Evaluating the Construct Validity of the Charité Alarm Fatigue Questionnaire using Confirmatory Factor Analysis

Maximilian Markus Wunderlich
Institute for Medical Informatics, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin

Henning Krampe
Department of Anesthesiology and Intensive Care Medicine, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin

Kristina Fuest
Technical University Munich, Klinikum Rechts der Isar

Dominik Leicht
Justus Liebig University

Moritz Benedikt Probst
Vivantes Klinikum im Friedrichshain

Julian Runge
Marien Hospital Herne – Universitätsklinikum der Ruhr-Universität Bochum

Sebastian Schmid
University Hospital Ulm, Ulm University

Claudia Spies
Department of Anesthesiology and Intensive Care Medicine, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin

Björn Weiß
Department of Anesthesiology and Intensive Care Medicine, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin

Felix Balzer
Institute for Medical Informatics, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin

Akira-Sebastian Poncette (akira-sebastian.poncette@charite.de)
Institute for Medical Informatics, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin

Article

Keywords: patient monitoring, intensive care unit, alarm fatigue, alarm management, patient safety, ICU, alarm system, alarm system quality, medical devices, clinical alarms

Posted Date: January 2nd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3744296/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
Abstract

Background. The Charité Alarm Fatigue Questionnaire (CAFQa) is a 9-item Questionnaire that aims to standardize how alarm fatigue (AF) in nurses and physicians is measured. We previously hypothesized that it has two correlated scales, i.e. one on the psychosomatic effects of AF and the other on staff's coping strategies in working with alarms.

Objective. We aimed to validate the hypothesized structure of the CAFQa and thus underpin the instrument's construct validity.

Methods. We conducted two independent studies with nurses and physicians from ICUs in Germany ($N_{\text{Study 1}} = 265, N_{\text{Study 2}} = 1212$). Responses to the questionnaires were analyzed using confirmatory factor analysis with the unweighted least-squares algorithm based on polychoric covariances. Convergent validity was assessed by participants' estimation of their own AF and exposure to false alarms in percent.

Results. In both studies the chi-square test reached statistical significance ($\chi^2(26) = 44.932, p = 0.012$ and $\chi^2(26) = 92.416, p < 0.001$ for Study 1 and 2, respectively). Other fit indices suggested a good model fit (in both studies RMSEA < 0.05, SRMR < 0.08, RNI > 0.95, TLI > 0.95, and CFI > 0.995). Participants' mean scores correlated moderately with self-reported AF ($r_{\text{Study 1}} = 0.45; r_{\text{Study 2}} = 0.53$) and weakly with self-perceived exposure to false alarms ($r_{\text{Study 1}} = 0.3; r_{\text{Study 2}} = 0.33$).

Conclusion. The questionnaire measures the construct of alarm fatigue as proposed in our previous study. Researchers and clinicians can rely on the CAFQa to measure the AF of nurses and physicians.

Introduction

Background

Alarm fatigue is a phenomenon where healthcare workers in ICUs become desensitized to alarms of medical devices [1]. It can make ICU staff feel stressed, and is a substantial risk to patient safety, as it can lead to alarms being missed or acknowledged with delay [2]. When implementing interventions or IT-solutions [3] that try to remedy alarm fatigue, clinicians and clinical alarm researchers need a reliable way to assess whether they were successful. However, they have not yet agreed on a standardized way of measuring alarm fatigue [4, 5], even though it was recognized more than two decades ago [6].

Solely analyzing an ICU's alarm log data cannot serve as a measure of staff's alarm fatigue. While it is a valuable method for designing alarm management interventions [7], there is no clear association between the number of alarms on an ICU and staff's subjective alarm fatigue (examples can be found in Sowan et al. [8], Hüskes-Kraus et al. [4] and Wilken et al. [9]).

