Leininger & McFarland, 2002; Purnell & Paulanka, 2003). Often, these tools are translated and transferred to new cultural contexts uncritically, as is the case with current Swedish studies on cultural care (Gebru & Willman, 2003; Lundberg, 2000). However, the reliability and validity of an instrument, and its conceptual equivalence, must be scrutinized after translation. Furthermore, as validity is context bound, it is impossible to guarantee that use of the instrument in a new context will measure that which it was originally intended to (Beaton, Bombardier, Guillemin, & Ferraz, 2000; Geisinger, 1994; Maneesriwongul & Dixon, 2004; Spector, 1992).

In this article, the process of translating and adapting a tool for measuring cultural competence into the Swedish language and culture is presented, as well as the subsequent testing of the validity and reliability of the Swedish version. The instrument studied was the Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals–Revised (IAPCC-R), which was originally designed to measure cultural competence in transcultural nursing settings, though it can be applied in a range of arenas (Campinha-Bacote, 2003a).

The instrument, the IAPCC-R, originates from the model, The Process of Cultural Competence in the Delivery of Healthcare Service (Campinha-Bacote, 2003a), which was designed for use within the nursing community, for example, registered nurses, advanced practice nurses, nursing students, and nursing faculties. The instrument contains 25 items evenly distributed into five subscales. The five subscales reflect each area of the model and are described as follows: Cultural Desire (CD) refers to a genuine interest in, and a motivation to become involved in the process of developing cultural competence, that is "wanting to" rather than "having to" develop cultural competence. Cultural Awareness (CA) is the

[1]Karolinska Institutet, Stockholm, Sweden
[2]College of Nursing, Seattle University, Seattle, WA USA

Corresponding Author:
Helen Olt, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, SE 141 83, Sweden
E-mail: helen.olt@ki.se

understanding of one's prejudices toward others and the awareness of one's own culture. Cultural Knowledge (CK) refers to seeking knowledge about other cultural groups and an understanding that culture affects perceptions, values, beliefs, and views on health, illness, and treatment. Cultural Skill (CS) refers to the actual ability to collect and assess culturally based information when conducting physical assessments of people from different cultures. Finally, Cultural Encounter (CE) refers to an opportunity to gain experience from face-to-face meetings with patients from other cultures.

Background

Contemporary health care must be continually improved on to create a system that is as user-friendly as possible. Nurses can play an important role in this endeavor by creating conditions that promote the well-being of patients and their significant others. Specifically, nurses must maintain a holistic perspective of their patients and consciously recognize that experiences and understandings are shaped by a person's social and cultural background. Misunderstandings frequently occur when health personnel encounter patients from other cultures (Bäärnhielm, 2003). Emami (2000) suggests that such misunderstandings often arise because health personnel approach situations from a biomedical-oriented perspective, whereas the patient's understanding of the same situation is based on their life world, which is in turn shaped by life experiences that take place in a sociocultural context. Other literature suggests (Campinha-Bacote, 2003a; Emami, Benner, & Ekman, 2001; Giger & Davidhizar, 2004; Leininger & McFarland, 2002; Purnell & Paulanka, 2003) that to provide good-quality health care, personnel must combine biomedical knowledge with involvement, empathy, flexibility, and an understanding of the patient's social and cultural background.

A transcultural approach to nursing care can be traced to Leininger's research during the 1950s, which aimed at developing humanistic and scientific knowledge to provide culture-specific care (Leininger & McFarland, 2002). Building on her work, several instruments were developed and are now used in transcultural nursing to measure cultural competence. These tools include the Cultural Competence Scale (CCS; Kim-Godwin et al., 2001), the Cultural Self-efficacy Scale (CSES; Bernal & Froman, 1987), the Transcultural Self-Efficacy Tool (TSET; Jeffreys, 2000), IAPCC-R (Campinha-Bacote, 2003a), and the Cultural Competence Assessment instrument (CCA; Schim, Doorenbos, Miller, & Benkert, 2003).

Among the instruments listed above, the IAPCC-R is commonly used in nursing education. Given that our intention was to use an instrument pre- and postteaching interventions in cultural competence to measure the effects of training and education, and because the IAPCC-R was initially intended for use with those target groups that our study aimed to approach, this instrument was an appropriate choice. However, this choice also had to be weighed against the possible problems that could be encountered with the instrument, particularly in relation to the lack of directions for respondents and the combination of positive- and negative-phrased items, as well as the interchangeable use of the terms *cultural* and *ethnic*.

Very little scientific results were available regarding the validity and reliability of the IAPCC-R generally or following adaptation to a new context. Koemple conducted a study in 2003 with 275 certified nurse practitioners and established a Cronbach's alpha of .85 and Guttmann's split half of .83. Spencer and Cooper-Brathwaite conducted a study with 50 public health–registered nurses establishing a Cronbach's alpha of .90 (Campinha-Bacote, 2003a). Given the limited number of studies and information about the validity and reliability of the instrument, there was a need to investigate these matters in relation to the Swedish context before conducting the intended research with it. The aim of this study is to describe the translation, adaption, and psychometric evaluation process in relation to validity and reliability of the Swedish version of the instrument, IAPCC-R.

