



A brief measure of guilt and shame: validation of the Guilt and Shame Questionnaire (GSQ-8)

Thole H. Hoppen^a, Pascal Schlechter^b, Arnoud Arntz^c, Sophie A. Rameckers^c, Thomas Ehring^d and Nexhmedin Morina^a

^aInstitute of Psychology, Division of Clinical Psychology and Psychotherapy, University of Münster, Münster, Germany; ^bDepartment of Psychiatry, University of Cambridge, Cambridge, UK; ^cDepartment of Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands; ^dDepartment of Psychology, Division of Clinical Psychology and Psychotherapy, LMU Munich, Munich, Germany

ABSTRACT

Background: Guilt and shame regulate basic human processes such as social cognition and relations. Both emotions are also involved in the aetiology and maintenance of trauma-related mental disorders such as posttraumatic stress disorder (PTSD). However, a concise scale that adequately captures these constructs is currently lacking, impeding research efforts to understand them more thoroughly.

Objective: To this end, we developed the eight-item Guilt and Shame Questionnaire (GSQ-8) in English, German, and Dutch.

Method: We examined the reliability and validity of the GSQ-8 in a clinical sample of adults seeking treatment for childhood-trauma-related posttraumatic stress disorder ($n = 209$), a sample of adults who had suffered at least one traumatic life event reporting different levels of PTSD symptoms ($n = 556$), and a non-clinical sample of adults ($n = 156$).

Results: Theory-driven confirmatory factor analyses confirmed two correlated latent factors guilt and shame with four items for each factor. Across all samples, two-factor models yielded better model fit than one-factor solutions. Measurement invariance across the three samples, gender, and Dutch and German language was mostly established. Guilt and shame composite scores were associated with PTSD symptoms, depressive symptoms, life satisfaction, mental health-related quality of life, and self-blame, thus supporting scale validity. Importantly, both subscales predicted PTSD symptoms, depression, life satisfaction, and mental health-related quality of life over and above cognitions of self-blame.

Conclusions: The GSQ-8 is a parsimonious, reliable, and valid tool to assess guilt and shame in clinical, sub-clinical, and non-clinical populations, allowing applications across a broad range of research questions.

Una breve medida de culpa y vergüenza: validación del Cuestionario de Culpa y Vergüenza (GSQ-8)

Antecedentes: La culpa y la vergüenza regulan procesos humanos básicos como la cognición y las relaciones sociales. Ambas emociones también están implicadas en la etiología y el mantenimiento de los trastornos mentales relacionados con el trauma, como el trastorno de estrés postraumático (TEPT). Sin embargo, actualmente se carece de una escala concisa que capture adecuadamente estos constructos, lo cual impide los esfuerzos de investigación para comprenderlos más a fondo.

Objetivo: Para ello, desarrollamos el Cuestionario de Culpa y Vergüenza de ocho ítems (GSQ-8) en inglés, alemán y holandés.

Método: Examinamos la confiabilidad y validez del GSQ-8 en una muestra clínica de adultos que buscaban tratamiento para el trastorno de estrés postraumático relacionado con traumas infantiles ($n = 209$), una muestra de adultos que habían sufrido al menos un evento de vida traumático que reportaron diferentes niveles de síntomas de TEPT ($n = 556$), y una muestra no clínica de adultos ($n = 156$).

Resultados: Los análisis factoriales confirmatorios basados en la teoría confirmaron dos factores latentes correlacionados: la culpa y la vergüenza, con cuatro ítems para cada factor. En todas las muestras, los modelos de dos factores produjeron un mejor ajuste del modelo que las soluciones de un factor. La medición de invarianza mayormente establecidas entre las tres muestras fueron el género, y el idioma holandés y alemán. Las puntuaciones compuestas de culpa y vergüenza se asociaron con síntomas de TEPT, síntomas depresivos, satisfacción con la vida, calidad de vida relacionada con la salud mental y el culpabilizarse a si mismo, lo que respalda la validez de la escala. Es importante

ARTICLE HISTORY

Received 17 September 2022

Revised 4 November 2022

Accepted 4 November 2022

KEYWORDS

Guilt and Shame Questionnaire; guilt; shame; measurement; depression; PTSD

PALABRAS CLAVE

Cuestionario de Culpa y Vergüenza; culpa; vergüenza; medición; depresión; PTSD

关键词

内疚和羞耻问卷; 内疚; 羞耻; 测量; 抑郁; PTSD

HIGHLIGHTS

- We present the 8-item Guilt and Shame Questionnaire (GSQ-8), a parsimonious tool to assess guilt and shame.
- We found strong psychometric properties of the GSQ-8 across three samples and different languages.
- The GSQ-8 can be reliably used to assess shame and guilt in clinical and non-clinical work.

CONTACT Thole H. Hoppen [✉ thoppen@uni-muenster.de](mailto:thoppen@uni-muenster.de) Institute of Psychology, Division of Clinical Psychology and Psychotherapy, University of Münster, Fliegerstr. 21, Münster 48149, Germany

Supplemental data for this article can be accessed online at <https://doi.org/10.1080/20008066.2022.2146720>.