Therefore, we recently developed the Charité Alarm Fatigue Questionnaire (CAFQa), which is a 9-item questionnaire that measures alarm fatigue in nurses and physicians [10]. Using exploratory factor analysis, we identified two correlated factors: one revolving around the psychophysiological effects of alarms (e.g., headaches and feelings of distraction), and one revolving around ICU staff's alarm management strategies (e.g., customization of alarm limits). We named the former the “alarm stress scale” and the latter the “alarm coping scale”. The alarm coping scale consists of items that are reversely scored. Hence, a high score on either scale is indicative of alarm fatigue.

When developing a new questionnaire, it is essential to establish construct validity, i.e., whether the questionnaire truly measures what it attempts to measure. One way to test an instrument's construct validity is to administer it to a different sample and test whether the originally proposed factor structure re-emerges, using confirmatory factor analysis [11, 12] (for a recent example see Canivez et al. [13]).

Aim

We aim to validate the exploratively derived factor structure of the Charité Alarm Fatigue Questionnaire and thus underpin the instrument's construct validity.

Methods

Ethics Approval

The ethical approval for this study was granted by the ethics committee of the Charité – Universitätsmedizin Berlin (EA4/218/20) and, if required, confirmed by the local ethics committee at the participating hospital. This study was conducted in compliance with the relevant guidelines and regulations. All participants provided informed consent after receiving a full briefing on the study's purpose, procedures, and potential implications.

Participants

Study 1
We recruited participants from nine ICUs of five large German hospitals. The questionnaire was administered online using REDCap between October 2021 and July 2022. As a reward for completing the questionnaire, we offered participants the chance to enter a drawing where they could win a €50 voucher for online shopping. Participants were asked to consent to have their data collected, analyzed, and stored anonymously.

**Study 2**

Using a mailing list, we invited all members of the German Society of Anaesthesiology and Intensive Care Medicine [14] to fill out the online questionnaire (again using REDCap) between March 2023 and July 2023.

**Questionnaire**

The questionnaire used in both studies was identical and consisted of all nine items from the CAFQa [10] and five general questions about the alarm situation in participants’ ICUs. These general questions were not part of the analysis for this report. All 14 items were pseudo-randomly arranged and required responses on a Likert scale ranging from 1 to 7 (indicating “I do not agree at all”) to 7 (indicating “I very much agree”). Items with negative valences were reverse scored. Demographic items asked participants about their average number of workdays in an intensive care or monitoring area, their number of years/months of ICU experience, their workplace (campus and unit), and their profession. We made small adjustments to the original wording of two items (items 8 and 9, see Table 1 below) to improve readability: In item 8 we used “situation” instead of “urgency”. In item 9 we used the phrase “clinical pictures” instead of “clinical symptoms”.

**Statistical Analysis**

All analyses were conducted in R (Linux version 4.2.1) [15] using the following packages: Tidyverse [16], reshape2 [17], psych [18], semPlot [19], and lavaan [20]. For Study 1, we pooled the data from the participating hospitals. The data is openly available at zenodo.org/record/8296934.

**Missing Data**

In accordance with Heymans and Eekhout [21], we employed predictive mean matching via the mice package [22] to impute missing data that was assumed to be missing at random (MAR). We did not impute questionnaires that were either completely empty or terminated prematurely (presumably due to survey fatigue), as the assumption of MAR was not met in these cases. We assumed that survey fatigue occurred if a participant failed to respond to at least the final 20% of the questionnaire (i.e., the last three or more of the 9 items of CAFQa plus the 5 general questions). In total, 0.3% of the data were missing at random.

**Testing Assumptions of Confirmatory Factor Analysis**

In both studies, the results of Mardia’s test indicated that the multivariate skew did not come from a normal distribution with p < .001. Outliers were identified using Mahalanobis distances, with none being detected in Study 1 and four being detected in Study 2 (for both studies: χ²(9) cutoff = 27.88, p < .001). Visual inspection of the data from all four cases revealed no unusual response patterns. Given the large sample size, we decided not to remove any outliers. The KMO statistic [23] in Study 1 = 0.76, and in Study 2 = 0.8. In both studies, Bartlett’s test of sphericity [24] rejected the null hypothesis that the correlation matrix was an identity matrix (Study 1: χ²(36) = 438.27, p < .001; Study 2: χ²(36) = 2495.44, p < .001). There was no evidence of multicollinearity in either study as the determinant of both R matrices was greater than 0.00001 [25] and no correlations were greater than |0.7|. Overall, these results suggest that the data of both studies is suitable for factor analysis.