Method

Instrument

The 25 items included in the IAPCC-R are based on a 4-point Likert-type response scale with five different response categories, that is, measurement scale anchors. One measurement anchor is *strongly agree*, *agree*, *disagree*, and *strongly disagree*. The scores given for the items range from 4 to 1 and five of the items (#1, #3, #11, #17, and #21) are inverted and given scores of 1 to 4. The total scores indicate whether a health care worker is culturally incompetent (25-50 points), culturally aware (51-74 points), culturally competent (75-90 points), or culturally proficient (91-100 points; Campinha-Bacote, 2003a).

Translation and Adaptation

The IAPCC-R was translated and adapted to a Swedish context in autumn 2003 with permission from the developer of the original version and in accordance with Geisinger's guidelines (1994). Three authors (HO, MJ, and AE) participated in this process. After the initial individual translation and adaptation, the authors compared documents and agreed on a final joint document. The translated instrument was sent to three other persons, together with the original instrument. These three individuals were selected based on their professional and academic backgrounds (doctoral students and researchers knowledgeable in instrument design and cultural issues originating from Sweden, United Kingdom, and North America), their fluency in English and Swedish, and their

earlier experiences of working with assessment instruments. Written instructions were enclosed, asking informants to record their thoughts and evaluation of the translated instrument's items as well as their reactions to cultural and linguistic issues.

A meeting was arranged for all 6 persons and each item was discussed word by word, on the basis of each person's notes. Each item's content and construct was discussed. E-mails were sent to Campinha-Bacote for clarification when difficult questions arose. The next step, in accordance with Geisinger (1994), was to adjust the first draft of the instrument. This was based on the group members' comments after clarification from the developer. Some words were adapted, for example, the word "client." In Sweden, the word "patient" is used instead of "client." Questions also emerged regarding the number of response categories. There were five different measurement scale anchors in the original instrument. After a discussion with Campinha-Bacote, only two measurement scale anchors were chosen for the Swedish instrument. The revised instrument was then sent back to the group for comments before the translation and adaptation process was completed. The translated instrument was then back translated in two steps, first by a person fluent in both American English and Swedish, and then by a professional translator. The final back translation and the Swedish instrument were sent to the developer of the original instrument (Campinha-Bacote, 2003a).

Ethical Considerations

All respondents were given written and verbal information about the study, and informed consent was obtained from the respondents. They were told that they could withdraw from the study whenever they wanted. We communicated with the developer of the instrument throughout the whole process to keep her informed during the different stages. Ethical approval was obtained from the ethical committee, and the study was also approved by the local institutional review board.

Testing of the Response Processes

Fifteen respondents were convenience sampled (Polit & Beck, 2004). These included 5 lecturers, 5 nursing students, and 5 registered nurses, a total of 2 men and 13 women. One way to collect evidence on the validity of an instrument is to explore the response process, that is, by asking the respondents to read the questions out loud. Participants were then asked to discuss the questions openly so that their responses could be tape-recorded (Goodwin, 2002). The respondents were interviewed in spring 2004. The interviews were inspired by the cognitive method "Think aloud," whereby respondents are asked to verbalize what they are thinking as they read and answer the items (Sudman, Bradburn, & Schwarz, 1996). Data were analyzed with Conrad and Blair's (1996) taxonomy for analyzing "think aloud data," which gives the researcher an idea of the possible problems that might emerge and, if any, where these lie. Conrad and Blair include five classes of problems in their taxonomy: lexical problems, inclusion/exclusion problems, temporal problems, logical problems, and computational problems. Lexical problems are identified when the respondent does not know the meaning of a word or a phrase, and this leads to problems when he or she is asked to give an answer. Inclusion/exclusion problems are identified when respondents have difficulties in determining what to include or exclude in a word used in an item. Temporal problems mean incompatibilities with the item and the response option. Logical problems occur when the item has more than one focus, or includes negations, contradictions, or tautologies. Finally, computational problems are a residual class where problems that do not match the other four classes are placed.

Content Validity Testing of the Instrument

A field test was conducted in spring 2004 to validate the content of the instrument. One way of measuring validity of the content is to use an expert panel to evaluate whether the items measure what they were intended to measure (Goodwin, 2002). Seven experts on cultural issues were asked to participate. An article written by the developer (Campinha-Bacote, 2003b) describing the model was given to the experts together with the instrument. They were requested to relate each item to the subscale they thought the item belonged to, that is, which of the five subscales in the model the item measured. After data collection, each expert's responses were analyzed in accordance with data from Campinha-Bacote, who had provided us with the key (answers) as to which subscale each of the 25 questions measured.