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

destacar que ambas subescalas predijeron los síntomas del TEPT, la depresión, la satisfacción con la vida y la calidad de vida relacionada con la salud mental por encima de las cogniciones de culpabilidad.

Conclusiones: GSQ-8 es una herramienta parsimoniosa, confiable y válida para evaluar la culpa y la vergüenza en poblaciones clínicas, subclínicas y no clínicas, lo que permite aplicaciones en una amplia gama de preguntas de investigación.

一个内疚和羞耻的简短测量：内疚和羞耻问卷 (GSQ-8) 的验证

背景: 内疚和羞耻调节基本的人类过程，如社会认知和关系。这两种情绪也与创伤相关精神障碍如创伤后应激障碍 (PTSD) 的病因和维持有关。然而，目前缺乏能够充分捕捉这些结构的简洁量表，阻碍了更彻底理解它们的研究工作。

目的: 为此，我们开发了英语、德语和荷兰语的八条目内疚和羞耻问卷 (GSQ-8)。

方法: 我们在一个寻求治疗童年期创伤相关创伤后应激障碍的成人临床样本 ($n = 209$)，一个至少经历过一次创伤性生活事件并报告了不同程度 PTSD 症状的成人样本 ($n = 556$)，和一个成人的非临床样本 ($n = 156$) 中检验了 GSQ-8 的信效度。

结果: 理论驱动验证性因素分析证实了两个相关的潜在因素内疚和羞耻，每个因素有四个条目。在所有样本中，双因素模型比单因素解决方案拟合更好。三个样本、性别、荷兰语和德语的测量不变性基本成立。内疚和羞耻综合得分与 PTSD 症状、抑郁症状、生活满意度、心理健康相关生活质量和自责相关，从而支持量表的效度。重要的是，除自责认知之外，这两个分量表都预测了 PTSD 症状、抑郁、生活满意度和心理健康相关生活质量。

结论: GSQ-8 是一个简洁、可靠且有效的工具，用于评估临床、亚临床和非临床人群的内疚和羞耻，可以在广泛的研究问题中应用。

1. Introduction

Guilt and shame are fundamental human emotions involved in the regulation of basic human cognition and behaviour including consciousness, moral cognition, social cognition, and social relations (Tangney et al., 2007; Tangney & Dearing, 2003). While guilt revolves around internal, unstable attributions for perceived failure, shame revolves around internal, stable and uncontrollable attributions for perceived failure (Tracy & Robins, 2006). Although guilt and shame serve adaptive functions such as driving prosocial behaviour they can also be maladaptive. The *social-adaptive hypothesis* argues that shame is an inherently maladaptive emotion since failure is perceived as uncontrollable whereas guilt is proposed to be an inherently adaptive emotion since failure is perceived as controllable (Dempsey, 2017). However, according to the well-supported *functionalist hypothesis*, the adaptiveness (vs. maladaptiveness) of guilt and shame depends on various factors such as the frequency and intensity of guilt and shame and their situational appropriateness (Dempsey, 2017; Leach & Cidam, 2015; Muris & Meesters, 2014; Tignor & Colvin, 2016). Understanding differential effects of shame and guilt in the context of adaptive behaviour, dysfunctional beliefs, and psychopathology is crucial (DeCou et al., 2016; Fletcher, 2011; Held et al., 2015; Pugh et al., 2015; Street & Arias, 2001). However, a granulated understanding of guilt and shame is impeded by a lack of a concise scale that captures these constructs in their entire complexity. Here, we therefore present and validate the 8-item Guilt and Shame Questionnaire (GSQ-8), a parsimonious tool to assess guilt and shame.

Guilt and shame are highly prevalent in trauma survivors suffering posttraumatic stress disorder (PTSD)

or major depression (Beck, Reich et al., 2015; DeCou et al., *in print*; Kim et al., 2011; Lee et al., 2001; Tangney et al., 1992; Wilson et al., 2006). It has been proposed that guilt and shame are involved in the aetiology and maintenance of trauma-related disorders for instance by driving behavioural avoidance and fostering dysfunctional beliefs about the trauma and its consequences (DeCou et al., 2016; Fletcher, 2011; Held et al., 2015; Pugh et al., 2015; Street & Arias, 2001). Some studies have directly compared the relative contribution of guilt vs. shame to PTSD (for an overview, see Cunningham, 2020), indicating that guilt and shame independently predict PTSD severity, with a somewhat higher association for shame than for guilt (Bannister et al., 2018; Cunningham, 2020; Shi et al., 2021). Further, shame but not guilt mediated PTSD symptom reduction during psychotherapeutic group treatment (Ginzburg et al., 2009). Similarly, guilt and shame could clearly be distinguished in a meta-analysis focusing on depression, with shame being more strongly related to depression severity than guilt (Kim et al., 2011). However, guilt and shame are usually highly correlated in clinical samples (Bannister et al., 2018).