**Confirmatory Factor Analysis**

For both studies, we specified the model in line with our previous findings [10] with two correlated latent factors labeled “alarm stress” and “alarm coping”. Items 1–5 were assigned to “alarm stress”. Items 5–9 were assigned to “alarm coping”. Since all CAFQa items are ordered categorical variables (due to being measured on a 5-point Likert scale) and because Mardia’s tests indicated that the multivariate skew of both studies did not come from a normal distribution, we used the unweighted least-squares (ULS) algorithm based on polychoric covariances for estimating factor loadings [26–28]. We assessed the goodness-of-fit of the model using chi-square, and the following fit indices in line with the cutoff criteria defined by Hu and Bentler [29]: root mean square error of approximation (RMSEA), relative non-centrality index (RNI), Tucker-Lewis index (TLI), standardized root mean squared residual (SRMR), and comparative fit index (CFI).

**Convergent Validity**

At the end of the questionnaire in both studies, we provided participants with a brief description of alarm fatigue and asked them to estimate their personal alarm fatigue in percent (0% indicating no alarm fatigue and 100% indicating extreme alarm fatigue). We also asked participants to provide their perceived rate of false alarms in their ICU in percent (0% indicating no false alarms, 100% indicating no true alarm). To measure convergent validity, we correlated the participants’ mean scores on the questionnaire with the self-provided alarm fatigue and false alarm rate estimations (in total, and per factor).

**Internal Consistency**

As a measure of internal consistency, we report Cronbach’s coefficient alpha, McDonald’s coefficient omega [30], and the mean inter-item correlation for both factors.
Results

Participants

Study 1

We received 363 submissions. Of these, 23 did not consent to have their data analyzed, 67 questionnaires were empty, and 8 showed signs of survey fatigue. Therefore, the sample size for this study was N = 265. The number of participants was roughly similar for each hospital (n_G = 43, n_H = 50, n_M = 64, n_U = 57, n_V = 51). Most participants were nurses (56.6%) and 35.8% were physicians. Few participants (3.4%) were supporting nurses, nurses in training, medical students, or interns, while 4.2% did not state their profession.

Study 2

Of the 1564 submissions we received, 69 participants refused to consent to have their data processed and 223 submitted empty questionnaires. We suspected survey fatigue in 60 cases. Hence, the sample size of Study 2 was N = 1212. Contrary to Study 1, more participants were physicians (82.7%) than nurses (15.3%). Again, the group of supporting nurses, nurses in training, medical students, and interns was a minority (0.5%). 1.5% did not state their profession.

Confirmatory Factor Analysis

Descriptive statistics of both studies are presented in Table 1 for each item. Figure 1 provides an overview of the final model of both studies.

Study 1

While the chi-square test was significant at alpha = 0.05 with $\chi^2 (26) = 44.932$, $p = 0.012$, indicating that the model did not fit the data, all fit indices suggested a good model fit: RMSEA = 0.03, SRMR = 0.052, RNI = 0.989, TLI = 0.985 and CFI = 0.989. All items loaded onto their hypothesized factors as expected with factor loadings that were statistically significant at $p < .001$, ranging from 0.35 to 0.73. The factors were moderately correlated with 0.4 ($p < .001$; 95% CI = 0.21–0.59).

Study 2

As in study one, the chi-square test was significant ($\chi^2 (26) = 92.416$, $p < 0.001$), indicating that the model did not fit the data, while the fit indices show a good model fit (RMSEA = 0.046, SRMR = 0.041, RNI = 0.982, TLI = 0.975 and CFI = 0.982). Again, all items loaded onto their hypothesized factors as expected with factor loadings that were statistically significant at $p < .001$, ranging from 0.43 to 0.81. The factors were moderately correlated with 0.44 ($p < .001$; 95% CI = 0.36–0.51).
Table 1
Descriptive statistics for each item and the pattern coefficients found in the confirmatory factor analysis of the two-factor model in both studies. All loadings were statistically significant at p < 0.001. F = factor, CI = confidence interval, SD = standard deviation, Kurt. = kurtosis.

