Male guilt – and shame-proneness: The Personal Feelings Questionnaire (PFQ-2 Brief)

Simon M. Rice, David Kealy, Matt S. Treeby, Olivier Ferlatte, John L. Oliffe, John S. Ogrodniczuk

Summary
Extant research points to shame and guilt as salient affective experiences for men’s mental health outcomes. As the constructs of shame and guilt gain increasing research attention in relation to at-risk men, including those with recent military combat experience, history of sexual abuse, substance misuse and suicidality, valid and reliable assessment is needed. The present psychometric validation study had four arms (Study 1a, 1b, 1c and Study 2) and aimed to validate a short-form of the Personal Feelings Questionnaire (PFQ-2) for assessing guilt and shame. Data were collected from four independent samples of men across community and clinical populations (total N=1042). In Study 1a (n=333) the factor structure of the original 16-item PFQ-2 was rejected. In Study 1b (n=332) a 7-item PFQ-2 Brief was calibrated. This was validated using confirmatory factor analysis in Study 1c (n=335; CFI=0.986, TLI=0.978, RMSEA=0.060, SRMR=0.026). Finally, PFQ-2 Brief properties were evaluated in 42 men attending outpatient psychiatric care (Study 2). The PFQ-2 Brief appears to provide a valid and reliable measure for assessing guilt – and shame-proneness in men and as such, should aid further investigations of the manner in which the two affect styles impact help-seeking, treatment engagement, treatment outcomes, and men’s overall mental health.

Research concerning various aspects of men’s mental health continues to expand in recognition of the unique emotional experiences of men. An important realm of emotional experience that has yet to be fully explored with regard to men’s mental health concerns are the self-conscious emotions of guilt and shame [1]. When experienced at great intensity, chronicity or marked imbalance, guilt and shame can have a deleterious effect on men’s subjective sense of emotional well-being [2]. For some men, guilt and shame may become intertwined with socialized masculine ideals regarding attitudes and behaviors [3,4]. Moreover, these affects can exert a powerful influence on men’s behavioral functioning, including coping strategies, interpersonal relationships and help-seeking behavior. Further efforts to understand the ways in which men experience and manage guilt and shame are thus warranted in order to develop a more sophisti-
cation repertoire of responses to men’s mental health concerns.

While shame and guilt are considered to be self-conscious moral emotions and tend to be experienced in the contexts of interpersonal transgression or failure [5,6], a key difference between them lies in the perceived role of the self in the problematic behavior [7]. With experiences of shame, the focus of the individual’s negative evaluation is squarely on the self, whereas in experiences of guilt, the focus is on the problematic behavior and the ways in which the individual ought to make amends for their perceived failure or transgression [8,9]. A moderate level of transient guilt following a transgression can thus promote empathic perspective-taking and motivate the repair of an interpersonal rupture [10].

While niggling pangs of guilt are often helpful in that they can promote adaptive behavioral change [8], experiences of guilt can be considered problematic in situations when an individual accepts and exaggerates responsibility for negative events that are likely to have been beyond their control [5]. For example, ‘survivor guilt’, or guilt experienced as a consequence of surviving trauma, appears to be maladaptive and is associated with psychological maladjustment [11]. It may also be problematic to experience guilt in situations where personal responsibility for negative events or transgressions is clearly ambiguous [12]. The hypothesized relationship between guilt and psychopathology is featured in diagnostic nosology, with the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, p. 125) [13] listing ‘excessive or inappropriate feelings of guilt’ as part of the diagnostic criteria for a major depressive episode.

A large body of research has established that shame is consistently associated with a range of psychological disorders and symptomatology [5]. Chronic experiences of shame may denote poor self-esteem and feelings of abject inadequacy, inferiority and deficiency [14]. For some men, chronic feelings of shame may be related to a perceived failure to live up to socialized masculine ideals [15]. Shame can fuel social disconnection as a means of avoiding anticipated negative evaluation, or conversely, aggressive behavior to deflect perceptions of weakness [16].

Extant research points to guilt and shame as salient affective experiences for men’s mental health outcomes. Overall, guilt and shame have both been found to be associated with depressive symptoms [17]. Shame-proneness has been linked with substance abuse [18,19], hypersexual behavior [20], anger and hostility [21], and suicidality [22,23]. Proneness to guilt and shame has been linked with gender role stress among men [24], while shame appears to moderate the relationship between conformity to masculine norms, self-esteem and self-compassion [1]. Research suggests that men tend to feel guilt and shame in connection with perceived lack of status and achievement, leading to subsequent suicidal behavior [25,26]. Indeed, a study of suicide notes found guilt to be a prominent theme among men’s reasons for attempting suicide [27].