Despite the importance of understanding guilt and shame in both human flourishing and psychopathology, the assessment of guilt and shame is impeded by the complexity of these constructs, as reflected in current measurement approaches (Ferguson & Crowley, 1997). Although several existing guilt and shame scales have contributed to a better understanding of these constructs, these existing scales have their limitations (for a review, see Lear et al., 2022). The Personal Feelings Questionnaire-2 (PFQ2; Harder et al., 1987) and the Adapted Shame and Guilt Scale (ASGS,

Hoblitzelle, 1982) are among the psychometrically soundest measures of shame and guilt. The PFQ-2 is a 22-item measure assessing shame-proneness and guilt-proneness, comprising 10 items assessing 'shame-proneness', 6 items assessing 'guilt-proneness' and 6 filler items. Despite good psychometric properties of the PFQ-2 in US samples (Harder et al., 1992; Harder & Greenwald, 1999) along good test-retest reliability (Di Sarno et al., 2019), a recent systematic review reported mixed evidence for its structural validity (Lear et al., 2022). Three studies found adequate model fit of a two-factor solution (Di Sarno et al., 2019; Harder & Zalma, 1990; Rice et al., 2018) and one study found poor model fit for this factor solution (Eterović et al., 2020). Importantly, the 22-item scale may not be considered as parsimonious in some research settings (Harder & Zalma, 1990).

The 24-item ASGS captures guilt and shame with good evidence for the structural validity of the intended two-factor structure of the ASGS (Harder & Zalma, 1990; Hoblitzelle, 1982), good internal consistency (ASGS Shame $\alpha = 0.83$; ASGS Guilt $\alpha = 0.89$) and good 2-week test-retest reliability (ASGS Shame $r = 0.93$; ASGS Guilt $r = 0.95$). However, evidence for the two-factor structure of the ASGS is limited by the exclusive use of principal component analysis, potential conflation of guilt and shame constructs within this measure, small sample sizes and lacking clarity in terms of the description of study procedures (Lear et al., 2022).

Similarly, other measures of guilt and shame come with their short-comings (Lear et al., 2022). Few scales have evidence concerning measurement invariance. To our knowledge, only the Experience of Shame Scale (ESS) has proven metric (but not scalar) measurement invariance across clinical vs. non-clinical samples (Vizin et al., 2016) and the PFQ-2 metric measurement invariance across gender (Di Sarno et al., 2022). This is an important omission in the literature because it remains unknown whether differences in guilt and shame across gender or clinical vs. non-clinical samples can be truly attributed to differences in the latent construct or whether they represent an artefact of biased measurement properties (Meredith, 1993). Moreover, few scales have been validated in multiple languages.

2. The present study

In sum, current measurement approaches have their limitations in terms of their parsimony, psychometric properties, content validity, and utility. Therefore, we developed the 8-item Guilt and Shame Questionnaire (GSQ-8), which assesses the feelings of guilt and shame during the last four weeks. As such, guilt and shame were defined in the present project as emotions, rather than cognitions or

behaviours. To develop a measure that addresses previous short-comings, we merged the best-performing items of PFQ2 and ASGS into a concise 8-item scale in three languages (i.e. English, German, and Dutch). We chose items based on their content validity, psychometric performance and translatability and performed a thorough psychometric validation in a clinical sample (i.e. treatment seeking sample with PTSD), traumatized sample and non-clinical sample. Following a theory-driven approach, we used confirmatory factor analyses (CFA) to test the factor structure of the GSQ-8. We hypothesized that a two-factor solution with the two correlated latent factors guilt and shame describes the data better than a one-factor solution representing one construct. We further scrutinized measurement invariance across gender, languages, and subsamples to ensure that the GSQ-8 measures the same underlying latent constructs across these groups. Based on previous literature, we expected both subscales to significantly correlate with one another (Shi et al., 2021) and with PTSD symptoms (Shi et al., 2021), depressive symptoms (Kim et al., 2011), life satisfaction (Bugay & Demir, 2011), mental health-related quality of life (Owens et al., 2009; Persons et al., 2010), and self-blame (Duncan & Cacciatore, 2015). We further tested whether both GSQ-8 subscales predicted PTSD symptoms, depression, life satisfaction, and mental health-related quality of life beyond cognitions of self-blame.

3. Method

3.1. Participants

3.1.1. Clinical sample

In total, 209 individuals seeking treatment for childhood trauma-related posttraumatic stress disorder participated in this study within the scope of an international multicentre clinical trial (Wibbelink et al., 2021). Data were collected in five, three and two mental healthcare centres in the Netherlands, Germany, and Australia (Wibbelink et al., 2021). Participants were eligible if they (1) were between 18 and 70 years old; (2) had a primary diagnosis of PTSD as assessed by the Structural Clinical Interview for DSM-5-Clinician Version or Research Version (SCID-5-CV/RV); (3) had experienced the worst traumatic event before the age of 16; and (4) had been experiencing PTSD symptoms for longer than three months. This sample (76.56% females) was middle-aged (mean = 37.13; $SD = 14.01$) and highly educated (57.42% finished high school, 12.44% had a university degree). Data were collected between August 2018 and April 2022 in diagnostic interviews mostly face-to-face or seldomly via telephone or online due to the COVID-19 lockdowns.

