The Fear of Pain Questionnaire-III and the Fear of Pain Questionnaire-Short Form: a confirmatory factor analysis

Vambheim, Sara M.; Lyby, Peter Solvoll; Aslaksen, Per M.; Flaten, Magne Arve; Asli, Ole; Martinussen, Laila Marianne

Published in:
Journal of Pain Research

Link to article, DOI:
10.2147/JPR.S133032

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
The Fear of Pain Questionnaire-III and the Fear of Pain Questionnaire-Short Form: a confirmatory factor analysis

Sara M Vambheim¹
Peter Solvoll Lyby¹,²
Per M Aslaksen¹
Magne Arve Flaten³
Ole Åsli¹
Laila M Martinussen⁴

¹Department of Psychology, UiT, The Arctic University of Norway, Tromsø;
²CatoSenteret Rehabilitation Center, Son;
³Department of Psychology, NTNU, The Norwegian University of Science and Technology, Trondheim, Norway;
⁴Management Engineering, DTU, Technical University of Denmark, Lyngby, Denmark

Introduction

Fear of pain (FOP) is associated with emotional reactions occurring in the anticipation of pain or during pain.¹ FOP is related to increased pain perception²–⁴ and reduced capacity for pain inhibition,¹ and is a central mechanism in several pain disorders.⁵,⁶ Females have higher FOP⁷–⁹ and are more sensitive to pain and less sensitive to pain inhibition than males.¹,⁵,10,11 Furthermore, the prevalence of several chronic pain conditions is higher in females than in males.¹ Thus, the association between FOP and pain and sex differences in FOP and pain are well established.

FOP is related to avoidant pain behaviors, for example, avoidance of situations that involve or may involve pain experience. It has been argued that FOP is more disabling than the actual pain.¹² Testing the applicability of FOP questionnaire is, therefore, of
great importance. In order to accurately capture FOP, the FOP questionnaires should mirror FOP across sex and age. Furthermore, the validation of FOP questionnaires is crucial for the development of personalized medicine and treatments.

The Fear of Pain Questionnaire-III (FPQ-III) was developed by McNeil and Rainwater in 1998. Since then, it has been translated into several different languages, and has become widely used in both clinical and nonclinical samples. The FPQ-III consists of 30 items and three subscales: fear of Severe Pain, Minor Pain and Medical Pain. Asmundson et al examined the factor structure, consistency and validity of the FPQ-III in a Canadian sample of 589 undergraduates, and concluded that the model had unacceptable fit. Similar findings have been made by others. The Fear of Pain Questionnaire-Short Form (FPQ-SF) was, therefore, suggested as an alternative and better adjusted questionnaire for measurements of FOP. This questionnaire is reduced to 20 items and is divided into four subscales: fear of Severe Pain, Minor Pain, Injection Pain and Dental Pain.

It has been argued that the FPQ-SF enables factor structure invariance across sex, and thus is better suited for measuring FOP in both males and females. The psychometric properties of the FPQ-III have been investigated previously, and sex differences were reported in FPQ-III Total and the three subscales: Severe Pain, Minor Pain and Medical Pain. However, van Wijk and Hoogstraten developed a German version of the FPQ-III and found no sex differences in their sample. Thus, translational, linguistic and cultural issues may contribute to increase or decrease sex differences in endorsed FOP. Therefore, the main aim of this study was to investigate the FPQ-III’s and FPQ-SF’s model fit, reliability and validity in a Norwegian sample, in order to explore the questionnaires’ applicability. The secondary aim was to investigate the models’ sex neutrality, that is, how well the questionnaires measure FOP across sex groups. We hypothesized that the FPQ-SF would have a better model fit than the FPQ-III. We also expected sex differences in the model fit of the FPQ-III, but not the FPQ-SF. Higher fit was expected for women than for men in the FPQ-III model. Furthermore, higher FOP scores were expected in females than in males.

