Reviewer's report

Title: Assessing the Safety Attitudes Questionnaire (SAQ), German language Version in Swiss University Hospitals - A Validation study

Version: 3 Date: 11 October 2012

Reviewer: Barbara Hoffmann

Reviewer's report:

Thank you for having again the opportunity to review the re-revised manuscript. It improved again. Unfortunately, some concerns still remain. I still appreciate the value of the study and that the article should be published but it still needs improvement.

Major Compulsory Revisions

Generally, I am concerned about the stated objectives of the study (to explore the psychometric properties of the translated questionnaire) and the somewhat problematic method applied (in particular the collapsing of a symmetric 5-point-likert-scale into a three-point-asymmetric scale in order to meet prerequisites for confirmatory factor analysis). In addition, normalisation procedures were stated to have been made but without any details about which procedures have been conducted and why they failed. Since this is the base of the analysis it may be helpful if this aspect could be reviewed by an expert in this field.

Minor Essential Revisions

# Abstract/background: It still says that the SAQ measures patient safety DIMENSIONS and not safety climate.

# Background: page 3, 2.paragraph last sentence: "Although, conceptual differences between safety culture and safety climate EXIST(?) the two terms are often used interchangeable in the literature."

# The SAQ measures safety climate and therefore the PERCEPTION of workers of safety attitudes but not the attitudes of workers (Sexton et al (2006)BMC Helath Serv Res 6:44; Shtynberg et al AHRQ technical report). Climate is an aspect of a group but not of individuals.

# Method/measures page 5: “We decided to use the 30 items version of the SAQ because of its usability, good psychometric properties as shown in previous studies (Raykov’s rho 0.9, Cronbach’s 0.85 and adjusted goodness of fit indices between 0.871) and broad implementation.” Is there something missing between 0.871 and ???.

# I did not find an answer to research question R2: Are the items clear and easy to understand?

# Results/content validity page 10: “The scale-level content validity index (S-CVI)
was 0.83, indicating good content validity (research questions 1 and 2).” Is this the overall S-CVI or a mean of the six S-CVIs?

# Results and discussion: I did not find an explanation for the variable modification index. Please give details how this index is assessed and information on interpretation of values.

# Discussion page 12: “The results showed moderate to strong reliability and acceptable to good validity indices, with the exception of few items, especially in the perception of management sub scale (Table 6), which require additional investigation and refinement.” Good structural validity indices or goodness of fit indices? Then, later they are said to be moderate. Please rewrite the phrase more clearly.

# Discussion page 13, end of page: “Item 25 about “problem personnel” shows high loadings for all factors and seems to have an impact on safety culture, but also on perception of management, job satisfaction as well as on teamwork climate.” What does the impact on safety culture and the other scales mean, what is the meaning of this sentence?

# Discussion page 14, head of page: „This result differed from other studies and needs to be discussed within an expert panel and translation experts.” What are the results of these discussions? What is the opinion of the authors? Otherwise, please delete the phrase.

# Discussion page 14: “The high and significant factor loadings for 5 of the 6 previously identified factors suggests that with minimal translational improvements and cultural adaptations (targeting especially the perception of management scale), the SAQ German version will fit the original factor structure even better, a hypothesis which will be tested in future studies.” I would like to recommend to be more cautious: The SAQ MAY fit the original factor structure…

# Discussion page 14: “In addition to acceptable internal validity, the model showed strong reliability.” What is meant by internal validity and by reliability – internal consistency (cronbach alpha, item-item-correlation, item-factor-correlation), structural validity (fit indices)?

# Discussion, limitation: There seem to be a number of limitations to the study that are dependent on each other. The questionnaire had been applied to a 1) highly selected 2) convenience sample, that 3) had been smaller than expected. In addition, 4) the likert-scale had been collapsed due to highly positively skewed item responses to a scale with one negative end and two positive ones.

# Therefore, this study seems to be one step on the way from questionnaire development to analysis of psychometric properties but the final analysis will be the next one to follow.

Discretionary Revisions

Method/measures page 5: May be it is less irritating if you treat the items that are responded to regarding the unit and the hospital level as separate items leading to 34 items of the core safety climate scale of the SAQ.
Please make sure the manuscript will be proof-read in the end. I still found some orthographic mistakes (e. g. abstract/conclusions), words missing and some abbreviations used in the manuscript without explanations (e. g. RN, MD, PSCL).

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.