3.1.2. Traumatized sample

Data were collected between October 2019 and January 2020 as part of a larger project (Hoppen & Morina, 2021). Data were collected online via psyweb (<https://psyweb.uni-muenster.de/>), a non-commercial online panel for psychology research. Inclusion criteria were: (1) being at least 18 years old, (2) experience of at least one traumatic life event as defined by the Life-Events-Checklist (Blake et al., 1995), (3) fluent German literacy, (4) no acute suicidality, and (5) no lifetime history of a psychotic disorder. This sample (61.84% females, 0.18% diverse) was middle-aged (mean = 51.01; $SD = 14.28$) and highly educated (76.62% finished high school, 52.70% had a university degree).

3.1.3. Non-clinical sample

The data of the non-clinical control sample were collected online between March and April 2020. Participants were recruited via online websites, such as Facebook groups. Participants were included when they (a) were between 18 and 70 years old, (b) did not have a current DSM-5 diagnosis, and (c) gave informed consent. This sample (79.49% females, 0.64% diverse) was middle-aged (mean = 36.31; $SD = 16.30$) and highly educated (96.15% finished high school, 27.56% had a university degree). See Table A1 in the [supplementary material](#) for an overview of sample characteristics across the three samples.

3.2. Measures

3.2.1. Guilt and shame

To develop the GSQ-8, we defined guilt and shame as emotions, rather than cognitions or behaviours with guilt being defined as an emotion arising from remorse for having done something wrong and shame as an emotion arising from thinking that one deserves contempt as a person in relation to other people's opinions. The GSQ-8 assesses the *frequency* of experiencing these emotions over a period of four weeks, mimicking the time interval often used to assess psychopathological symptoms (e.g. four weeks for PTSD symptoms). We aimed to develop a short questionnaire that would be easy to administer in different research designs. To this end, we chose the best-performing items from the PFQ-2 (Harder et al., 1987) and the ASGS (Hoblitzelle, 1982). Both scales display good psychometric properties overall (Harder & Zalma, 1990; Lear et al., 2022). However, not all items perform well and the scales are not parsimonious. Moreover, some items did not seem to represent an *emotional* experience, and others were judged to be not specific to guilt or shame, and yet others showed translation problems into German and Dutch. Our item selection was therefore based on psychometric properties, content validity, and

translatability into German and Dutch. The decision to include Item 2 ('I felt guilty') and Item 8 ('I felt intense guilt') despite being similar was based on the premise that our scale tapped into the frequency of endorsing these emotions. By including both items, we aimed at increasing variance on the severity spectrum of this construct. Accordingly, we anticipated that both items would differentiate well on the latent construct of guilt, particularly in populations with intense endorsement of guilt. In a first step, four authors of the manuscript selected an initial pool of 20 items. In several rounds of discussions and revisions, these authors selected 8 items based on psychometric properties, content validity and translatability. The items were first formulated in English, and then translated into Dutch and German. We used a standard forwards and backwards translation procedure. We chose a five-point response format with the response options (0) 'never', (1) 'once or twice', (2) 'three to four times', (3) 'more than once per week', and (4) 'daily'.

3.2.2. Posttraumatic stress disorder symptoms

The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5, Weathers et al., 2013) assesses DSM-5 PTSD symptom severity during the past months with 20 self-report items scored on a 5-point Likert scale from 0 (not at all) to 4 (extremely). The PCL-5 has good psychometric quality in traumatized populations (Blevins et al., 2015; Boeschoten et al., 2018; Bovin et al., 2016; Krüger-Gottschalk et al., 2017; Wortmann et al., 2016). The index-trauma version of the PCL-5 was utilized for the clinical sample (rather than the all trauma version). The internal consistency of the PCL-5 in the present study was good (clinical: $\alpha = 0.95$; traumatized: $\alpha = 0.87$).

3.2.3. Depression

The Beck Depression Inventory (BDI-II) is a 21-item self-report measure assessing endorsement of depressive symptoms during the past two weeks (Beck, Steer et al., 1996). Items are scored on a four-point Likert scale from zero to three with higher scores indicating worse endorsement of the depressive symptom. The BDI-II is one of the most widely used self-report measures for the assessment of depressive symptomatology with good psychometric properties in various contexts and populations including English-, Dutch- and German-speaking populations (Dobson & Ahnberg, 1998; Wang & Gorenstein, 2013). The internal consistency of the BDI-II in the present study was good (clinical: $\alpha = 0.91$; non-clinical: $\alpha = 0.87$).