Internal Validity Testing of the Instrument

The field test on the internal structure of the instrument was conducted throughout 2005 until spring 2006. The instrument was convenience distributed to three groups of respondents: nursing students (n = 138) in their third semester, nursing students in their fourth semester (n = 102), and registered nurses participating in a specialist nursing program (n = 94), in total 334 persons. Item analyses were conducted (Clark & Watson, 1995; Spector, 1992), testing each item's association with (a) the total summary measure, (b) the subscale it belonged to, and (c) the other subscales. Items that correlated above .40 with the total scale and the subscale they belonged to, while having lower correlations with the other subscales, were regarded as suitable items. If all items in a subscale possessed this quality, this was interpreted as evidence of a good internal structure. Data were analyzed with SPSS, Version 14.0.

Item analysis is a good starting point when scrutinizing an instrument (Spector, 1992). However, it does not test one of the most fundamental assumptions in classical test theory, that is, the assumption of unidimensionality. This assumption states that items in a scale should only measure one common phenomenon. In addition, specific phenomena assessed by items and not attributable to a common factor should not be correlated. The best test of this is implemented in the confirmatory factor analytic (CFA) framework (Brown, 2006). Thus, to test the proposed structure of the instrument, a confirmatory factor analysis (Brown, 2006) was first conducted, following the procedure for Likert-type items outlined by Jöreskog (2005). Two alternative models were compared, the one-dimensional model including one latent factor explaining all common variance in all 25 items, and a five-dimensional structure corresponding to the five subscales.

To assess model fit, different indices were used to capture various aspects (Brown, 2006). Absolute fit was assessed by the standardized root mean square residual (SRMR). Parsimonious fit was assessed by the root mean square error of approximation (RMSEA). Finally, incremental fit was assessed by the comparative fit index (CFI). These indices and proposed cutoff points were chosen on the basis of their performance in previous Monte Carlo simulations and recommendations based on these simulations (Brown, 2006). Good model fit is indicated specifically with an SRMR <0.08, an RMSEA around 0.05, and values >0.95 for CFI.

As a second procedure, an explorative factor analysis was conducted to highlight possible sources of model misfit and to find alternative common factors. A principal axis factor analysis with oblique rotation (Gorsuch, 1983) was also performed. Oblique rotation was used as the factors were expected to be correlated. Data were analyzed with LISREL 8.7 (Jöreskog, 2005) and SPSS, Version 14.0.

Reliability Testing (Cronbach's Alpha)

Geisinger (1994) recommends reliability testing on a larger population. Therefore, Cronbach´s alpha was calculated with the data from the 334 instruments completed during the field test of the instrument's internal structure. Data were analyzed with the SPSS, Version 14.0.

Results

The results are presented according to the problem classes in Conrad and Blair's (1996) taxonomy.

Lexical Problems

A lexical problem is identified when the respondent does not know the meaning of a word, a combination of words, or a phrase in the item. This problem was found in items belonging to subscales CD, CA, CK, and CS. The word "culture" was used in some items, and this was discussed with regard to its content, as in Item 5: "I feel that there is a relationship between culture and health" (Campinha-Bacote, 2003a, p. 109).

> Do they mean cultural or peoples' ethnic background or music or culture . . . what do they mean? (Lecturer 2, CS)

The combination of culture and competence in *cultural competence* was problematic for several respondents, as in Item 1: "Cultural competence mainly refers to one's competency concerning different ethnic groups" (Campinha-Bacote, 2003a, p. 109).

> This one was hard . . . the word cultural competence, what do they actually mean? (Nurse 4, CA)

Inclusion/Exclusion Problems

Inclusion/exclusion problems refer to when respondents encounter difficulties in determining what to include/exclude in a word used in an item. This was found in all five subscales, as with Item 19: "I have a passion for caring for clients from culturally/ethnically diverse groups" (Campinha-Bacote, 2003a, p. 110). The respondents found it difficult to understand whether culturally/ethnically diverse groups referred to all members of society or only members with a foreign background. One student said,

> I care for them as I do for others, it's not that I have a passion for them more than others, you can't think like that when you are working. (Student 4, CD)

When items referred to "the others," the respondents were confused as to whom this referred to and, therefore, it was difficult to know who to include/exclude, as in Item 21: "It is more important to conduct a cultural assessment on ethnically diverse clients than with other clients" (Campinha-Bacote, 2003a, p. 110).

> Other than my own, but what is my own group, a European or what? (Lecturer 1, CS)

Temporal Problems

Temporal problems arise when the item and the response choices are incompatible. This occurred in items belonging to subscales CA and CK. It was difficult for the respondents to answer some items with the proposed response category as the items asked for a specific quantity. For example, Item 15: "I am aware of at least two institutional barriers that prevent cultural/ethnic groups from seeking health care services" (Campinha-Bacote, 2003a, p. 110).

