Journal of Psychosomatic Research 75 (2013) 167–172

Contents lists available at ScienceDirect

Journal of Psychosomatic Research

Shortening the S-STAI: Consequences for research and clinical practice

Peter M. Kruyen [a,*], Wilco H.M. Emons [b], Klaas Sijtsma [b]

[a] Radboud University Nijmegen, Institute for Management Research, The Netherlands
[b] Department of Methodology and Statistics, Tilburg University, The Netherlands

ARTICLE INFO

Article history:
Received 7 August 2012
Received in revised form 21 March 2013
Accepted 24 March 2013

Keywords:
Assessment of individual change
Individual diagnosis
Measurement precision
State-Trait Anxiety Inventory
Test shortening
Total-score reliability

ABSTRACT

Objective: Several authors proposed a shortened version of the State scale of the State-Trait Anxiety Inventory (S-STAI) to obtain a more efficient measurement instrument. Psychometric theory shows that test shortening makes a total score more vulnerable to measurement error, and this may result in inaccurate and biased research results and an increased risk of making incorrect decisions about individuals. This study investigated whether the reliability and the measurement precision of shortened versions of the S-STAI are adequate for psychological research and making decisions about individuals in clinical practice.
Methods: Secondary data analysis was used to compare reliability and measurement precision between twelve shortened S-STAI versions and the full-length 20-item S-STAI version. Data for the 20-item version came from a longitudinal study performed previously in the Netherlands and included 377 patients and 375 of their family members. This was our master data set. A literature study was conducted to identify shortened S-STAI versions that are used in research and clinical practice. Data for each shortened version were obtained from the master data set by selecting the relevant items from the 20-item version. All analyses were done by means of classical test theory statistics.
Results: The effect of test shortening on total-score reliability was small, the effect on measurement precision was large, and the effect on individual diagnosis and assessment of individual change was ambiguous.
Conclusion: We conclude that shortened versions of the S-STAI seem to be acceptable for research purposes, but may be problematic in clinical practice.

Introduction

The state scale of the State-Trait Anxiety Inventory (S-STAI) [1,2] is one of the most widely used measures of state anxiety, both in scientific research and clinical practice [2–4]. Over the years, several authors [5,6] have proposed shortened versions of the original 20-item S-STAI. Shortened S-STAI versions may be preferred for at least three reasons. First, test shortening may alleviate the respondent's burden and save testing time [7]. Second, shortened S-STAI versions may raise response rates in large-scale anxiety surveys and epidemiological research [8]. Third, asking respondents to answer items from the 20-item S-STAI having overlapping content may be considered unethical.

The use of shortened S-STAIs also has serious drawbacks. When removing psychometrically sound items, total scores tend to be less reliable and provide less precise measures of the underlying attribute [9–12]. Using unreliable scores may also jeopardize test validity. In clinical applications, using less reliable, shortened S-STAI versions to diagnose highly-anxious individuals who are in need of special treatment [13] increases the risk that patients do not receive the best treatment available. Practitioners also use S-STAI scores to monitor change in the person's anxiety level during treatment [14]. Because test shortening reduces the reliability of both pre-test and post-test scores, the impact test shortening has on reliability will be even greater for change scores than for the separate pre-test and post-test scores. As a result, observed change within a person may not be distinguishable from measurement error and sound conclusions about treatment effects are impossible.

The goal of this study was to compare reliability and measurement precision of the full-length 20-item S-STAI version and shortened S-STAI versions. In particular, we investigated the consequences of test shortening for individual decision-making, which received little attention thus far. To accomplish our goal, we examined shortened S-STAI versions that were reported in the literature. The following research questions were addressed:

1. What are the differences between total-score reliability and measurement precision of the original 20-item S-STAI and total-score reliability and measurement precision for the shortened S-STAI versions?
2. To what extent does shortening the S-STAI increase the risk of drawing incorrect conclusions about an individual's state anxiety?
3. To what extent does shortening the S-STAI reduce the power to find significant changes in anxiety over time?

* Corresponding author at: Radboud University Nijmegen, Institute for Management Research, PO Box 9102, 6500 HC Nijmegen, The Netherlands. Tel.: +31 243611934.
E-mail address: p.m.kruyen@fm.ru.nl (P.M. Kruyen).

0022-3999/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jpsychores.2013.03.013

Method

Participants

We analyzed an existing data set to address the three research questions. The data came from a longitudinal study at the Erasmus Medical Center, Rotterdam, the Netherlands [15]. The Rotterdam study aimed at finding explanations for fluctuations in state anxiety and device-related concerns in patients after receiving an implantable cardioverter-defibrillator (ICD). It was found that anxiety levels were stable across time, but device-related concerns fluctuated. Personality characteristics and amount of social support explained individual differences in anxiety and device-related concern trajectories.

Participants in the Rotterdam study were cardiac patients who received an ICD between 2003 and 2008 at the Erasmus Medical Center, and for each patient one of their family members (often their partner). An ICD nurse approached patients and their family members for participation in the study [15]. Patients and partners were excluded when they had insufficient knowledge of the Dutch language, were on the waiting list for heart transplantation, had a life expectancy of less than a year, or had a history of psychiatric illness other than affective anxiety disorders. All respondents provided written informed consent. The ethical commission of the Erasmus Medical Center in Rotterdam approved the study.