Methods
Participants
A total of 807 healthy participants were included in the study: 339 males, age range 18–40 years, mean 23.4, standard deviation (SD)=4.1, and 468 females, age range 18–40 years, mean =22.1, SD=3.4. The participants were undergraduate students recruited on campus at the University of Tromsø and screened for medical history of serious diseases or injuries. Somatic or psychiatric disorders and use of prescription-based medications or allergy medications led to exclusion from participation. Pregnant women were excluded. Participants were instructed to abstain from use of nicotine- and caffeine-containing substances 3 hours before participation. The participants had to speak Norwegian as Norwegian language was used in the questionnaires, instructions, consent and measurements of pain, stress and activation collected in the experiments. Data from seven different study samples were pooled. All participants filled in the FPQ-III and an informed consent form. The studies were approved by the Regional Committee for Medical Research Ethics North Norway (project numbers: 2013/966, 2012/1888, 2610.000001, 49/2005, 5.2006.2452; 20277, 17/2006).

Measures
The FPQ-III measures pain-related fear on a 30-item questionnaire. Each question presents a pain-related situation, and participants are asked to rate FOP on a 5-point Likert scale (1= not afraid at all, 5= extremely afraid). The FPQ-III is further divided into three factorial distinct subscales consisting of 10 questions: Severe Pain (eg, being involved in a car accident), Minor Pain (eg, biting your tongue) and Medical Pain (eg, receiving an injection in your mouth). McNeil and Rainwater reported that the psychometric properties of the FPQ-III were considered satisfactory, with good internal consistency and test–retest reliability. Similar findings have been reported by others. A Norwegian version of the FPQ-III, translated by Lyby et al, was used in this study.

The FPQ-SF is a reduced version of the FPQ-III. The FPQ-SF was developed by Asmundson et al, who reported good internal consistency and construct validity. This reduced version of the FPQ-III consists of 20 items and has four factorial distinct subscales: Severe, Minor, Injection (having a blood sample drawn) and Dental Pain (having a tooth pulled). The questionnaire is organized in a similar format as the FPQ-III, with a 5-point Likert scale. To date, there is limited knowledge about the FPQ-SF’s reliability and validity.

Procedure
The participants were recruited from the University of Tromsø. All subjects were participants in pain studies and filled in the FPQ-III and a written consent form prior to testing. Data obtained from the pain experiments are published elsewhere.

Statistical analyses
Firstly, SPSS version 24 was used to run independent samples t-tests to examine whether there were sex differences on
subscale level. $p$ values from the analysis of variance were corrected for multiple comparisons by the Holm–Bonferroni procedure. Secondly, IBM SPSS AMOS version 22 (IBM Corporation, Armonk, NY, USA), maximum likelihood estimation was used to conduct confirmatory factor analysis (CFA) in order to examine the fit of the two models to the data and the possible sex differences in model fit. The fit of the models to the data was evaluated by the $\chi^2$/degrees of freedom ratio, the root mean square error of approximation (RMSEA), the goodness-of-fit index (GFI), the expected cross-validation index (ECVI) and the comparative fit index (CFI). Traditionally, a good fit model should have 2:1 or 5:1 $\chi^2$/degrees of freedom ratio, RMSEA <0.08 or 0.10 (preferably <0.05), GFI >0.90, lower values of ECVI indicating a closer fit and CFI >0.90 (preferably >0.95). Two multigroup CFAs (chi-square difference test) were used to examine if the measurement models differed across the two sex groups. Three (FPQ-III) and four (FPQ-SF) factors respectively were held constant simultaneously across the sex groups (i.e., configural invariance) and the factor loadings were constrained to be equal across the sex groups (i.e., a test if there are equal factor loadings or metric invariance) in the multigroup analysis. Lastly, IBM SPSS version 24 (IBM Corporation) was used to calculate Cronbach’s $\alpha$ values for the factors in the two models as well as to determine the intercorrelations between the sum scores of factors in the FPQ-III and the FPQ-SF. Acceptable high Cronbach’s $\alpha$ is 0.7 and above.