Efforts to understand male proneness to guilt and shame would be greatly facilitated by a wider range of assessment options. While the Test of Self-Conscious Affect [28] is the most widely used measure of guilt and shame-proneness [29], its scenario-based measure response format is not appropriate for studies where scale brevity is essential. This is relevant, given a recent move towards models of next-generation ecological momentary assessment (EMA) of emotions and risk states, including suicidality [30], whereby individuals are prompted by smartphone apps to respond to scale items multiple times per day over a period of week(s). An alternative to TOSCA-3 is the Personal Feelings Questionnaire (PFQ-2) [31]. An adjective checklist designed to measure guilt and shame-proneness, its efficient layout makes it more amenable to studies where item brevity is important. Though there has been criticism of its adjective checklist approach, as checklist measures rely on individuals to explicitly and accurately differentiate between experiences of shame and guilt experiences, the PFQ-2 remains a widely used measure of guilt and shame within clinical [32] and non-clinical populations [33].

As the constructs of guilt and shame gain increasing research attention in relation to at-risk men, including those with recent military combat experience [22], history of sexual abuse [34], substance misuse [19] and suicidality [35], valid and reliable assessment is needed. An important psychometric property for any construct measure is support for its underlying factor structure [6].
The present study contributes to the literature on the assessment of guilt and shame by conducting a confirmatory factor analysis (CFA) of the PFQ-2. Despite its widespread use, we were unable to locate any CFA-based studies of the scale. A secondary aim was to develop a brief version of the PFQ-2 to enable wider application of the scale, and the assessment of guilt and shame, in EMA studies, where measures are necessarily brief. Finally, we sought to examine the clinical sensitivity of the guilt and shame-proneness subscales of a brief version of the PFQ-2 through comparisons with a sample of help-seeking men relative to the general community.

STUDY 1

Materials and methods

Participants

Data presented in Studies 1a, 1b and 1c were collected as part of a larger online survey of Canadian men’s mental health in April 2016 (N=1000). Respondents were sourced from a Canadian online survey provider and screened to ensure they met survey eligibility requirements (≥ 19 years, internet access, able to read English). The sample was stratified to ensure that the composition reflected the underlying distribution of the English speaking Canadian male population by age and province, as determined by 2011 Census data. Respondents were reimbursed for their time with proprietary panel points, which could later be exchanged for various rewards. The sample was randomly partitioned into three separate groups of equivalent size, corresponding to three data-sets used in Study 1a, 1b and 1c in order to facilitate:

- model evaluation (Study 1a);
- model refinement (i.e. calibration; Study 1b); and
- model validation (Study 1c).

Participant demographics for the three cohort data-sets are summarized in Table 1, which shows equivalence across groups.

| Table 1. Participant demographics across Study 1a, 1b and 1c |
|-----------------|-----------------|-----------------|-----------------|----------------|
|                | Study 1a n=333  | Study 1b n=332  | Study 1c n=335  | Statistic      | p        |
| Age, years: M (SD) | 48.79 (14.64)  | 49.84 (13.98)  | 50.26 (15.31)  | F              | 0.406    |
| Employment, n (%) |                |                |                |                |          |
| Employed full time | 186 (55.9)     | 160 (48.2)     | 174 (51.9)     | χ²             | 0.141    |
| Employed part time | 21 (6.3)       | 31 (9.3)       | 30 (9.0)       | χ²             | 0.300    |
| Self-employed     | 23 (6.9)       | 38 (11.4)      | 28 (8.4)       | χ²             | 0.110    |
| Job-seeking       | 16 (4.8)       | 19 (5.7)       | 19 (5.7)       | χ²             | 0.841    |
| Unable to work    | 17 (5.1)       | 10 (3.0)       | 9 (2.7)        | χ²             | 0.191    |
| Unemployed, not seeking | 3 (0.9)     | 4 (1.2)        | 5 (1.5)        | χ²             | 0.781    |
| Stay-at-home parent | 0 (0)         | 4 (1.2)        | 4 (1.2)        | χ²             | 0.134    |
| Retired           | 71 (21.3)      | 75 (22.6)      | 80 (23.9)      | χ²             | 0.731    |
| Education, n (%) |                |                |                |                |          |
| Some secondary    | 7 (2.1)        | 11 (3.3)       | 16 (3.4)       | χ²             | 0.161    |
| Secondary graduate | 40 (12.0)     | 32 (9.6)       | 43 (12.8)      | χ²             | 0.406    |
| Some college/trade school | 35 (10.5) | 46 (13.9)     | 35 (10.4)      | χ²             | 0.291    |
| College/trade school graduate | 81 (24.3) | 66 (19.9)     | 58 (17.3)      | χ²             | 0.076    |
| Some university   | 24 (7.2)       | 30 (9.0)       | 35 (10.4)      | χ²             | 0.337    |
| University graduate | 86 (25.8)   | 82 (24.7)      | 82 (24.5)      | χ²             | 0.911    |
| University postgraduate | 60 (18.0) | 65 (19.6)     | 66 (19.7)      | χ²             | 0.827    |
MATERIALS