The Depression Anxiety Stress Scale (DASS) is a 21-item self-report measure assessing depression, anxiety and stress symptoms with seven items per subscale (Lovibond & Lovibond, 1995). The DASS has good psychometric properties across various clinical

and non-clinical populations and contexts including German-speaking populations (Brown et al., 1997; Crawford & Henry, 2003). The 7-item depression subscale of the DASS utilized in the present study yielded good internal consistency in the traumatized sample ($\alpha = 0.91$).

3.2.4. Satisfaction with life

The Satisfaction With Life Scale (SWLS, Diener et al., 1985) is a 5-item self-report measure assessing global satisfaction with life. Items are scored on a 7-point Likert scale. The SWLS is the most widely used measure for general life satisfaction and good psychometric properties have been reported across various contexts and populations including German-speaking populations (Glaesmer et al., 2011; Pavot & Diener, 1993). The internal consistency of the SWLS in the present study was good (traumatized sample: $\alpha = 0.91$).

3.2.5. Mental health-related quality of life

The Mental Health Quality of Life (MHQoL) is a self-report measure assessing quality of life in individuals with mental health problems (van Krugten et al., 2022). The MHQoL assesses seven mental health-related quality of life-related domains on a 4-point Likert scale: self-image, independence, mood, relationships, daily activities, physical health, and hope. The validity of the MHQoL has been cross-validated with good psychometric qualities in a Dutch sample of mental healthcare service users and a non-clinical sample of the general Dutch population (Eekers et al., 2021). The internal consistency of the MHQoL in the present study was good (clinical sample: $\alpha = 0.79$).

3.2.6. Posttraumatic cognitions – self-blame

The Posttraumatic Cognitions Inventory (PTCI) is a 33-item self-report measure assessing trauma-related cognitions across the following three domains: negative cognitions about the self, negative cognitions about the world and self-blame (Foa et al., 1999). Each item is rated on a 7-point Likert scale with higher scores indicating higher endorsement of maladaptive posttraumatic cognitions. The PTCI has been used widely with good psychometric properties in traumatized populations (Beck, Coffey et al., 2004; Müller et al., 2010; van Emmerik et al., 2006). Here, we utilized the 5-item self-blame subscale of the PTCI which had a good internal consistency in all applicable samples (clinical: $\alpha = 0.79$; traumatized: $\alpha = 0.83$).

3.3. Transparency and openness

The English, Dutch and German versions of the GSQ-8 and the analysis code to reproduce our results are available on the open science framework (https://osf.io/t3dh7/?view_only=442bdfeda8ec4a429c07dee864d6e442). The pre-registration of the present study was

part of the clinical trial registration (<https://www.trialregister.nl/trial/6965>) and published as part of the design article of the larger multi-centre trial (i.e. clinical sample; Wibbelink et al., 2021).

3.4. Ethical approval

All study protocols were approved by the respective local institutional ethics committees.

3.5. Analysis procedure

3.5.1. Factor solutions

Analyses were performed in R (R Core Team, 2013) version 4.01. Factor models were analysed using the *lavaan* package (Rosseel, 2012). Our data analytical strategy was informed by theory. Accordingly, we used confirmatory factor analyses (CFA) to test the model fit of two-factor structures that were derived from prior literature (Harder & Zalma, 1990). First, we examined a unidimensional factor solution with all items loading on one underlying latent factor. Second, we tested a two-factor solution with the two correlated latent factors *guilt* and *shame*. Both factor solutions were tested separately for the clinical sample, traumatized sample and non-clinical sample.

Given that our response options were not equidistant, data were treated as ordinal in our factor models. Accordingly, we used the weighed least squares mean and variance adjusted (WLSMV) estimator (Asparouhov & Muthén, 2010). The following criteria were used to evaluate overall model fit: Comparative Fit Index (CFI) and Tucker–Lewis Index (TLI) values should be > 0.95 to indicate good fit and at least > 0.90 to indicate acceptable fit. The Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residuals (SRMR) should be $< .05$ and $< .08$ to indicate good or acceptable fit, respectively (Browne & Cudeck, 1992; Hu & Bentler, 1999). For all three samples, the one-factor solution was tested against the two-factor solution with a scaled chi-square difference test. A non-significant test indicates model equivalence, and that the simpler model should be used. These two models were treated as nested as a two-factor solution with a perfect correlation of one between the two-latent factors resembles the unidimensional factor solution.