<table>
<thead>
<tr>
<th>Item</th>
<th>Study 1</th>
<th>Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>1 With too many alarms on my ward, my work performance, and motivation decrease.</td>
<td>0.730</td>
<td>0.643–0.818</td>
</tr>
<tr>
<td>2 Too many alarms trigger physical symptoms for me, e.g., nervousness, headaches, and sleep disturbances.</td>
<td>0.706</td>
<td>0.612–0.800</td>
</tr>
<tr>
<td>3 Alarms reduce my concentration and attention.</td>
<td>0.725</td>
<td>0.635–0.814</td>
</tr>
<tr>
<td>4 My or neighboring patients' alarms or crisis alarms frequently interrupt my workflow.</td>
<td>0.432</td>
<td>0.318–0.547</td>
</tr>
<tr>
<td>5 There are situations when alarms confuse me.</td>
<td>0.488</td>
<td>0.384–0.593</td>
</tr>
<tr>
<td>6 In my ward, procedural instruction on how to deal with alarms is regularly updated and shared with all staff.</td>
<td>-</td>
<td>0.434</td>
</tr>
<tr>
<td>7 Responsible personnel respond quickly and appropriately to alarms.</td>
<td>-</td>
<td>0.587</td>
</tr>
<tr>
<td>8 The acoustic and visual monitor alarms used on my ward floor and in my nurses' station allow me to assign the patient, the device, and the situation clearly.</td>
<td>-</td>
<td>0.349</td>
</tr>
</tbody>
</table>
Study 1

<table>
<thead>
<tr>
<th>Item</th>
<th>F1</th>
<th>F2</th>
<th>95% CI</th>
<th>Mean</th>
<th>SD</th>
<th>Kurt.</th>
<th>Skew</th>
<th>F1</th>
<th>F2</th>
<th>95% CI</th>
<th>Mean</th>
<th>SD</th>
<th>Kurt.</th>
<th>Skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>-</td>
<td>0.581</td>
<td>0.428–0.734</td>
<td>-0.38</td>
<td>0.93</td>
<td>-0.24</td>
<td>0.26</td>
<td>-</td>
<td>0.575</td>
<td>0.508–0.641</td>
<td>-0.35</td>
<td>0.94</td>
<td>-0.12</td>
<td>0.40</td>
</tr>
</tbody>
</table>

* Item with a negative valence that is reversely scored.

Convergent Validity

In Study 1, the participants’ mean scores correlated moderately with self-reported alarm fatigue (r(242) = 0.45, p < .001, 95% CI = 0.34–0.54), and weakly with the perceived percentage of false alarms: r(247) = 0.3, (p < .001, 95% CI = 0.18–0.41). Similar patterns were observed in Study 2 (see Table 2 for full details).