Personal Feelings Questionnaire-2

The PFQ-2 [31] is a 16-item measure designed to assess guilt and shame proneness that uses a global adjective checklist measurement approach. Respondents indicate the frequency of their experiences consistent with guilt and shame-related affective descriptors on a scale from 0 (never) to 4 (continuously or almost constantly). The PFQ-2 includes six items assessing guilt-proneness (e.g., remorse) and 10 items assessing shame-proneness (e.g., feeling disgusting to others).

Patient Health Questionnaire

The Patient Health Questionnaire PHQ-9 [36] is a self-report rating scale of the nine symptoms of major depressive disorder, as specified in DSM-5. Respondents rate PHQ-9 items relative to the preceding 2-week period, on a scale from 0 (not at all) to 3 (almost every day). The PHQ-9 is a well-validated and commonly used clinical and research measure of depression severity. In the current set of studies, data for the PHQ-9 were used in Study 1c and Study 2.

PROCEDURE AND DATA ANALYSIS

The study was approved by the institutional behavioral research ethics board at the University of British Columbia. The topic of the online survey was not disclosed in the initial invitation to potential participants, and only potential respondents who went to the survey introduction page were advised that men’s mental health was the focus.

Confirmatory factor analysis (CFA) for Study 1a, 1b and 1c was undertaken with AMOS Version 22.0 using maximum likelihood estimation. Typically accepted fit indices for assessing model fit were reported; the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR) with the conventional model fit criteria used to assess model acceptance (CFI >0.95, SRMS <0.08, RMSEA <0.09, TLI >0.95) as recommended by Hu and Bentler [37]. It is further recommended that researchers refer to broader sources of information most relevant to their particular situation when assessing model fit in large samples [38], including theory and other psychometric considerations [39]. As such, model re-specification was based on examining standardized regression weights to ensure sub-scale items loaded strongly on the overarching latent construct. Items reporting low weights (<0.70), were considered to be poorly reflective of the overarching latent construct of either guilt – or shame-proneness, and were subsequently omitted.

RESULTS

Table 2. Model fit indices for competing models of the PFQ-2 in samples from Studies 1a, 1b and 1c

<table>
<thead>
<tr>
<th>Data-set</th>
<th>Model</th>
<th>d.f.</th>
<th>χ²</th>
<th>Sig.</th>
<th>χ²/d.f.</th>
<th>SRMR</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>LO 90</th>
<th>HI 90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1a</td>
<td>Model 1</td>
<td>103</td>
<td>329.87</td>
<td>&lt;0.001</td>
<td>3.203</td>
<td>0.043</td>
<td>0.927</td>
<td>0.915</td>
<td>0.081</td>
<td>0.072</td>
<td>0.091</td>
</tr>
<tr>
<td>Study 1b</td>
<td>Model 1</td>
<td>103</td>
<td>419.53</td>
<td>&lt;0.001</td>
<td>4.073</td>
<td>0.066</td>
<td>0.876</td>
<td>0.856</td>
<td>0.096</td>
<td>0.087</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>13</td>
<td>41.56</td>
<td>&lt;0.001</td>
<td>3.197</td>
<td>0.037</td>
<td>0.975</td>
<td>0.960</td>
<td>0.081</td>
<td>0.054</td>
<td>0.110</td>
</tr>
<tr>
<td>Study 1c</td>
<td>Model 2</td>
<td>13</td>
<td>28.65</td>
<td>&lt;0.001</td>
<td>2.204</td>
<td>0.026</td>
<td>0.986</td>
<td>0.978</td>
<td>0.060</td>
<td>0.030</td>
<td>0.090</td>
</tr>
</tbody>
</table>