3.5.2. Measurement invariance across samples and gender

We tested whether any differences between groups in the observed scores are indeed attributable to true-score differences (Meredith, 1993). Hence, we systematically examined measurement invariance across the three samples (i.e. clinical, traumatized and non-

clinical samples), language of the applied GSQ-8 version (German and Dutch¹), and gender (males & females) with multigroup CFA. To demonstrate measurement invariance, increasingly constrained models were tested against each other. The respective invariance constraints on the model parameters were added at each step in addition to the constraints introduced in the step before. First, the factor structure was constrained to be equal across the three samples/language/gender to investigate whether the construct is equally represented among these groups (configural invariance). Second, the factor loadings were additionally constrained to be equal across the three samples/language/gender to discern whether the items equally relate to the underlying factors (weak/metric invariance). Third, item thresholds were additionally constrained to be equivalent to gauge whether the observed thresholds conditional on the latent factor do not differ across the three samples/language/gender (strong/scalar invariance). Fourth, the residual variances of the items were set equal to scrutinize whether the variance in the items not explained by the latent factor does not differ across the three samples/language/gender (strict/residual invariance; Meredith, 1993). To detect violations of the measurement invariance assumptions, we evaluated changes (Δ) in the CFI and RMSEA. A violation of measurement invariance is indicated when Δ CFI exceeds .010 and Δ RMSEA exceeds .007 (Chen, 2007; Meredith, 1993). When measurement invariance was violated, we tested partial measurement invariance models by iteratively freeing parameters according to their unconstrained between-group discrepancies (Byrne et al., 1989; Guenole & Brown, 2014).

3.5.3. Associations with external constructs

First, we tested convergent validity of the GSQ-8 by correlating *total* scale, *guilt* subscale, and *shame* subscale composite scores with external constructs. In the clinical subsample, we correlated these scores with PTSD symptoms, life satisfaction, and self-blame. In the traumatized sample, we used PTSD symptoms, mental health-related quality of life, and self-blame, while we used depressive symptoms in the non-clinical sample. Going beyond single associations, we additionally tested whether the GSQ-8 would be associated with clinically relevant variables beyond post-traumatic cognitions of self-blame in the clinical and the traumatized sample. Here, we ran multiple regression models using both the *guilt* and *shame* subsample scores while adjusting for self-blame. PTSD symptoms, depression, life-satisfaction, and mental health-related quality of life served as criterion variables in these regression models.

4. Results

4.1. Descriptive statistics

There were no missing data in the three samples. Descriptive statistics for all items and all three samples are shown in Table 1. Descriptively, item mean values were highest in the clinical sample, followed by the traumatized and the non-clinical sample. In the clinical and the non-clinical sample, Item 2 ('I felt guilty') had the highest mean endorsement. In the traumatized sample, Item 3 ('I felt regret') had the highest reported frequency.

4.2. Factor solutions

4.2.1. Clinical sample

Table 2 depicts the model fit of all factor models for all three samples. The unidimensional factor model had good fit according to the CFI and TLI, and acceptable fit according to the SRMR. RMSEA suggested non-acceptable model fit. Table 3 displays all standardized factor loadings. All items loaded well on the unidimensional factor (all $\lambda \geq .65$). The unidimensional scale yielded excellent internal consistency ($\alpha = .90$; $\omega_{total} = .90$). The two-factor model had good model fit according to the CFI, TLI, and SRMR, while the RMSEA indicated acceptable fit. All items loaded well on their respective factors. Internal consistencies were good for both the *guilt* factor ($\alpha = .87$; $\omega_{total} = .87$) and the *shame* factor ($\alpha = .84$; $\omega_{total} = .85$). The latent factors were positively correlated with $r = .79$. The two-factor model displayed better model fit than the unidimensional model, $\chi^2(1) = 39.25$, $p < .001$.

4.2.2. Traumatized sample

The unidimensional factor model had good fit according to the CFI and TLI, and acceptable fit according to the SRMR and RMSEA. Again, all items loaded well on the unidimensional factor (all $\lambda \geq .54$). Internal consistency was good ($\alpha = .87$; $\omega_{total} = .87$). The two-factor model yielded good model fit according to the CFI, TLI, and SRMR, and acceptable fit according to the RMSEA. All items had good factor loadings on their respective factors. Internal consistencies were good for the *guilt* factor ($\alpha = .81$; $\omega_{total} = .82$) and acceptable for the *shame* factor ($\alpha = .76$; $\omega_{total} = .77$). The latent correlation between the factors was strong ($r = .88$). The two-factor model was superior to the unidimensional model, $\chi^2(1) = 30.87$, $p < .001$.

4.2.3. Non-clinical sample

For the unidimensional factor model, the CFI and TLI suggested good fit and the RMSEA and SRMR acceptable fit. Factor loadings were good (all $\lambda \geq .54$). Internal consistency was good ($\alpha = .83$; $\omega_{total} = .83$). For the two-factor model CFI, TLI, and RMSEA indicated good

Table 1. Descriptive statistics of the GSQ-8 in the three samples.