Table 2

<table>
<thead>
<tr>
<th>Correlation Measure</th>
<th>DF</th>
<th>r</th>
<th>p-value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS-SRAF</td>
<td>242</td>
<td>0.45</td>
<td>&lt; 0.001</td>
<td>0.34–0.54</td>
</tr>
<tr>
<td>F1S-SRAF</td>
<td>242</td>
<td>0.42</td>
<td>&lt; 0.001</td>
<td>0.31–0.52</td>
</tr>
<tr>
<td>F2S-SRAF</td>
<td>242</td>
<td>0.29</td>
<td>&lt; 0.001</td>
<td>0.17–0.4</td>
</tr>
<tr>
<td>MS-PPFA</td>
<td>247</td>
<td>0.30</td>
<td>&lt; 0.001</td>
<td>0.18–0.41</td>
</tr>
<tr>
<td>F1S-PPFA</td>
<td>247</td>
<td>0.20</td>
<td>0.0016</td>
<td>0.08–0.32</td>
</tr>
<tr>
<td>F2S-PPFA</td>
<td>247</td>
<td>0.29</td>
<td>&lt; 0.001</td>
<td>0.17–0.4</td>
</tr>
<tr>
<td>Study 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS-SRAF</td>
<td>1180</td>
<td>0.53</td>
<td>&lt; .001</td>
<td>0.49–0.57</td>
</tr>
<tr>
<td>F1S-SRAF</td>
<td>1180</td>
<td>0.49</td>
<td>&lt; .001</td>
<td>0.45–0.53</td>
</tr>
<tr>
<td>F2S-SRAF</td>
<td>1180</td>
<td>0.34</td>
<td>&lt; .001</td>
<td>0.29–0.39</td>
</tr>
<tr>
<td>MS-PPFA</td>
<td>1182</td>
<td>0.33</td>
<td>&lt; .001</td>
<td>0.28–0.38</td>
</tr>
<tr>
<td>F1S-PPFA</td>
<td>1182</td>
<td>0.26</td>
<td>&lt; .001</td>
<td>0.21–0.32</td>
</tr>
<tr>
<td>F2S-PPFA</td>
<td>1182</td>
<td>0.28</td>
<td>&lt; .001</td>
<td>0.23–0.33</td>
</tr>
</tbody>
</table>

Internal Consistency

In Study 1, Cronbach’s Alpha of factor 1 was 0.72 and 0.49 for factor 2. Cronbach’s alpha across factors was 0.67. The mean inter-item correlation on factor 1 was 0.38 and 0.23 on factor 2. McDonald’s coefficient omega for factor 1 was 0.77 and for factor 2 0.55. The overall coefficient omega for the assessment was 0.8.

Results were similar in Study 2: Cronbach’s Alpha of factor 1 was 0.77 and 0.55 for factor 2. Cronbach’s alpha across factors was 0.72. The mean inter-item correlation on factor 1 was 0.44 and 0.27 on factor 2. McDonald’s coefficient omega for factor 1 was 0.8 and for factor 2 0.59. The overall
The coefficient omega for the assessment was 0.85.

**Discussion**

We aimed to underpin the construct validity of the Charité Alarm Fatigue Questionnaire by submitting the exploratively derived factor structure from our previous study to confirmatory factor analysis in two independent studies. While the chi-square test rejected the model in both studies, all fit indices indicated a good model fit. The factor loadings ranged from 0.35 to 0.73 in Study 1 and from 0.43 to 0.81 in Study 2 and were all statistically significant. Overall, these results support the hypothesized factor structure. The questionnaire seems to measure the construct of alarm fatigue as proposed in our previous work [10].

The chi-squared test is known to be sensitive to large sample sizes [31], which might explain its statistical significance. We did not modify the model because all fit indices indicated a good fit and because model modifications, no matter how small or plausible, can make a model less generalizable.

In both studies, the first factor, i.e. the alarm stress scale, and the overall questionnaire were internally consistent. However, the second factor, i.e. the alarm coping scale, seems to have issues with its internal consistency. Here, Cronbach's Alpha and McDonald's Omega were 0.49 and 0.55, in Study 1, respectively, and 0.55 and 0.59, in Study 2, respectively. A similar pattern can be found in our previous study, where Cronbach's Alpha of factor 2 = 0.57 [10]. An internally consistent questionnaire is desirable. However, it can also mean that items are very similar. It was our ambition to create a questionnaire that is brief while measuring the many facets of alarm fatigue. Future studies using the CAFQa should routinely assess the internal consistency of both factors. If the second factor continues to show medium internal consistency, research should be done on how it can be improved (e.g., by means of adding additional items).

Participants who had a high mean score on the questionnaire also rated themselves as more alarm fatigued ($r = 0.45$ in Study 1 and $r = 0.53$ in Study 2). This positive correlation indicates the convergent validity of the questionnaire. Similarly, both studies demonstrated that participants with a high mean score on the questionnaire perceived more alarms to be false in their ICU. In Study 1 this association was stronger for factor 2 than for factor 1. This makes sense since a high score on factor 2 (i.e., the alarm coping scale) indicates that alarms are not properly managed (e.g., by means of patient-specific threshold customizations), which typically leads to more false alarms [32]. However, Study 2 could not replicate this pattern. Future research should find an answer to the question: Does ICU staff with a high perceived percentage of false alarms tend to develop stronger alarm fatigue, or does staff that is more alarm fatigued tend to perceive more alarms as being false?