> Either I know of two or not, the answer can only be yes or no. (Lecturer 2, CA)

Respondents said that some response categories were unsuitable, as with Item 10: "I am knowledgeable about biological variations among different ethnic groups" (Campinha-Bacote, 2003a, p. 109). One nurse said,

> The response options are difficult to grasp. If I agree/disagree with this or not, it's difficult to answer, . . . it feels wrong. Isn't there just a yes or no? (Nurse 3, CK)

Logical Problems

Logical problems refer to items with more than one focus, or items containing negations, contradictions, or tautologies as in subscales CD, CA, and CK. Logical problems were found in the instrument's construction. Items 4, 7, and 19 were considered to be identical. Item 4 asked for the respondent's commitment, Item 7 for motivation, and Item 19 about passion when caring for clients from culturally/ethnically diverse groups, as in Item 19: "I have a passion for caring for clients from culturally/ethnically diverse groups" (Campinha-Bacote, 2003a, p. 109).

> It feels like you are asking the same questions over and over again but in a different way, committed, motivated and passionate, they are all the same. (Nurse 5, CD)

Some items contained negations that were confusing, as in Item 3: "Factors such as geographical location, gender, religious affiliation, sexual orientation, and occupation are not considered areas of concern when seeking cultural competence" (Campinha-Bacote, 2003a, p. 109).

> I think this is difficult to answer, I do not really know how I am supposed to answer. I should "not" think of these factors to develop cultural competence, is that what it means? (Student 5, CA)

Computational Problems

Computational problems are a residual class—for problems that do not match any of the other four classes. These arose in subscales CD, CA, CS, and CE. Some of the respondents had trouble answering some items, as in Item 22: "I feel comfortable in asking questions that relate to a client's ethnic/cultural background" (Campinha-Bacote, 2003a, p. 110).

> I haven't asked a lot of questions yet as I haven't begun my work as a nurse yet. (Student 5, CS)

Another design problem identified in one item was the fact that it guided respondents on how to interpret other items, as in Item 3: "Factors such as geographical location, gender, religious affiliation, sexual orientation, and occupation are not considered areas of concern when seeking cultural competence" (Campinha-Bacote, 2003a, p. 109).

> I was caught red-handed here because everything refers to cultural competence: gender, religious affiliation, and sexual orientation. But at first, I was just generalizing all ethnic groups to immigrants. (Nurse 4, CA)

Some respondents found it difficult to answer certain items. Six of the 15 respondents refused to answer one item each (#3, #11, #15, #17, #21, and #25). Twelve respondents skipped items and answered them when they had completed the other items in the instrument (#1, #3, #10, #11, #21, and #25) after they had time to reconsider.

Content Validity Testing of the Instruments

Data collected from seven experts were analyzed, to evaluate to what extent each item matched the definition of the subscale it belonged to. The highest possible score of correctly related items in each subscale was 35 (7 participants multiplied with 5 items = 35; see Table 1). In the subscale CD, respondents correctly related 30/35 items (86%); in CA, 14/35 items (40%); in CS, 18/35 items (51%); in CK, 23/35 items (66%); and in CE, 12/35 items (34%). Five items (#6, #7, #19, #20, and #23) were correctly related by all respondents, whereas 15 items were only correctly related by 4 (or less) respondents. The subscale CD had the maximum number of correctly related items. Four of the items in this subscale were correctly placed by 6 or 7 respondents. The items in subscales CA and CE showed the weakest results; only one item in these two subscales was correctly placed by 4 (or more) respondents. In subscale CS, there was one item (#22) that no respondent managed to place correctly. Two of the five items within subscale CS were correctly placed by 6 or 7 respondents.

Internal Validity Testing

To evaluate to what extent each item correlates with (a) the total summary measure, (b) the subscale it belongs to, and (c) its association with the other subscales, an item analysis on the internal structure was conducted. The results from the item analysis are presented in Table 2. The corrected item-total scale correlation varied between –.09 and .50. Only five items had a correlation of .40 and above, three belonged to subscale CD, and two to subscale CE, indicating weak correlation.

The corrected item-total correlation in each subscale also showed weak correlation. Subscale CD was the strongest subscale, varying between .11 and .52. Four items correlated >.40. Item 24, which had a correlation of .11 with the subscale CD, had a stronger correlation of .15 with subscale CE.

Table 1. Content Validity of the IAPCC-R Swedish Version[a]

Cultural Desire		Cultural Awareness		Cultural Knowledge		Cultural Skill		Cultural Encounter	
Item No.	Correctly Related Items	Item No.	Correctly Related Items	Item No.	Correctly Related Items	Item No.	Correctly Related Items	Item No.	Correctly Related Items
4	6	1	3	6	7	5	2	14	1
7	7	2	2	8	4	9	6	16	1
13	4	3	1	10	4	20	7	17	1
19	7	15	3	11	3	21	3	23	7
24	6	18	5	12	5	22	0	25	2
Total	30	Total	14	Total	23	Total	18	Total	12

Note: IAPCC-R = Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals–Revised). Each item has a maximum of seven successfully related items. Each subscale has a maximum of 35 successfully related items.
a. Successfully related items for the five subscales ($N = 7$).