The participants filled out the full-length 20 item-version of the S-STAI at five time points: one day before ICD implantation, and ten days, three months, six months, and one year after implantation. Because more than 50% of the respondents did not complete the S-STAI at all five measurement occasions, we only used data collected at the first two time points (i.e., one day before and ten days after ICD implantation). For these two time points, we had access to data from 377 patients (298 males; 79 females; mean age = 57.69 years, age ranged from 17 to 81 years) and 375 partners (81 males; 294 females; mean age = 55.57 years, age ranged from 20 to 90 years). Few respondents had missing item scores; 0.51% of the scores in the patient sample were missing, and 0.36% in the partner sample. We used two-way imputation to estimate missing item scores [16]. The resulting data set was the master data set.

Measures

Full-length 20-item S-STAI. The data were collected by means of a Dutch translation of the second edition of the S-STAI published in 1983 [17]. The scale comprised ten items that address state anxiety (e.g., "I feel tense") and ten counter-indicative items of state anxiety (e.g., "I feel calm"). Respondents rated their answer on a 4-point scale ranging from 1 (i.e., *not at all*) to 4 (i.e., *very much so*). The 1983 edition of the S-STAI succeeded the first edition that was released in 1970. Because six items in the original 1970 edition had a stronger relation to depression than to anxiety or had poor psychometric properties, in the 1983 edition of the S-STAI [2] those items were replaced by new items. Nowadays, in practical anxiety measurement both editions are used [18,19].

Procedure for selection of shortened S-STAI versions

We did a literature review to identify shortened S-STAI versions that were either used for research purposes or in clinical practice. Because we used data from the 1983 S-STAI version, we only included shortened versions of the 1983 edition in our study. Moreover, we only included shortened versions for which we could exactly retrieve the constituent items. Four databases were explored, CINAHL, Google Scholar, MEDLINE, and PsycINFO (on October 22, 2012). We used the Boolean expression ((state trait anxiety inventory) and (shortened OR short OR abbreviated)) to find articles which, in addition to "State-Trait Inventory", included at least one of the key words "shortened" "abbreviated", or "short" in the article body. This query returned a large number of hits (i.e., over 400 results in CINAHL, even 77,000 hits in Google Scholar, over 200 hits in MEDLINE, and more than 1000 results in PsycINFO).

To reduce the number of hits to a manageable size, we searched for the articles that included the key words in their title. This search strategy retained only seven articles, two of which proposed a shortened version of the 1983 edition of the S-STAI. However, given the widespread use of the S-STAI and the popularity of shortened tests in general [20], we presumed that more shortened S-STAI versions would circulate. Therefore, we retrospectively examined the seven articles from the title-based search for references to other shortened S-STAI versions. For example, we looked for phrases such as "another shortened S-STAI version was developed by…" and "short versions of the S-STAI were also proposed by…". This strategy yielded another fourteen articles. Inspection of the articles returned another ten shortened 1983 S-STAI versions for which the constituent items were reported. Examination of these fourteen articles for references to other shortened versions did not reveal more shortened S-STAI versions used in practice. Therefore, we addressed the research questions for the twelve shortened S-STAI versions, the results of which we believe provide a representative picture of the psychometric properties of shortened S-STAI versions used in practice.

Data analysis

For each shortened S-STAI version, we identified the relevant items that appeared in the 20-item test and selected the corresponding data columns from the master data set. The selected data sets were used to address the research questions. Because the psychometric properties of scales, and thus the effects of test shortening, often depend on the population envisaged [9], we addressed the research questions separately for patients and their partners.

Total-score reliability and measurement precision

To compare total-score reliability of the shortened versions, we used coefficient alpha and Guttman's [21] lambda2. Coefficient alpha is the most frequently used measure of total-score reliability in applied research [22]. Both coefficients alpha and lamda2 are lower-bounds to the tests' population (true) reliability, but lambda2 is at least as large as coefficient alpha and thus in the population it is closer to the test's true total-score reliability [22].

For the comparison of measurement precision across shortened versions, we used 90% confidence intervals (CIs) [22]. The CI reflects the error margins users have to take into account when they draw conclusions about the individual's true score T based on total score X_+. We used coefficient alpha to compute the standard error of measurement (SEM). To fully appreciate the relationship between test length and measurement precision, it is important to take the total score range into account [22]. Removing items reduces both the score range and the SEM. However, under reasonable conditions the score range decreases faster than the SEM [22]. Hence, as the scale grows shorter, the SEM becomes large relative to the score range, and CI that fully depends on SEM grows larger relative to the score range. This means that total score X_+ provides a less precise estimate of the true score T [22,23]. To assess the impact of test shortening on measurement precision, we computed relative CI [23], which is defined by the ratio of CI length and score range. As relative CI increases, measurement precision decreases.

Classification consistency

To study the increased risk of drawing incorrect conclusions about an individual's state anxiety, we compared the extent to which shortened S-STAI versions classify individuals with at least 90% certainty into one of two diagnostic categories. To do so, we assumed that (the hypothetical event of) independently administering the test