Model 1, 16-item PFQ-2; Model 2, 7-item PFQ-2 Brief. Values in bold achieve model fit cut-off criteria, SRMR=Standardized Root Mean Square Residual, CFI=Comparative Fit Index, TLI=Tucker Lewis Index, RMSEA=Root Mean Square Error of Approximation, LO 90 = Low 90% Confidence Interval, HI 90 = High 90% Confidence Interval.
Results of the $\chi^2$ model fit test and fit indices are presented in Table 2. Fit indices for the 16-item, 2-factor model of the PFQ-2 were below the accepted range for the CFI and TLI. Standardized regression weights for the PFQ-2 are presented in Figure 1. Based on model fit indices, the 16-item model was rejected due to inadequate fit. The two-factor model was re-examined in a separate sample (Study 1b), and consideration given to evaluating an alternate model.

Results of the $\chi^2$ model fit test and fit indices are presented in Table 2. Indices for model 1 (16-item, 2-factor model) replicated results from Study 1a, demonstrating inadequate indices for the CFI and TLI. Given this, changes were made to the structural model in order to improve fit, based on inspection of the standardized regression weights (low loading items <0.70 were omitted), and modification indices generated by AMOS (Table 3). As can be seen in Figure 1, there were nine items with standardized regres-
A total of 40.5% were employed in some capacity, 52.4% were unemployed and the remainder were students or stay-at-home parents. Most (66.7%) had sought mental health treatment in the past. Diagnostic information is not available, as comprehensive case formulations, rather than formal diagnoses, are used to guide patients’ treatment in the clinic.

### Procedure and data analysis

The study was approved by the behavioral research ethics board at the University of British Columbia. Upon admission, patients provided informed consent and participated in the study by completing a survey consisting of the PHQ-9 and the PFQ-2. Bivariate and partial correlations were calculated to evaluate shame-free guilt (the correlation between guilt and depression controlling for shame) and guilt-free shame (the correlation between shame and depression controlling for guilt). One-sample \( t \)-tests were conducted to compare mean scores for the PHQ-9 and the guilt and shame-proneness subscales relative to Study 1c.

---

**STUDY 1C**

Results of the \( \chi^2 \) model fit test and fit indices for model 2 in the sample for Study 1c replicated those from the calibration sample in Study 1b (Table 2). In fact, all model fit indices in Study 1c were an improvement on those reported in Study 1b. The final model of the revised PFQ-2 (henceforth referred to as the PFQ-2 Brief) included seven items: three items for the guilt-proneness subscale and four items for the shame-proneness subscale. Descriptive statistics for the PFQ-2 Brief items are displayed in Table 3. See supplemental online material for PFQ-2 Brief items.

**STUDY 2**

**Materials and methods**

**Participants**

The participants in Study 2 were 42 consecutiely admitted male outpatients of a public psychiatric clinic based in Surrey, Canada. Mean age was 36.14 years (SD=12.23), most (53.7%) were single, and 14.6% were separated or divorced.
RESULTS

Mean scores on the PHQ-9 indicated that participants were experiencing moderate depression. Significant moderate bivariate correlations were observed between the PHQ-9 total score, and scores of the subscales for guilt-proneness ($r=0.45$, $p=0.004$) and shame-proneness ($r=0.35$, $p=0.030$), with the guilt – and shame-proneness subscales also moderately correlating with each other ($r=0.51$, $p=0.001$). One-sample t-tests were undertaken to evaluate differences between those in the Study 1c community sample ($n=335$) and those in the Study 2 clinical sample ($n=42$). Strong effects were observed. As expected, the clinical sample reported significantly higher depression ($M=16.87$ (SD=5.44), 95% CI 15.11–18.64) than those in the community ($M=5.20$ (SD=5.49), 95% CI 4.61–5.79; $t(39)=13.40$, $p<0.001$, $d=2.14$). Similarly, subscale mean scores for guilt-proneness were significantly higher for men in the clinical sample ($M=2.56$ (SD=1.01), 95% CI 2.39–2.74), relative to those in the community sample ($M=1.40$ (SD=0.82), 95% CI 1.32–1.50; $t(41)=5.04$, $p<0.001$, $d=1.26$). Those in the clinical sample also reported greater shame-proneness ($M=2.10$ (SD=0.89), 95% CI 1.94–2.25) than those in the community sample ($M=1.38$ (SD=1.04), 95% CI 1.27–1.49; $t(41)=6.86$, $p<0.001$, $d=0.074$).