Results

Descriptive statistics

Descriptive statistics for FPQ-III and FPQ-SF total and subscale mean scores and standard deviations for males and females are presented in Table 1. Independent samples t-tests revealed that females scored significantly higher than males on all subscales of the FPQ-III: Severe Pain ($t$ (1.805) = −8.60, $p < 0.001$), Minor Pain ($t$ (1.805) = 3.49, $p=0.001$) and Medical Pain ($t$ (1.805) = −4.25, $p<0.001$) and the Holm–Bonferroni adjusted $\alpha$ level=0.01. Furthermore, females scored significantly higher than males on all subscales of the FPQ-SF: Severe Pain ($t$ (1.805) = −7.99, $p<0.001$), Minor Pain ($t$ (1.805) = −3.93, $p=0.001$), Injection Pain ($t$ (1.805) = 4.00, $p<0.001$) and Dental Pain ($t$ (1.805) = −3.15, $p=0.002$).

Fit indices

Fit indices for the two tested models are presented in Table 2. In line with the main hypothesis, the FPQ-SF had the best overall fit. The second hypothesis was not fully confirmed, as the results revealed that model fit differed across sex for both the FPQ-III and the FPQ-SF. The FPQ-III had a slightly better fit for males, and the FPQ-SF had a slightly better fit for females. Further analysis (multigroup CFA) showed that the differences in model fit across the sex groups were significant in the FPQ-III model ($p<0.000$) and nonsignificant in the FPQ-III model ($p=0.054$) as hypothesized. Additionally, the FPQ-SF had a better overall fit, as indicated by the fit indices presented in Table 2. However, neither of the models showed very good fit, although the fit was adequate according to some of the predetermined criteria (RMSEA and ECVI). The factorial structures with standardized parameter estimates (factor loadings and squared multiple correlations) for the FPQ-III and FPQ-SF in the whole sample are presented schematically in Figures 1 and 2. Standardized parameter estimates (factor loadings and squared multiple correlations) for men and women separately
in the two models are presented in Table 3. The factor loadings presented in Figures 1 and 2 and Table 3 show the items correlation with the latent factor. As can be seen in Figures 1 and 2 and Table 3, for the whole sample, none of the items loaded under the recommended cut-off of 0.3; however, some loaded below 0.5. In the female subgroup, items 18 and 27 in the Severe FPQ-III subscale loaded below 0.3, and in male and female subgroups several items loaded below 0.5 in both models. None of the items in FPQ-III had loadings above 0.7, but in FPQ-SF several items loaded above 0.77.

Reliability and intercorrelations
The Cronbach’s $\alpha$ values showed high internal consistency for both the FPQ-III and the FPQ-SF (Tables 4 and 5). Acceptable high Cronbach’s $\alpha$ is 0.7 and above. However, the values are affected by the number of items in each factor. Even though the FPQ-SF has fewer items than the FPQ-III, it still had acceptable $\alpha$ values. All intercorrelations between the factors were significant (Table 5). Positive and significant correlations were found between all the subscales.

Discussion
This study investigated the model fit, reliability, validity and sex neutrality of the FPQ-III and the FPQ-SF factor models. CFA revealed that none of the models had good fit, but the FPQ-SF provided a better overall fit than the FPQ-III. The ECVI, CFI and GFI favored the FPQ-SF, even though both CFI and GFI were below the recommended threshold. RMSEA was within the recommended threshold for both models; however, it was not within a good model fit range. The poor fit of the models is also mirrored in the standardized parameter estimates (factor loadings and squared multiple correlations), in which some items were below the recommended factor loading threshold. There were sex differences in the model fit which were significant for the FPQ-III model, but not for the FPQ-SF, as hypothesized.
Figure 2 The factor loadings and squared multiple correlations of the FPQ-SF.

Abbreviations: FPQ-SF, Fear of Pain Questionnaire Short Form; FPQ-SF MD, FPQ-SF Dental Pain; FPQ-SF I, FPQ-SF Injection Pain; FPQ-SF M, FPQ-SF Minor Pain; FPQ-SF S, FPQ-SF Severe Pain.