**Limitations**

The fit indices RMSEA, CFI, and TLI have been shown to overestimate model fit when using the ULS estimator [33, 34], potentially leading researchers to accept a bad-fitting model. Yet, in our case, other fit indices indicate a good model fit. As in our previous work [10], the assumption that participants can accurately reflect and express their own alarm fatigue in percent is likely flawed (otherwise it would not be necessary to develop a questionnaire in the first place). However, we believe that it is a valuable method for assessing convergent validity when no other instrument is available. Most ICU nurses and physicians have heard of alarm fatigue, and we provided them with a brief recapitulation on the concept before having them answer the self-report-item in each study.

**Conclusion**

Our results from two independent studies underpin the construct validity of the Charité Alarm Fatigue Questionnaire. All items consistently loaded onto the factors as we proposed in a previous publication [10]. When conducting research or quality improvement projects in ICUs, clinical alarm researchers and clinicians can rely on this instrument to measure, compare and benchmark the alarm fatigue of nurses and physicians.

**Abbreviations**

CAFQa - Charité Alarm Fatigue Questionnaire
CI - Confidence Interval
CFA - Confirmatory Factor Analysis
CFI - Comparative Fit Index
DF - Degrees of Freedom
EFA - Exploratory Factor Analysis
e.g. - For example
Declarations

Acknowledgements

ASP is a participant in the Charité Digital Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin and the Berlin Institute of Health. We are grateful to Nicolas Frey for supervising the logistical aspects of the survey distribution, to Sandro Amende-Wolf for his valuable clinical insights, and to Dr. Daniel Schulze for his critical examination of our statistical analysis. We are grateful to the German Society of Anaesthesiology and Intensive Care Medicine for the distribution of the questionnaire to its members. This article was extracted from the doctoral thesis of MMW.

Members of the Study Group

In alphabetical order: Dr. med. Mirza Aghamov 7, Prof. Dr. med. Manfred Blobner 3, Prof. Dr. med. Ulrich Frey 6, Prof. Dr. Christian von Heymann 5, Jacqueline Holan 5, Prof. Dr. med. Bettina Jungwirth 7, Dr. med. Dragutin Popovic 5, Prof. Dr. med. Michael Sander 4.

Author Contributions

MMW: Conceptualization, Investigation, Methodology, Validation, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration. HK: Methodology, Validation, Writing - Review & Editing. KF: Investigation, Resources, Writing - Review & Editing. DL: Investigation, Resources, Writing - Review & Editing. MBP: Investigation, Resources, Writing - Review & Editing. JR: Investigation, Resources, Writing - Review & Editing. SS: Investigation, Resources, Writing - Review & Editing. CS: Conceptualization, Resources, Writing - Review & Editing. BW: Validation, Writing - Review & Editing. FB: Conceptualization, Supervision, Writing - Review & Editing. ASP: Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision, Project administration.

Data Availability Statement

The datasets generated and analyzed during the current study are available in the Zenodo repository: https://zenodo.org/record/8296934.

Conflict of Interest

The authors received no financial funding for this article’s research and publication. There are no competing interests for any author.
References

8. Sowan AK, Gomez TM, Tarriela AF, Reed CC, Paper BM. Changes in Default Alarm Settings and Standard In-Service are Insufficient to Improve Alarm Fatigue in an Intensive Care Unit: A Pilot Project. JMIR Hum Factors 2016;3(1). PMID:27036170


Figures

![Study 1, N = 265](image1)

![Study 2, N = 1212](image2)
Diagram of the final model from each study. Factors are shown in circles (factor 1: alarm stress, factor 2: alarm coping), and items 1-9 in squares. In both studies, the variance of the factors was constrained to 1. The arrows connecting the factors with the items are the factor loadings and arrows pointing toward the items show the residuals. The arrow between the two factors shows their correlation.