Table 2. Item Analyses of the Internal Validity of the IAPCC-R Swedish Version for the Total Score and Five Subscales of Cultural Competence ($N = 334$)

	Item-Total Correlation					
IAPCC-R Item No.	Total Scale Corrected	Cultural Desire	Cultural Awareness	Cultural Skill	Cultural Knowledge	Cultural Encounter
Cultural desire (CD)						
4	**0.46**	**0.51**	0.11	0.25	0.08	0.39
7	**0.40**	**0.52**	0.07	0.17	0.10	0.40
13	**0.39**	**0.44**	0.16	0.22	0.09	0.31
19	**0.46**	**0.49**	0.05	0.28	0.17	0.39
24	**0.15**	**0.11**	0.12	0.10	–0.07	0.15
Cultural awareness (CA)						
1	–**0.08**	–0.05	**0.02**	–0.06	–0.06	–0.04
2	**0.15**	0.22	**0.08**	0.12	–0.09	0.19
3	–**0.06**	–0.01	**0.07**	–0.10	–0.05	0.06
15	**0.33**	0.30	**0.00**	0.24	0.17	0.27
18	–**0.06**	–0.08	**0.10**	–0.02	–0.07	0.02
Cultural skill (CS)						
5	**0.09**	0.20	0.04	–**0.01**	0.08	0.03
9	**0.31**	0.10	0.13	**0.15**	0.28	0.18
20	**0.14**	0.10	–0.13	–**0.01**	0.30	0.06
21	–**0.09**	0.03	0.00	–**0.20**	–0.13	0.10
22	**0.27**	0.19	0.03	**0.09**	0.15	0.21
Cultural knowledge (CK)						
6	**0.22**	0.03	–0.12	0.29	**0.42**	0.06
8	**0.39**	0.29	0.07	0.27	**0.29**	0.40
10	**0.37**	0.13	0.02	0.27	**0.48**	0.13
11	**0.13**	0.01	0.06	0.01	**0.19**	0.11
12	**0.20**	0.05	–0.05	0.20	**0.30**	0.08
Cultural encounter (CE)						
14	**0.43**	0.45	0.14	0.21	0.13	**0.29**
16	**0.06**	0.17	0.18	–0.02	–0.06	–**0.04**
17	**0.13**	0.14	–0.04	0.13	0.12	**0.10**
23	**0.50**	0.41	0.12	0.26	0.31	**0.32**
25	**0.16**	0.11	0.09	0.08	0.04	**0.07**

Note: IAPCC-R = Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals–Revised. The numbers in boldface represent the corrected item-total correlation.

In subscale CA, no item reached a correlation >.40, and four of five items correlated stronger with other subscales. No item in subscale CS correlated above .40, and four of five items correlated stronger with other subscales than with CS.

In CK, two items correlated >.40, and one item correlated >.40 with a different subscale from CK. In subscale CE, none of the items correlated ≥.40, and all the items correlated stronger with other subscales.

Moreover, the internal structure was further tested using CFA. First, a one-dimensional model was tested. The fit indices indicated poor model fit with RMSEA = 0.121, CFI = 0.595, SRMR = 0.106. Second, the five-dimensional model was tested, as defined by the five subscales. Again, the fit indices indicated poor model fit with RMSEA = 0.100, CFI = 0.691, SRMR = 0.0947. An inspection of the factor loading estimates revealed that an extremely low amount of variance was explained by the measurement model (more than 50% of the items share less than 10% of variance with the common factors).

As the hypothesized structure of the instrument could not be confirmed with confirmatory factor analysis, an exploratory factor analysis was conducted to uncover alternative common factors as those proposed by the cultural competence model. A principal axis factoring procedure with oblique rotation was performed, and this analysis suggested that the instrument contained eight factors (with eigenvalues >1), rather than five proposed by the model. However, these eight factors did only explain less than 50% of the variance, and the interpretation of the factors is not straightforward. Two factors may be highlighted here. The first factor includes items from all but one of the original five scales, that is, CD, CA, CS, and CE, and the items constituting the remaining scale are found in the second factor, that is, CK. This may be an indication that the first factor covers a large dimension of cultural competence and that cultural knowledge is independent of that general cultural competence factor. The other six factors are more difficult to interpret. The factor loadings are presented in Table 3.

Reliability Coefficient (Cronbach's Alpha)

Cronbach's alpha was .65 when summarizing all the items into one total scale. When summarizing the items into the five subscales, Cronbach's alpha varied between −.01 and .65 (see Table 4).

Discussion

In the present study, we identified three different problem areas in relation to validity and reliability. The problems were related to the *content* and *design* of the instrument and to the *sample.* Regarding the content of the instrument, we found that the instrument succeeded in measuring certain areas, but was less successful in measuring others. In subscales CE and CA, the respondents only managed to correctly place 12 and 14 items, respectively, of the possible 35, which indicates that the items did not measure what they were intended to. Problems with subscales were also established in the exploratory factor analysis, which revealed eight factors instead of the intended five. Some published studies are available on the instrument where evidence of acceptable content validity was found (Brathwaite, 2005; Campinha-Bacote, 1999). However, there was no description on how the content validity was conducted, which makes it difficult to draw any concrete conclusions from the results. It is worth noting that the latter study was conducted with the original instrument containing 20 questions (not measuring CD) and not with the revised version containing 25 questions.