Table 3 The factor loadings and squared multiple correlations of the FPQ-III and the FPQ-SF among males and females separately

<table>
<thead>
<tr>
<th>FPQ-III</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>In</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>12</td>
<td>12</td>
<td>19</td>
<td>19</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>Mi fl</td>
<td>0.54</td>
<td>0.61</td>
<td>0.62</td>
<td>0.64</td>
<td>0.53</td>
<td>0.61</td>
<td>0.51</td>
<td>0.55</td>
<td>0.70</td>
<td>0.72</td>
<td>0.52</td>
<td>0.53</td>
<td>0.53</td>
<td>0.55</td>
</tr>
<tr>
<td>r</td>
<td>0.29</td>
<td>0.38</td>
<td>0.38</td>
<td>0.41</td>
<td>0.28</td>
<td>0.37</td>
<td>0.26</td>
<td>0.30</td>
<td>0.49</td>
<td>0.52</td>
<td>0.27</td>
<td>0.28</td>
<td>0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>S fl</td>
<td>0.48</td>
<td>0.59</td>
<td>0.89</td>
<td>0.75</td>
<td>0.54</td>
<td>0.66</td>
<td>0.91</td>
<td>0.77</td>
<td>0.50</td>
<td>0.69</td>
<td>0.48</td>
<td>0.67</td>
<td>0.50</td>
<td>0.56</td>
</tr>
<tr>
<td>r</td>
<td>0.23</td>
<td>0.35</td>
<td>0.80</td>
<td>0.57</td>
<td>0.29</td>
<td>0.43</td>
<td>0.82</td>
<td>0.60</td>
<td>0.25</td>
<td>0.48</td>
<td>0.23</td>
<td>0.45</td>
<td>0.25</td>
<td>0.31</td>
</tr>
<tr>
<td>MP fl</td>
<td>0.71</td>
<td>0.65</td>
<td>0.82</td>
<td>0.72</td>
<td>0.76</td>
<td>0.71</td>
<td>0.41</td>
<td>0.50</td>
<td>0.32</td>
<td>0.33</td>
<td>0.72</td>
<td>0.75</td>
<td>0.52</td>
<td>0.59</td>
</tr>
<tr>
<td>r</td>
<td>0.50</td>
<td>42</td>
<td>0.67</td>
<td>0.51</td>
<td>0.57</td>
<td>0.51</td>
<td>0.17</td>
<td>0.25</td>
<td>0.10</td>
<td>0.11</td>
<td>0.52</td>
<td>0.57</td>
<td>0.27</td>
<td>0.35</td>
</tr>
<tr>
<td>Covariates F</td>
<td>Mi-S 0.33</td>
<td>Mi-MP 0.47</td>
<td>S-MP 0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariates M</td>
<td>Mi-S 0.47</td>
<td>Mi-MP 0.62</td>
<td>S-MP 0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FPQ-SF</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
<th>F</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>In</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>19</td>
<td>19</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>23</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Mi fl</td>
<td>0.55</td>
<td>0.67</td>
<td>0.67</td>
<td>0.72</td>
<td>0.51</td>
<td>0.58</td>
<td>0.75</td>
<td>0.74</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.48</td>
</tr>
<tr>
<td>r</td>
<td>0.30</td>
<td>0.45</td>
<td>0.45</td>
<td>0.51</td>
<td>0.26</td>
<td>0.34</td>
<td>0.56</td>
<td>0.55</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.29</td>
<td>0.29</td>
<td>0.14</td>
</tr>
<tr>
<td>S fl</td>
<td>0.44</td>
<td>0.52</td>
<td>0.92</td>
<td>0.86</td>
<td>0.50</td>
<td>0.57</td>
<td>0.93</td>
<td>0.87</td>
<td>0.46</td>
<td>0.63</td>
<td>0.44</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.19</td>
<td>0.27</td>
<td>0.85</td>
<td>0.74</td>
<td>0.25</td>
<td>0.32</td>
<td>0.86</td>
<td>0.75</td>
<td>0.22</td>
<td>0.40</td>
<td>0.19</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariates F</td>
<td>I-S 0.10</td>
<td>Mi-S 0.29</td>
<td>Mi-D 0.42</td>
<td>D-I 0.65</td>
<td>Mi-I 0.26</td>
<td>D-S 0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: D, dental; F, female; fl, factor loading; FPQ-III, Fear of Pain Questionnaire-III; FPQ-SF, Fear of Pain Questionnaire Short Form; I, injection; In, item number; M, men; Mi, minor; MP, medical pain; r, squared multiple correlations; S, severe.
The internal consistency of the FPQ-III subscales was good (ranging from 0.82 to 0.85). The internal consistency of the FPQ-SF subscales was good for the Severe and Injection Pain subscales (0.84) and acceptable for the Minor and Medical Pain subscales (0.70 and 0.78). Thus, the findings from the FPQ-III are in line with previous studies. However, the finding of good internal consistency on all FPQ-SF subscales by Asmundson et al was not replicated in this study. Few studies have examined the FPQ-SF. In this study, the FPQ-III was administered to all participants. Future investigations should administer and examine the FPQ-SF, as other results may emerge when subjects respond to this shorter FOP questionnaire.