DISCUSSION

The aim of the present study was to validate the factor structure of the widely used PFQ-2 in a large community sample of men, develop a short form of the measure and validate it in a clinical sample. In doing so, the study was the first to apply confirmatory validation techniques to the PFQ-2. Confirmatory factor analysis demonstrated poor factor structure for the 16-item PFQ-2 (Study 1a). A follow-up CFA with an independent sample (Study 1b) also revealed poor factor structure of the PFQ-2, demonstrating significant cross-loadings for both the shame and guilt subscales. A parsimonious model of guilt and shame (the seven-item PFQ-2 Brief) was found to have good model fit indices in a separate sample (Study 1c) and convergent validity was demonstrated in a clinical sample of help-seeking men (Study 2).

The PFQ-2 Brief may have some advantages over the PFQ-2 in that it offers a more efficient and more psychometrically robust self-report assessment of guilt – and shame-proneness. Use of the PFQ-2 Brief revealed significantly higher scores of guilt and shame in a clinical sample of help-seeking men relative to men in the general community. Furthermore, scores on both subscales of the PFQ-2 Brief were moderately correlated with a measure of depression symptomatology.

There has been some debate in the literature as to whether guilt and shame are equally related to psychopathology, and a large meta-analysis found a stronger relationship between shame and depression than between guilt and depression [17]. This appears in contrast to the present findings among help-seeking men (Study 2). Excessive and inappropriate feelings of guilt form part of the diagnostic criteria for a major depressive episode, and as such, the elevated level of guilt would be expected in the present clinical population. Nonetheless, while the association was greater between guilt and depression than shame and depression in Study 2, this difference was not statistically significant.

Given that shame and guilt are often encountered in treatment settings [40], a brief self-report measure that assesses the levels of shame and guilt an individual is experiencing is likely to have some utility for clinicians. For instance, if the measure reveals that a man is experiencing a significant level of shame, the treating therapist may need to focus on diffusing this aversive experience [18,41]. This is particularly the case given that experiences of shame are associated with a reluctance to disclose therapy-relevant material, which can impede the treatment process [42].

There is a compelling need for healthcare professionals to be cognizant of men’s distinct gendered healthcare needs [15,43]. In the context of men’s mental health, there may be particular value in clinicians modifying approaches to assessment and therapy with guilt – or shame-prone men, mindful of their preferences and values. Future work should look to more fully apprehend men’s shame and guilt processes as potential by-products of being in treatment. For example, needing help, which transgresses masculine ideals of autonomy and self-reliance, may
elicit guilt and shame in some men. A more thorough understanding of these processes might guide the efforts of health care professionals for framing help-seeking and being in treatment as a conduit to self-management.

LIMITATIONS AND FUTURE DIRECTIONS

The limitations of the present study include the use of a cross-sectional design. Subsequent research should look to determine the longitudinal predictive and discriminant validity of the PFQ-2 Brief as a measure of factors that may impede men’s mental health. The study offered a gendered analysis and a gender lens might further elaborate on some of our findings to advance ideas for effectively intervening with men. A next step for future research is to determine whether the factor structure of the PFQ-2 Brief can be replicated in a female population. Additionally, discriminant validity of the PFQ-2 Brief is yet to be determined, and future research should seek to determine the relationships between the shame and guilt subscales of the measure with other well-established measures of shame and guilt, such as the TOSCA-3 [28] and Compass of Shame Scale [44].

CONCLUSIONS

In summary, this study has demonstrated that the factor structure of the original PFQ-2 is not supported by CFA models, and a revised and brief measure with more sound psychometric properties is proposed in the PFQ-2 Brief. What the present study affords are important contributions to the assessment of guilt and shame in men. The PFQ-2 Brief appears to provide a valid and reliable measure for assessing guilt – and shame-proneness in men and as such it should aid further investigations of the manner in which the two affect styles impact help-seeking, treatment engagement, treatment outcomes, and men’s overall mental health.

REFERENCES


