PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Hospital Survey on Patient Safety Culture: Psychometric evaluation in Kuwaiti Public Healthcare Settings
AUTHORS	Alsalem, Gheed; Bowie, Paul; Morrison, Jill

VERSION 1 - REVIEW

REVIEWER	Stig Harthug University of Bergen, Norway
REVIEW RETURNED	01-Feb-2019

GENERAL COMMENTS	The is an important contribution to the field of measuring safety culture or climate. In contrary to several other studies, the present study applies both EFA and CFA in randomely sampeled halves of the studypopulation. The use of these methods is well performed and all relevant statistical information is provided. The authors have choosen HOSPSC because it seems to used in many contries and therefor it should be possible to compare the results worldwide. However their data from the responders do not fit with the original datastructure and by appllying EFA they find an more feasible and simplified model comprising 8 factors buildt on 22 items. There are two main concerns: two factors are bases on only two items each and the new factor structure makes it difficult to compare their results with results from other contries. Both concerns are properly discussed. Questions to the authors: 1: The authors provides different definitions for safety culture and safety climate, but claim that thei have a pragmatic view and use both words with the same meaning. It would be better to use only one of them consequently in the paper. 2: The authors metion that different health care system settings could give different results, but cuold be more presise. 3: The have a high response rate >80% but do not discuss that especially the studies from USA have much lower response rate. Could the low respons rate have impact on the results and the differenses in stucture? 4: It is not menitioned in which language the present questionnaire was presented for the HCWs. English of a local language?

REVIEWER	Zenewton A. S. Gama Professor of Collective Health. Federal University of Rio Grande do Norte, Brazil.
REVIEW RETURNED	07-Feb-2019

GENERAL COMMENTS	The study is about a very relevant topic for the quality of health services, the objectives are well defined and the method is appropriate. I have only a few concerns, which can be clarified or corrected by the authors. Major concerns: 1- The final version presented is the smallest version of HSOPSC (8 dimensions and 22 items) among all those mentioned in Appendix 5. The large number of items discarded (20 items) raises doubts about the reasons for discarding intermediate versions with more items and dimensions , but the study does not have enough data on these intermediate versions. I therefore ask you to consider these suggestions: • Page 9, lines 48-55 and page 10 lines 3-9: The sample size of the CFA used to test the original HSOPSC (12-factor) model was not presented. I imagine you have used the whole sample, as well as the internal consistency analysis, but it is important to include this information. • Page 10, lines 43-54. Specify in the text which was the main statistical criterion and cut-off point used to exclude certain item (factorial load of item <0.40?). • Table 3. Expand data with additional rows with EFA results for models with 9 factors, 10 factors, 11 factors and 12 factors. Also include an additional column that presents a rationale, based on statistics or other factors, of not choosing the model. 2- In the discussion (page 15, lines 50-56), consider that a Brazilian study confirmed the 12-dimensional structure of the HSOPSC (Andrade et al., 2017 http://scielo.iec.gov.br/pdf/ess/v26n3/en_2237 -9622-ess-26-03-00455.pdf). This study was published a few months after reference 74 by Reis et al., Used an independent transcultural translation and adaptation process, a different data collection method (electronic questionnaire) and a sample size approximately 3 times higher. Minor concerns: 1- Page 4 of 35, lines 23-25: On "Lack of reporting of explicit psychometric data in some comparative studies used in the comparative analysis was another challenge faced in our study", this is not a limitation of this study, but of other studies. 2- Page 25 of 25, box 1. You split two boxes with the same title. I suppose the title of the second one is different from the first one (service and hospital?). 3- Specify in which language the questionnaire is evaluated. Arabic?

VERSION 1 – AUTHOR RESPONSE

Questions to the authors:

1: The authors provide different definitions for safety culture and safety climate, but claim that they have a pragmatic view and use both words with the same meaning. It would be better to use only one of them consequently in the paper.

The literature surrounding safety culture is vague with no clear definitions, dimensions and theoretical basis. There are numerous definitions of safety culture and safety climate. Yet, despite their distinctive terminologies, they are commonly used interchangeably in the literature. Safety culture and safety climate were used interchangeably in certain parts due to the fact that the main tool adopted in my paper uses the term safety culture in its name; Hospital Survey on Patient Safety Culture (HSOPSC). Based on the reviewer's recommendation, I have used safety climate instead of safety culture where I found it to be more appropriate (P.5).

2: The authors mention that different health care system settings could give different results, but could be more precise.

Due to space limitations, I was not able to fully expand on the raised point. I have mentioned two paragraphs in p. 16 and p.17 where I pointed out that the discrepancy in results could be attributed to the neglect of crucial elements, including context, processes and actors involved, when attempting to adapt an instrument in a different setting might lead to conflicting results and might weaken the validity of the instrument 77. The following section was added: "This might be due to the contextual specificity of the construct of safety culture (Coyle et al., 1995). Other factors include unique country characteristics, types of health systems and settings, staff groups, and cultural differences (Ginsburg et al., 2009, Pfeiffer and Manser, 2010). Hedsköld et al. (2013) pointed out that such differences might weaken the validity of the instrument.".