This study used a Norwegian version of the FPQ-III. Therefore, translation or linguistic issues may have influenced the results. The Norwegian version of the FPQ-III was developed by Lyby et al by translation and back-translation. It has been argued that back-translation is highly important in the translation process as it enables identification of translation or linguistic errors or imprecisions. However, retranslation was not conducted in the translation process with the Norwegian version of the FPQ-III. Future studies could use translation, back-translation and retranslation to ensure an optimal final translation.

A probable explanation of the discrepancies between our findings and findings made in previous studies is that cultural issues influence FOP scoring. Social, cultural and psychologic factors mediate pain behavior, and some cultures value expression of emotional and physiologic distress more than other cultures. Cultural variations have been found in the expression of pain and in the effectiveness of analgesic medications. Additionally, it has been demonstrated that there are cross-cultural sex differences in attitudes toward the expression of pain. One study reported that Indian research participants disapproved pain expression by males less appropriate than pain expression by females. Hobara included a sample of Japanese and Euro-American participants, and found that Japanese participants were less accepting of pain expression than Euro-American participants. Furthermore, participants from the Japanese culture disapproved male pain expression to a larger extent than participants from the Euro-American culture. Such cultural differences may also apply to sex differences in FOP observed in this and other studies, and may be related to differing gender role stereotypes or gender role expectations between cultures. The significant impact of gender role stereotypes is supported by several studies demonstrating that males report lower pain to female experimenters than to male experimenters. Thus, cultural factors combined with psychosocial factors may increase or decrease pain expression, but the relative contribution of cultural factors remains unknown. Previous studies employing the German and English versions of the FPQ-III have reported considerably lower mean subscale scores compared to the present study. The mean subscale scores from our sample correspond to, for example, Roelofs et al, who used a Dutch version of the FPQ-III. Further investigations that include samples with different languages and/or samples from different cultures could help to clarify issues of linguistic and cultural differences. In summary, none of the FOP models seemed to capture FOP well, thus questioning the questionnaires’ applicability in Norway. In other words, these models do not capture the structural constructs of FOP well in Norway.

In line with our hypothesis, there were differences in model fit for males and females. The fit indices showed that FPQ-III had a slightly better fit for males than females, whereas the FPQ-SF had a slightly better fit for females than males. The multigroup CFA analysis confirmed the differences in the FPQ-III model, but not in the FPQ-III model. In the FPQ-SF, 10 items were removed from the original FPQ-III model. Elimination of some of the items where sex differences occur may contribute to sex neutrality. However, the analysis still revealed sex differences on the FPQ-SF subscales, albeit nonsignificant for the whole model. A recent study found that sex differences on the FPQ-III were mainly due to differences in male and female responses to items

### Table 4 Subscale intercorrelations (Pearson correlation) and internal consistency (Cronbach’s α) of the FPQ-III subscales (N=807)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Severe Pain</th>
<th>Minor Pain</th>
<th>Medical Pain</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe Pain</td>
<td></td>
<td></td>
<td></td>
<td>0.854</td>
</tr>
<tr>
<td>Minor Pain</td>
<td>0.387**</td>
<td>0.821</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Pain</td>
<td>0.359**</td>
<td>0.537**</td>
<td>0.828</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Acceptable Cronbach’s α <0.7. **p<0.001.
Abbreviation: FPQ-III, Fear of Pain Questionnaire-III.

