REVIEW ARTICLE

Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales

Christophe Goetz[a], Joël Coste[b], Fabienne Lemetayer[c], Anne-Christine Rat[a], Sébastien Montel[c], Sophie Recchia[c,d], Marc Debouverie[a], Jacques Pouchot[b], Elisabeth Spitz[c], Francis Guillemin[a,*]

[a]*Université Paris-Descartes, Sorbonne Paris Cité, Université de Lorraine, EA 4360 APEMAC, Nancy, France*
[b]*Université Paris-Descartes, Sorbonne Paris Cité, Université de Lorraine, EA 4360 APEMAC, Paris, France*
[c]*Université Paris-Descartes, Sorbonne Paris Cité, Université de Lorraine, EA 4360 APEMAC, Metz, France*
[d]*Département de Psychologie, Université du Luxembourg, Luxembourg, Luxembourg*

Abstract

Objective: To review current practice and update guidelines for the methodology of shortening composite measurement scales (CMSs).

Study Design and Setting: A literature review gathered data on 91 shortening processes from 1995 to 2009. The validity of the initial CMS, the shortening methods, and the validity of the derived short-form scales were examined. The results were compared with those from a previous literature review (articles from 1985 to 1995) to develop updated guidelines for CMS shortening.

Results: The literature review revealed a persisting lack of use of rigorous methodology for CMS shortening. Of the 91 cases of CMS shortening, 36 combined a content approach and a statistical approach; 45 used only a statistical approach and 10 (11%) only a content approach. The updated guidelines deal with the validity and conceptual model of the initial CMS, the preservation of content and psychometric properties during shortening, the selection of items, and the validation of the short form.

Conclusion: Item reduction based on a rigorous methodology is necessary if the short-form instrument aims to maintain the validity and other measurement properties of the parent instrument, which in turn supports application in research and clinical practice. © 2013 Elsevier Inc. All rights reserved.

Keywords: Composite measurement scale; Shortening; Short form; Psychometrics; Methodology; Recommendations

1. Introduction

Many health constructs are too complex to be captured by direct measurement. When these constructs need to be examined, one of the most popular methods requires the use of a composite measurement scale (CMS). The CMS generally consists of items or questions that assess one or several attributes scored by a scale.

Over the years, the measurement of health constructs has led to the production of a large number of scales, with often a high number of items [1].

The burden of long scales and the increasing need for multiple instruments in the same study have logically created a strong need to shorten CMSs. To shorten a CMS consists in reducing its number of items while trying to preserve or improve its psychometric properties. The methodology to develop new CMSs is well documented [2] and includes consideration of different psychometric properties but has limited applicability as a means of informing the shortening of instruments. Guidelines for shortening existing CMSs are scarce. In 1997, Coste et al. [3] noted that most articles reporting on scale shortening lacked rigorous methodology: shortening processes were often inadequately conceptualized, and excessive credit was given to statistical techniques. The authors recommended carefully choosing the original scale according to its content, its possibility for shortening, and its psychometric properties; focusing on criterion validity if the original scale can be considered the gold standard and if not, preferring an expert-based approach to content validity, possibly helped by statistical considerations; and finally, performing a validation study in an independent sample. In 2000, Smith et al. [4] also noted methodological pitfalls concerning CMS shortening in psychology and recommended using a validated original scale, clarifying the intended use of the short form, estimating

---

What is new?

- Current practice of shortening composite measurement scales (CMSs) still often lacks methodological rigor.

- An update of previous guidelines is proposed to improve the quality of the shortening process and thus the validity of the resulting short-form scales.

- The main items to check when shortening a CMS are:

  1. Document the validity of the original CMS and the objective of its shortening

  2. Take the conceptual model into account,

  3. Preserve content validity,

  4. Preserve psychometric properties,

  5. Document the reasons for item selection, and

  6. Validate the short-form CMS in an independent sample.

a priori the properties of the short form to balance resource or time savings against the loss of validity, preserving content, and using an independent sample to validate the short form. In 2002, Stanton et al. [5] also observed a lack of methodological recommendations for scale reduction. The authors argued that focusing on internal consistency should be avoided and proposed a set of item ''quality indices'' to help conceptualize the competing issues that influence item retention decisions. However, new approaches and new statistical methods such as item response theory (IRT) are now used for scale shortening [6], but we lack published recommendations integrating these methods. Another shortfall in current guidelines concerns content validity; retaining content during the shortening process is important, but only methodological guidelines for developing new CMSs focus on content analysis [7,8].

This article aims to describe the methodology currently used to shorten CMSs through a literature review and to compare with a previous review for proposing updated and structured guidelines for CMS shortening.

2. Methods

2.1. Search and identification of articles

Articles reporting on the development of a short form of an existing CMS concerning health or psychology domains, published between January 1, 1995 and December 31, 2009 and written in English, were selected from *MedLine* and *Psycinfo*. The following query was used to search title,

abstract, and keyword fields: (''*short form*'' or ''*brief form*'' or ''*short version*'' or ''*brief version*'') and (*questionnaire* or *scale* or *instrument*) and (*development* or *validation* or *reduction* or *shortening* or ''*item reduction*'' or ''*item selection*''). Reference lists of selected articles were also checked. The CMSs were included regardless of their measurement objective (evaluation, discrimination, or prediction) and completion type (self-, interviewer-, or computer-completed/assisted). The CMSs not related to health or psychology were excluded, as were short-form CMSs developed from two or more different scales and refinement of existing short-form CMSs.

To analyze the changes in scale-shortening methods over time, the results of this literature review were compared with those from a review of articles published between 1985 and 1995 (both studies conducted with similar methodology) [3].

2.2. Data collection

All articles reporting on the development or validation of an identified short scale were analyzed together. Data about the characteristics and psychometric properties of the original CMS, the reason and methodology for the shortening process, and characteristics and validity of the short form were collected. The psychometric properties were described according to four main types of quality criteria, namely (1) content validity, (2) criterion and construct validity (convergent and discriminant validity, factorial validity, or internal consistency), (3) inter-/intraobserver or test–retest reliability, and (4) sensitivity to change or responsiveness [2,9]. Additional attention was paid to the use of IRT methods in the shortening process and validation of the short form in an independent sample.

2.3. Data extraction and analysis

Seven reviewers were involved in the literature analysis: four were epidemiologists and three health psychologists. Each article was independently analyzed by a randomly chosen pair of reviewers, one from each field. Differences were discussed in a plenary session to achieve consensus.

A standardized extraction form (SEF) was elaborated on the model used for the previous literature review [3]. A pilot study was conducted to test and refine the SEF. Twelve papers were analyzed, and each reviewer was asked to point out any difficulties or make suggestions. This process led to minor changes of the SEF; some items were reformulated to clarify their objectives, and two items were added (refinement of response categories and linking between content and statistical analyses).

Agreement between readers before achieving consensus was examined by computing the percentage of studies with perfect concordance of the readers' answers and kappa coefficients. To assess the evolution of shortening methodology over time, we compared the time periods