We have identified some limitations to this study. The study sample originates from one large city in Sweden and, therefore, we cannot rule out the possibility that the instrument might work in a different context. Also, one could argue that the sample size of 334 might be quite small for some of the field tests. A larger sample could have produced different results. Therefore, we cannot be certain whether the instrument might work in other parts of Sweden or in specific hospital settings, such as elderly care. Also, no published material was available on the original instrument regarding interitem correlation, factor analysis, and so on, which makes it hard to draw any concrete conclusions about whether the results are due to our translation and adaptation of the original instrument or to the construction of the original instrument. Moreover, we did not have any comparable material from other international studies and, therefore, it is impossible to conclude whether the results are due to the fact that the instrument only works in a national rather than an international context.

We cannot draw any far-reaching conclusions from our results as there is a lack of information regarding validity of the original instrument. In this study, a factor analysis was conducted. However, we could not find any previous studies that had conducted the same analysis on the original instrument. Furthermore, very few similar studies were available from other countries where the instrument has been translated, adapted, and validity and reliability tested. This meant that we were unable to make comparisons with other reports. However, our conclusion is that this instrument is not valid and reliable for use in Sweden at this stage.

Furthermore, we only included nursing students and nurses in specialist programs, and this might be a further limitation as the weak results could be due to a lack of cultural competence in these two groups. However, the IAPCC-R instrument is constructed for use among various health care workers, such as nurses, nursing students, and nursing teachers (Campinha-Bacote, 2003a). Some studies have shown evidence of acceptable reliability (Brathwaite & Majumdar, 2006; Sargent, Sedlak, & Martsolf, 2005; Spencer & Cooper Brathwaite, 2003) and acceptable validity (Campinha-Bacote, 1999). However, a recent study from Taiwan reported weak reliability and a weak internal structure (Ho & Lee, 2007), whereas a study from the United States reported weak internal consistency regarding low item-total correlation (Vito, Roszkowski, & Wieland, 2005).

Table 3. Factor Loadings From the Principal Axis Factor Analysis With Oblique Rotation of the IAPCC-R Swedish Version (N = 334)

	Factor							
IAPCC-R Item No.	1	2	3	4	5	6	7	8
Cultural desire								
4	0.14	0.14	0.90	0.02	–0.03	0.04	–0.25	0.17
7	0.09	0.04	0.43	–0.18	0.18	0.09	–0.26	–0.05
13	0.58	–0.05	0.27	–0.29	0.68	0.20	–0.22	–0.17
19	0.65	0.06	0.30	–0.15	0.33	0.26	–0.21	–0.01
24	0.59	0.13	0.09	–0.23	0.49	0.42	0.08	–0.04
Cultural awareness								
1	–0.04	-0.07	-0.04	0.04	-0.04	-0.08	0.04	0.08
2	0.04	–0.01	0.04	0.02	0.28	–0.03	0.02	0.02
3	0.05	0.01	0.03	0.00	0.01	0.05	–0.01	0.38
15	0.64	0.10	0.21	–0.33	0.48	0.20	–0.23	–0.09
18	0.47	–0.08	–0.02	–0.36	0.46	0.31	0.01	0.20
Cultural skill								
5	0.07	–0.01	0.35	–0.13	0.13	0.12	0.14	–0.10
9	0.21	0.73	0.07	–0.12	0.15	0.04	–0.18	0.02
20	0.70	0.11	0.13	–0.51	0.16	0.18	–0.20	–0.16
21	0.55	0.08	–0.12	0.04	0.31	0.40	–0.08	0.26
22	0.70	0.22	0.09	–0.18	0.27	0.39	–0.12	0.02
Cultural knowledge								
6	0.11	0.30	0.12	–0.69	–0.04	0.15	–0.21	–0.02
8	0.17	0.22	0.20	–0.27	0.06	0.06	–0.60	–0.09
10	0.09	0.64	0.13	–0.23	–0.07	0.33	–0.24	0.15
11	0.12	0.07	0.05	–0.10	–0.03	0.51	0.02	0.04
12	0.10	0.60	0.04	–0.25	–0.07	0.13	–0.07	–0.24
Cultural encounter								
14	0.58	0.02	0.26	–0.26	0.56	0.22	–0.37	–0.07
16	0.57	0.08	0.04	–0.22	0.62	0.49	0.03	–0.00
17	0.64	-0.01	0.03	–0.17	0.32	0.31	–0.29	0.00
23	0.63	0.22	0.23	–0.23	0.35	0.55	–0.47	–0.04
25	0.71	0.27	0.08	–0.11	0.12	0.25	–0.06	0.26

Note: IAPCC-R = Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals–Revised.

Table 4. Cronbach's Alpha for the Total Score and the Five Subscale Scores of the IAPCC-R Swedish Version (N = 334)

Scale	Cronbach's α
Total score	0.65
Cultural desire (CD)	0.65
Cultural awareness (CA)	0.12
Cultural skill (CS)	–0.01
Cultural knowledge (CK)	0.56
Cultural encounter (CE)	0.31

Note: IAPCC-R = Inventory for Assessing the Process of Cultural Competence Among Healthcare Professionals–Revised.