Item during the last 4 weeks ...	Clinical (N = 209)				Traumatized (N = 566)				Non-clinical (N = 156)			
	M	SD	Sk	Kt	M	SD	Sk	Kt	M	SD	Sk	Kt
1 ... I felt embarrassed	1.48	1.18	0.50	-0.70	0.70	0.87	1.37	1.81	0.74	0.80	1.23	1.89
2 ... I felt guilty	2.29	1.39	-0.28	-1.21	1.05	1.21	0.98	-0.15	1.03	0.95	0.87	0.49
3 ... I felt regret	2.06	1.37	-0.06	-1.26	1.39	1.16	0.64	-0.45	0.85	0.94	1.33	1.87
4 ... I felt disgusting	1.96	1.51	0.00	-1.49	0.40	0.95	2.50	5.36	0.29	0.68	2.61	7.19
5 ... I felt remorse	1.56	1.38	0.46	-1.10	0.67	0.93	1.53	2.03	0.32	0.70	2.61	7.47
6 ... I felt humiliated	1.56	1.41	0.40	-1.20	0.64	1.01	1.69	2.12	0.25	0.59	2.95	11.22
7 ... I felt ashamed	2.16	1.44	-0.14	-1.35	0.68	1.03	1.54	1.56	0.74	0.86	1.55	3.08
8 ... I felt intense guilt	1.96	1.49	-0.01	-1.48	0.36	0.90	2.69	6.53	0.28	0.66	3.29	13.20

Note. M = Mean; SD = Standard Deviation; Sk = Skewness; Kt = Kurtosis. For all item, 0 was the Min and 4 the Max.

model fit. The SRMR suggested acceptable model fit. In addition, the χ^2 was non-significant. All items loaded well on their respective factors. Internal consistencies were acceptable for the *guilt* factor ($\alpha = .74$; $\omega_{total} = .75$) and the *shame* factor ($\alpha = .73$; $\omega_{total} = .74$). The latent correlation between the factors was $r = .82$. The two-factor model had better model fit than the unidimensional model, $\chi^2(1) = 15.09$, $p < .001$.

4.2.4. Measurement invariance

According to ΔCFI , strict measurement invariance could be established across the three samples for the two-factor solution (Table 4). For the one-factor solution, strong invariance was established. Changes in the RMSEA, however, indicated violations of measurement invariance for both factor solutions when transitioning from the configural to the metric invariance model. To test partial invariance, we set factor loadings of item 8 free ('I felt intense guilt') because it had the high factor loading discrepancy in the non-clinical sample compared to the clinical and traumatized sample. This led to partial strict measurement invariance for the one-factor solution but not the two-factor solution. Non-invariance was mainly attributable to somewhat lower yet still good factor loadings of item 8 in the non-clinical sample. Setting other factor loadings free did not lead to changes in our conclusions.

For both factor solutions (unidimensional/two-factor solution) we could demonstrate the highest level of measurement invariance across gender (strict invariance). Model fit did not deteriorate substantially for any of the tested facets when model parameters were increasingly constrained.

According to ΔCFI , strict measurement invariance could be established across languages for both factor solutions. Changes in the RMSEA indicated violations of measurement invariance for both factor solutions in the metric invariance model. Setting the factor loadings of item 8 free ('I felt intense guilt') led to partial strict measurement invariance for the one-factor solution and strong partial invariance for the two-factor solution.

4.2.5. Associations with external constructs

Table 5 shows single association of all three subsamples with external constructs. In all three subsamples, total

scores and the subscales correlated highly among each other. In the clinical sample, all three scores (i.e. total score, guilt score & shame score) correlated positively with PTSD symptoms, depression, self-blame, and poor mental health-related quality of life. In the traumatized sample, all three scores correlated positively with PTSD symptoms, depression, and self-blame, and negatively with life satisfaction. In the non-clinical sample, they were positively associated with depressive symptoms. We found no gender differences in the clinical sample. In the traumatized sample, females reported higher values on all three scores. In the non-clinical sample, we only found a weak negative association between male gender and *shame* but not the *total* scale or the *guilt* subscale.

4.2.6. Controlling for self-blame

Table 6 depicts multiple regression models, in which both subscale scores *guilt* and *shame* predict the outcome variables while adjusting for cognitions of self-blame in these regressions. After adjusting for self-blame both *guilt* and *shame* still predicted PTSD symptoms, depressive symptoms, and poor mental health-related quality of life in the clinical sample. For PTSD symptoms and poor mental health-related quality of life, the effect of self-blame was not significant. In the traumatized sample, both *guilt* and *shame* were still significantly associated with PTSD symptoms and depression (positively) as well as life satisfaction (negatively).

5. Discussion

We developed a brief measure of guilt and shame, the GSQ-8. Overall, the GSQ-8 yielded good psychometric properties in a clinical sample, a traumatized sample and a non-clinical sample. The two-factor solution distinctly capturing guilt and shame had superior model fit compared to a one-factor solution. Measurement invariance across the subsamples, gender and Dutch and German language was mostly established. Both GSQ-8 subscales predicted clinically relevant variables over and above cognitions of self-blame.

Leading theories on guilt and shame postulate that these are related but distinct self-conscious emotions

Table 2. Model fit for the unidimensional and two-factorial solution of the GSQ-8.