### Table 5 Subscale intercorrelations (Pearson correlation) and internal consistency (Cronbach’s α) of the FPQ-SF subscales (N=807)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Severe Pain</th>
<th>Minor Pain</th>
<th>Injection Pain</th>
<th>Dental Pain</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.841</td>
</tr>
<tr>
<td>Minor Pain</td>
<td>0.352**</td>
<td>0.789</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injection Pain</td>
<td>0.133**</td>
<td>0.284***</td>
<td>0.843</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dental Pain</td>
<td>0.251**</td>
<td>0.370**</td>
<td>0.505**</td>
<td>0.706</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Acceptable Cronbach’s α <0.7. **p<0.001.
Abbreviation: FPQ-SF, Fear of Pain Questionnaire-Short Form.
comprising the Severe Pain subscale. In that study, it was concluded that sex differences observed on the FPQ-III may have been due to different emotional reactions to or different interpretations of Severe Pain items in males and females. In females, such items may have elicited, for example, anxiety, while in males the items may have elicited, for example, fear. Anxiety may represent fear or worry about incidents that might occur in the future, while fear may be understood as an immediate and temporary reaction. The findings from this study support this, as the largest sex differences were observed on the Severe subscale.

Asmundson et al argued that the FPQ-SF was better adjusted for sex, compared to the FPQ-III. Our results support this, as the FPQ-SF has better overall fit as well as better fit among both men and women. The multigroup analysis provides further support that the FPQ-SF is better adjusted for sex than the FPQ-III.

The sex differences in mean subscale FOP scores in our sample, displayed by the independent samples t-tests, revealed higher FOP in females than in males. Similar findings have been made previously. In a study by McNeil and Rainwater, sex differences were found on overall, Severe, Medical and Minor Pain subscales. Osman et al reported that females scored significantly higher on FOP than males on two of three subscales. Lyby et al found significantly higher overall FOP scores in females than males. Thus, there are some inconsistencies in the existing literature regarding where the sex differences in FOP appear, with some studies reporting sex differences on overall FOP and others at one or several of the subscales. However, the direction of the sex difference in FOP is consistent, with higher FOP in females than in males.

Sex differences in FOP scores measured by the FPQ-III have been observed in several different translated versions, including the Norwegian, English, Dutch and French versions of the FPQ-III. However, absence of sex differences has also been observed, for example, in the responses to the German FPQ-III version by Van Wijk et al. The larger sex differences in FOP in some languages compared to others may thus be explained by translational or linguistic issues.

Further investigations could examine possible contributing factors for the observed sex differences in FOP by evaluating psychologic, physiologic, linguistic and/or cultural mechanisms. Refining the FPQ-III model may help on the proposed lack of sex neutrality, and thus develop it as a more reliable tool for measurement of FOP in both males and females.

**Limitations**

The present study has several limitations. First, all participants responded to the FPQ-III. Therefore, it is possible that other results would emerge if the FPQ-SF had been distributed. Future investigations should, therefore, include samples where the FPQ-SF, and not only the FPQ-III, is applied. Second, only nonclinical samples were included in this study, so the findings may not be generalizable to clinical samples. Third, the sample consisted of undergraduate students at the University of Tromso, thus contributing to sample homogeneity of age and education. The discrepancy in findings in this and other studies may, therefore, be due to inclusion of participants of different age or educational statuses. Fourth, a Norwegian translation of the FPQ-III was used. Thus, the findings may not be directly transferable to FPQ measures in other languages and/or cultures.

**Conclusion**

The present study shows that none of the FOP models capture FOP well in this Norwegian nonclinical sample. However, the FPQ-SF is a better model than the FPQ-III. Additionally, sex differences were found in the two models fit. The FPQ-SF is a slightly better measurement theory of FOP for females, whereas the FPQ-III is slightly better for males. When testing if the models significantly differed in model fit across sex, the FPQ-SF was not significantly different, but the FPQ-III was. As none of the models proves good fit, adjustment of the FPQ to increase usefulness and ensure accurate FOP measures in Norwegian samples is highly recommended.

**Acknowledgment**

The present study was supported by the Bial Foundation.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**