

Content validity and its estimation

Yaghmaie F., PhD

Faculty member of Nursing and Midwifery School, Shahid Beheshti University of Medical Sciences and Health Services

ABSTRACT

Background: Measuring content validity of instruments are important. This type of validity can help to ensure construct validity and give confidence to the readers and researchers about instruments. content validity refers to the degree that the instrument covers the content that it is supposed to measure. For content validity two judgments are necessary: the measurable extent of each item for defining the traits and the set of items that represents all aspects of the traits.

Purpose: To develop a content valid scale for assessing experience with computer usage.

Methods: First a review of 2 volumes of International Journal of Nursing Studies, was conducted with only 1 article out of 13 which documented content validity did so by a 4-point content validity index (CVI) and the judgment of 3 experts. Then a scale with 38 items was developed. The experts were asked to rate each item based on relevance, clarity, simplicity and ambiguity on the four-point scale. Content Validity Index (CVI) for each item was determined.

Result: Of 38 items, those with CVI over 0.75 remained and the rest were discarded resulting to 25-item scale.

Conclusion: Although documenting content validity of an instrument may seem expensive in terms of time and human resources, its importance warrants greater attention when a valid assessment instrument is to be developed.

Keywords: Content Validity, Measuring Content Validity

Journal of Medical Education Spring 2003; 3(1):25-27

Introduction

Some authors in their articles have reported the process of measuring content validity frequently, while others did not document this type of validity. Measuring and reporting content validity of instruments are important. This type of validity can also help to ensure construct validity and give confidence to the readers and researchers about instruments. Content validity is used to measure the variables of interest. It is also known as: content related validity, intrinsic validity, relevance validity, representative validity and logical or sampling validity. It can be used to measure the appropriate sampling of the content domain of items in a questionnaire.

Kerlinger (1) argues that content validity is representative of the content. Thus, content validity of an instrument depends on the adequacy of a specified domain of content that is sampled (2). Bush (3) pointed out that content validity refers to the degree that the instrument covers the content that it is supposed to measure. It also refers to the adequacy of the sampling of the content that should be measured (4). Therefore, content validity measures the comprehensiveness and representativeness of the content of a scale.

Nunnally(2) argued that there are two standards for ensuring content validity: firstly, the sampling of the items and secondly, the method of constructing the items. For content validity two judgments are necessary: the measurable extent of each item for defining the traits and the set of items that represents all aspects of the traits.

Measuring and analyzing content validity

Burns and Grove stated that content validity "is obtained from three sources: literature, representatives of the relevant populations, and experts".(5) Content validity could also be established in two stages; development and judgment stage.

It is also stated that addressing content validity should begin with instrument development(5). The first step of instrument development is to identify 'what domain of construct' should be measured. This can be determined through literature reviews, interviews and focus groups. By determining a precise definition of traits of interest, a more clear picture of limitations, dimensions, and components of the subject can be reached. The qualitative method can be helpful for determining the domain and concepts of construct that are of interest here. There is no complete objective method for determining the content validity of an instrument

nor is any statistical approach (4,6). However, content validity in the judgment stage is based on quantitative evidence (7).

To examine the content validity in judgment stage, professional subjective judgment is required to determine the extent to which the scale was designed to measure a trait of interest (2). Ghiselli pointed out that content validity depends on subjective or professional judgment (8). Content validity is a subjective judgment of experts about the degree of relevant construct in an assessment instrument. However, inclusion of at least five experts in that field (5) or five to ten experts (7) would be useful to judge the content domains of a scale through use of rating scales.

Material and Methods

First a review of 38 articles that published in the International Journal of Nursing Studies, 1995, 1996 (volumes 32 and 33), was conducted finding 13 (20.48%) articles that discuss content validity. Of these articles, 1 study's content validity was based on the previous studies. Six measured content validity only based on the opinion of experts (from 1 to 10 experts) for accuracy, completeness, clarity relevance, scoring system, compressive and length of questions. One of the study measured content validity only by the review of literature and an expert panel without any explanation about the process of the measuring. Finally, 1 article measured content validity by a 4-point content validity index (CVI) and the judgment of 3 experts. To generate a representative sample of items of computer experience, a host of procedures was used: literature review, focus groups and interviews. At this stage the scale was developed with 38 items.

To test the content validity of the scale, ten experts from the Department of Computer Science, the Department of Nursing at the University and the Community Health Centers of Australia. The participants used computer in their routine work or in teaching. The researcher gave a copy of the scale and explained the purpose and objectives of the study to them individually. The Content Validity Index (CVI) developed by Waltz and Bausell was used (8). The experts were then asked to rate each item based on relevance, clarity, simplicity and ambiguity on the four-point scale. (table1)

TABLE 1. Criteria for Measuring Content Validity

1.Relevance

- 1 = not relevant
- 2 = item need some revision
- 3 = relevant but need minor revision
- 4 = very relevant

2.Clarity

- 1 = not clear
- 2 = item need some revision
- 3 = clear but need minor revision
- 4 = very clear

3.Simplicity

- 1 = not simple
- 2 = item need some revision
- 3 = simple but need minor revision
- 4 = very simple

4.Ambiguity

- 1 = doubtful
- 2 = item need some revision
- 3 = no doubt but need minor revision
- 4 = meaning is clear

Results

The researcher analyzed the results of the content validity of the scale. The items that had CVI over 0.75 remained and the rest were discarded. the remaining items were modified, based on the experts' opinions. By discarding those items of the scale that were not related to the domain of computer experience, the number of items decreased from 38 to 25.

Discussion

Our review of 38 articles that published in the International Journal of Nursing Studies, 1995, 1996 (volumes 32 and 33) shows that content validity is not regarded so important, since only 20% of articles documented content validity.

Researchers should improve a more comprehensive view of content validity. Statements such as: 'the validity of the instrument had been tested by comments of experts' or "content validity was determined through a review of literature or panel experts' are unacceptable and invalid statements.

Content validity is an important factor in identifying the concept of measuring; however, it is not a sufficient indication that the instrument

actually measures what is that intended to measure. Finding from content validity could contribute to support the construct validity of an instrument. A single approach is insufficient and a variety of approaches should be tested. By documenting the content validity of the instrument that has been used, the reader can understand the process of measuring content validity. By measuring content validity, the interpretations of results are precise.

References

- 1-Kerlinger FN. Foundations of behavioral research. 3rd ed. New York: CBS Publishing; 1986.
- 2-Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill Book Company; 1978.
- 3.Bush CT. Nursing research. Virginia: Reston Publishing Company; 1985.
- 4-Polit DF, Hungler BP. Nursing research principles and methods. 4th ed. Philadelphia: JB Lippincott Company; 1991.
- 5.Burns N, Grove SK. The practice of nursing research conduct, critique, and utilization. 2nd ed. Philadelphia: WB Saunders Company; 1993.
- 6-Dempsey PA, Dempsey AD. The research process in nursing. 2nd ed. Boston: Jones and Bartlett Publishers; 1986.
- 7-Wilson HS. Research in nursing. 2nd ed. California: Addison-Wesley Publishing Company; 1989.
- 8-Waltz CF, Bausell RB. Nursing research: Design, Statistics and Computer Analysis. 2nd ed. Philadelphia: FA Davis Company; 1983.
- 9-Ghiselli EE. Theory of psychological measurement. New York: McGraw-Hill Book Company; 1964.