Other problems with the content were found in relation to the field test on the response process, which showed that the respondents found it difficult to understand exactly what was being asked for, on several occasions. One question that focuses on the participant's understanding of the term *cultural competence* was especially difficult to grasp. One reason might be that respondents lacked theoretical knowledge and/or experience of what cultural competence actually means. One can argue that this was due to the fact that two thirds of the participants were nursing students. However, a basic explanation of these concepts was provided, as theoretical literature in transcultural nursing used in nursing programs in Sweden mainly originates from North America. Another explanation could be that our sample included a great number of nursing students in the early phase of their training program. However, district nurses in another Swedish study also found the term *cultural competence* difficult to relate to (Henriksson, 2006).

Other problems regarding the content were linked to how respondents were expected to interpret the terms *culture* and *cultural background*. The results showed that the questions relating to the term *cultural* were unclear. Furthermore, some questions were based on a group perspective, whereas others were not. Ahmad (1996) emphasizes the importance

of adopting a clear definition of the concept of culture, as this is the point of departure for the interpretation.

Problems in relation to the design of the instrument were identified in Items 3 and 11 as the word "not" had been added. Usually, positive and negative words are used as a technique to reduce biases when designing an instrument (Spector, 1992). Nevertheless, similar problems arose in a recent study by Vito et al. (2005) where the item had an inter-item correlation <.30, and when deleting them, Cronbach's alpha increased. Another well-known problem linked to the design of an instrument concerns items containing multiple factors or ideas (Spector, 1992). This problem was pointed out by our respondents as well as in the study by Vito et al. (2005).

Other design problems were identified in the field test on the internal structure of the instrument. Only a limited number of items correlated ≥.40 regarding the total scale and the subscale they belonged to. Instead, several items showed higher correlation with other subscales. In subscales CE and CA, no items reached a correlation >.40 within their own subscale, but five and four of the items, respectively, correlated to a higher degree with other subscales, which indicates that the instrument has a weak internal structure. In general, the factor analyses also suggested poor fit between model and data. The one- and five-dimensional measurement models did not fit the data, and further analyses revealed that the amount of common variance among the items was too low.

Problems with the sample were identified in the field test on the response process, which revealed that some respondents found it difficult to answer certain questions as they lacked clinical experience. The instrument is designed for use with registered nurses, advanced practice nurses, nursing students, and nursing teachers (Campinha-Bacote, 2003a). In the present study, some respondents stated that the instrument had been adapted for nurses working in clinical settings and *not* for nursing students and nursing teachers, as suggested when presenting the original instrument (Campinha-Bacote, 2003a). This was because some of the students lacked clinical experience, and some of the nurse teachers had not worked with patients (in the clinic) for several years. This indicates that the instrument might not be suitable for those who have no clinical experience/recent clinical experience. In the study conducted by Vito et al. (2005), the results indicated that work experience may lead to greater internal consistency.

Many questions emerged from the present study regarding the content, the design, and the intended sample. The different field tests provided us with evidence of validity from various perspectives. These studies verified both weaknesses and strengths regarding validity, and also some of the contradictions that arose, for example, in subscale CD. This subscale showed a high correlation on the internal structure and had the highest result in the field test on the content of the instrument. However, the field test on the response process revealed that the respondents interpreted that the area was measured by only three, not five, items as three questions were perceived as identical. This is a technique used when designing an instrument to check whether respondents answer "correctly" (Spector, 1992). Our reflection is, however: Why use the technique with three questions in one subscale and not in the other subscales? The contradictory results of the validity studies exemplify why it is important to examine the validity of an instrument by using more than one method.

Conclusion

This study represents a thorough evaluation of the Swedish version of the instrument IAPCC-R, and the results from our study indicate that it is not possible to use the instrument in the context for which we intended, that is, to measure cultural competence in Swedish nursing students and practicing nurses. However, we cannot conclude whether this is due to linguistic aspects based on the translation and/or the context in which it was evaluated, and/or whether it is due to the original construction. The inconclusive results of this study show the importance of contextual evaluation and adaptation of an instrument when translating it for use in other contexts. The global diversity of language and culture, as well as different types of social, structural, and institutional systems around the world challenges the notion of using "a universal tool" that can work globally. Stringent and rigid methods must be used for evaluating an instrument before adapting it for use in other contexts. More than one validity test of an instrument is crucial as different tests provide different insights.

Acknowledgments

The authors thank Jenny Larson for her valuable scientific advice, Christer Helsing for statistical consultation, and Caroline Sutton for the linguistic revision.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research and/or authorship of this article:
AMF Pension
The Department of Neurobiology, Care Sciences and Society
The Division of Nursing at Karolinska Institutet
Stockholms Sjukhem Foundation.

References

Ahmad, W. (1996). The trouble with culture. In D. Kelleher & S. M. Hillier (Eds.), *Researching cultural differences in health* (pp. viii, 244). London: Routledge.