Safety climate questionnaires, therefore, need to be appropriately validated before being used in different healthcare contexts. The original 12 factor model was replicated in Belgian (Hellings et al., 2007), Portuguese (Eiras et al., 2014) and Scottish data (Sarac et al., 2011). Other studies reported 11 factor models for Dutch (Smits et al., 2008), Arabic (Najjar et al., 2013a), Croatian (Brborovic et al., 2014) and Norwegian data (Haugen et al., 2010); 10 factor models for French (Perneger et al., 2014), Turkish (Bodur and Filiz, 2010), Chinese (Zhu et al., 2014) and Brazilian data (Reis et al., 2016); 9 factor models for UK (Waterson et al., 2010) and Slovene data (Robida, 2013); 8 factor models for Swiss (Pfeiffer and Manser, 2010), Saudi (Alonazi, 2011), Kosovo (Brajshori and Behrens, 2016) and Kuwaiti data. This discrepancy in results could be attributed to differences in employing survey methods and psychometric analytical techniques, in addition to the various modifications made to adapt the original instrument to different healthcare settings (Sarac et al., 2011). Thus, the original HSOPSC will clearly be limited when used in other contexts without proper assessment of its psychometric properties. This point has been highlighted in a number of studies.

3: The have a high response rate >80% but do not discuss that especially the studies from USA have much lower response rate. Could the low response rate have impact on the results and the differences in structure?

I do not think it has an effect. The US response rate was around 64% which is lower than ours but it is not considered low. They had a big sample size (50,513 hospital staff respondents). Psychometric evaluation of the tool (HSOPSC) has been conducted around the world and most of the studies confirm our findings (different structures 11,10,9,8,7 factor structures, check appendix 6) that the tool needs to be evaluated before being adapted to any context as explained in page 18.

Also, relying exclusively on quantitative validation may not be sufficient to produce an instrument that is fully applicable to the local context.

4: It is not mentioned in which language the present questionnaire was presented for the HCWs. English of a local language?

Based on the reviewer's suggestion, I have mentioned that the English version of the HSOPSC tool was used in our study (P.7).

Major concerns:

1- The final version presented is the smallest version of HSOPSC (8 dimensions and 22 items) among all those mentioned in Appendix 5. The large number of items discarded (20 items) raises doubts about the reasons for discarding intermediate versions with more items and dimensions, but the study does not have enough data on these intermediate versions. I therefore ask you to consider these suggestions:

• Page 9, lines 48-55 and page 10 lines 3-9:

The sample size of the CFA used to test the original HSOPSC (12-factor) model was not presented. I imagine you have used the whole sample, as well as the internal consistency analysis, but it is important to include this information.

I have used the whole sample (n=1280) for testing the CFA and reliability (internal consistency) of the original HSOPSC model. I have added this point to clarify the statement (p.9).

• Page 10, lines 43-54. Specify in the text which was the main statistical criterion and cut-off point used to exclude certain item (factorial load of item <0.40?).

I have clarified the following point in page 10 with cut-off points: "…including identifying items with low communalities (<0.3), no or low loading (<0.4), items with cross loadings (>0.30) …"

• Table 3. Expand data with additional rows with EFA results for models with 9 factors, 10 factors, 11 factors and 12 factors. Also include an additional column that presents a rationale, based on statistics or other factors, of not choosing the model.

Table 3 contains the CFA, rather than the EFA, results of the final eight factor model.

In my study, a series of exploratory factor analyses were performed to identify an optimal model that fits the Kuwaiti hospital setting. Investigation of all possible solutions was undertaken including 12-11-10-9-8-7 number of factors. Due to space restrictions and the huge number of tables that examined the above solutions (40+) I have added a summary of the 12-11-10-9-8-7 models with a short rationale for my final choice in Appendix 2.

2- In the discussion (page 15, lines 50-56), consider that a Brazilian study confirmed the 12-dimensional structure of the HSOPSC (Andrade et al., 2017 http://scielo.iec.gov.br/pdf/ess/v26n3/en_2237 -9622-ess-26-03-00455.pdf). This study was published a few months after reference 74 by Reis et al., Used an independent transcultural translation and adaptation process, a different data collection method (electronic questionnaire) and a sample size approximately 3 times higher.

Due to space restrictions, I have added the new Brazilian study as one of the studies that confirmed the original 12-factor structure (p.16).

Minor concerns:

1- Page 4 of 35, lines 23-25: On "Lack of reporting of explicit psychometric data in some comparative studies used in the comparative analysis was another challenge faced in our study", this is not a limitation of this study, but of other studies.

I meant it limited our understanding of comparative studies results in comparison to ours. This point has been deleted.

2- Page 25 of 25, box 1.

You split two boxes with the same title. I suppose the title of the second one is different from the first one (service and hospital?).

The title is the same for both sections. I have corrected this mistake and made the sections merge into one table.

3- Specify in which language the questionnaire is evaluated. Arabic?

Based on the reviewer's suggestion, I have mentioned that the English version of the HSOPSC tool was used in our study (P.7).

VERSION 2 – REVIEW

REVIEWER	Zenewton A. da Silva Gama Federal University of Rio Grande do Norte, Brazil
REVIEW RETURNED	26-Mar-2019

GENERAL COMMENTS	Congratulations to the authors for the excellent work. I am satisfied with the changes and recommend accepting the work.

VERSION 2 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 2

Reviewer Name: Zenewton A. da Silva Gama

Institution and Country: Federal University of Rio Grande do Norte, Brazil

-Please state any competing interests or state 'None declared': None declared.

This has been declared in page 11.