Bäärnhielm, S. (2003). *Clinical encounters with different illness realities: Qualitative studies of illness meaning and restructuring of*

illness meaning among two cultural groups of female patients in a multicultural area of Stockholm. Unpublished doctoral dissertation, Karolinska Institutet, Stockholm.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*, 3186-3191.

Bernal, H., & Froman, R. (1987). The confidence of community health nurses in caring for ethnically diverse populations. *Image: Journal of Nursing Scholarship, 19*, 201-203.

Brathwaite, A. C. (2005). Evaluation of a cultural competence course. *Journal of Transcultural Nursing, 16*, 361-369.

Brathwaite, A. C., & Majumdar, B. (2006). Evaluation of a cultural competence educational programme. *Journal of Advanced Nursing, 53*, 470-479.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Campinha-Bacote, J. (1999). A model and instrument for addressing cultural competence in health care. *Journal of Nursing Education, 38*, 203-207.

Campinha-Bacote, J. (2003a). *The process of cultural competence in the delivery of healthcare services: A culturally competent model of care* (4th ed.). Cincinnati, OH: Transcultural C.A.R.E.

Campinha-Bacote, J. (2003b). Many faces: Addressing diversity in health care. *Online Journal of Issues in Nursing, 8*(1), 11.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309-319.

Conrad, F., & Blair, J. (1996). From impression to data: Increasing the objectivity of cognitive interviews. In N. Clyde Tucker (Chair) *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 1-10). Alexandria, VA: American Statistical Association.

Emami, A. (2000). *We are deaf, though we hear; we are dumb, though we talk; we are blind, though we see. Understanding Iranian late-in-life immigrants' perceptions and experiences of health, illness and culturally appropriate care.* Unpublished doctoral dissertation, Karolinska Institutet, Stockholm.

Emami, A., Benner, P., & Ekman, S. L. (2001). A sociocultural health model for late-in-life immigrants. *Journal of Transcultural Nursing, 12*, 15-24.

Gebru, K., & Willman, A. (2003). A research-based didactic model for education to promote culturally competent nursing care in Sweden. *Journal of Transcultural Nursing, 14*, 55-61.

Geisinger, K. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*, 304-312.

Giger, J. N., & Davidhizar, R. E. (2004). *Transcultural nursing: Assessment & intervention* (4th ed.). St. Louis, MO: Mosby.

Goodwin, L. D. (2002). Changing conceptions of measurement validity: An update on the new standards. *Journal of Nursing Education, 41*(3), 100-106.

Gorsuch, R. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Henriksson, M. (2006). *Cultural competence in Swedish primary care: Are some providers more prone to be culturally competent than others?* Unpublished master's thesis, Växjö University, Sweden.

Ho, M. J., & Lee, K. L. (2007). Reliability and validity of three cultural competency measures. *Medical Education, 41*, 519.

Jeffreys, M. R. (2000). Development and psychometric evaluation of the Transcultural Self-Efficacy Tool: A synthesis of findings. *Journal of Transcultural Nursing, 11*, 127-136.

Jöreskog, K. G. (2005). *Structural equation modelling with ordinal variables using LISREL* (Rev. ed.). Retrieved June 11, 2007, from http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf

Kim-Godwin, Y. S., Clarke, P. N., & Barton, L. (2001). A model for the delivery of culturally competent community care. *Journal of Advanced Nursing, 35*, 918-925.

Leininger, M., & McFarland, M. R. (2002). *Transcultural nursing: Concepts, theories, research and practice* (3rd ed.). New York: McGraw-Hill.

Lundberg, P. C. (2000). Cultural care of Thai immigrants in Uppsala: A study of transcultural nursing in Sweden. *Journal of Transcultural Nursing, 11*, 274-280.

Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: a methods review. *Journal of Advanced Nursing, 48*, 175-186.

Polit, D. F., & Beck, C. T. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia: Lippincott Williams & Wilkins.

Purnell, L. D., & Paulanka, B. J. (2003). *Transcultural health care: A culturally competent approach*. Philadelphia: F. A. Davis.

Sargent, S. E., Sedlak, C. A., & Martsolf, D. S. (2005). Cultural competence among nursing students and faculty. *Nurse Education Today, 25*, 214-221.

Schim, S. M., Doorenbos, A. Z., Miller, J., & Benkert, R. (2003). Development of a cultural competence assessment instrument. *Journal of Nursing Measurement, 11*, 29-40.

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, CA: Sage.

Spencer, W., & Cooper Brathwaite, A. (2003). *Testing the reliability and validity of Campinha-Bacote's inventory for assessing the process of cultural competence among healthcare professionals-revised.* Unpublished manuscript, University of Toronto, Ontario, Canada.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.

Vito, K., Roszkowski, M., & Wieland, D. (2005, September). *Measuring cultural competence as a curriculum outcome: What we learned from our experiences with two tools.* Paper presented at the National League for Nursing (NLN) Education Summit, Baltimore, MD.