	$\chi^2(df)$	<i>P</i>	CFI	RMSEA [95% CI]	SRMR	TLI
Unidimensional – Clinical	123 (20)	<.001	.986	.158 [.132; .185]	.080	.980
Two-factor – Clinical	40 (19)	.003	.997	.074 [.042; .105]	.049	.996
Unidimensional – Traumatized	82 (20)	<.001	.992	.075 [.059; .092]	.059	.989
Two-factor – Traumatized	55 (19)	<.001	.995	.059 [.041; .077]	.049	.993
Unidimensional – Non-clinical	32 (20)	.038	.990	.064 [.015; .102]	.070	.990
Two-factor – Non-clinical	20 (19)	.385	.999	.020 [.000; .074]	.056	.999

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residuals, TLI = Tucker-Lewis index; df = degrees of freedom, 95% CI = 95% confidence interval.

(Tangney & Dearing, 2003). In line with such theoretical notions, the two-factor solution for the GSQ-8 had better fit than the one-factor solution in all three samples (i.e. clinical, traumatized, and non-clinical sample), indicating that our scale captures guilt and shame as distinct yet correlated constructs. Although we found no differential effects of guilt and shame concerning their correlations with external constructs, the psychometrically supported distinction between these factors in our scale allows for important research applications building upon previous findings. For instance, shame but not guilt mediated PTSD symptom reduction during psychotherapeutic group treatment (Ginzburg et al., 2009), and guilt and shame could clearly be distinguished in a meta-analysis focusing on depression, with shame being more strongly related to depression severity than guilt (Kim et al., 2011). Accordingly, the GSQ-8 is well-suited to thoroughly study such important aspects in PTSD and depression research.

Notably, correlations in our study were comparable in size to those found in meta-analytic research on the relation between guilt, shame and trauma-related disorders. Shi et al. (2021), for instance, found large correlation between guilt and trauma-specific shame ($r = .544$, 95% CI = .468–.612) and guilt and generalized shame ($r = .504$, 95% CI = .404–.592) across various samples of patients suffering PTSD. Furthermore, good reliability was found for both factors across the three samples for the two-factor solution. In addition, factor loadings for all items were excellent, rendering them good discriminators between individuals on the latent construct. Item 8 ('I felt intense guilt') had the highest factor loadings in the clinical and the traumatized subsample and is therefore well-suited to differentiate between individuals

on the latent construct. This item implies a high level of guilt severity and may therefore be particularly pertinent to reflect interindividual differences in the construct of guilt in clinical or traumatized populations, which is in line with both theoretical notions (Lee et al., 2001) and empirical findings on the incidence and the role of guilt in traumatized populations (Pugh et al., 2015; Shi et al., 2021).

Moreover, the GSQ-8 demonstrated high levels of measurement invariance. Measurement invariance was established across gender and mostly established across the three samples (i.e. clinical, traumatized and non-clinical) and Dutch and German language. Item 8 ('I felt intense guilt') appeared to be the only somewhat problematic item in terms of measurement invariance with high discrepancy in factor loadings for the non-clinical sample compared to the clinical and traumatized samples with comparably low (yet still good) factor loading on item 8 for the former. The overall high levels of measurement invariance found in the present study enable the interpretation of differences between groups or gender as differences that can be attributed to differences in the latent construct.

In this regard, the GSQ-8 demonstrates evidence beyond other scale developments that did not test measurement invariance or only report evidence for metric measurement invariance of the PFQ-2 across gender (Di Sarno et al., 2019). We did not find evidence for gender differences in the frequency of guilt and shame in the clinical sample. Yet, females reported more frequent experiencing of guilt and shame than males in the traumatized sample and more frequent experiencing of *shame* (but not *guilt*) in the non-clinical sample. Results are in line with a meta-analysis on gender differences in guilt and shame which concluded that females tend to report

Table 3. Standardized factor loadings for the factorial solutions in the CFA.

Item during the last 4 weeks ...	Clinical			Traumatized			Non-clinical		
	1 F	Shame	Guilt	1 F	Shame	Guilt	1 F	Shame	Guilt
1 ... I felt embarrassed	.65	.69	–	.54	.56	–	.69	.72	–
2 ... I felt guilty	.90	–	.91	.86	–	.88	.81	–	.84
3 ... I felt regret	.75	–	.78	.65	–	.66	.78	–	.81
4 ... I felt disgusting	.71	.75	–	.81	.85	–	.63	.67	–
5 ... I felt remorse	.71	–	.74	.74	–	.76	.66	–	.68
6 ... I felt humiliated	.80	.83	–	.68	.71	–	.63	.66	–
7 ... I felt ashamed	.88	.96	–	.83	.87	–	.75	.82	–
8 ... I felt intense guilt	.91	–	.94	.93	–	.95	.60	–	.62

Table 4. Measurement invariance of the unidimensional and the two-factorial solution across subsample, gender and language.

Unidimensional	$\chi^2(df)$	CFI	RMSEA	ΔCFI	$\Delta RMSEA$	Two-factors	$\chi^2(df)$	CFI	RMSEA	ΔCFI	$\Delta RMSEA$
Subsample						Subsample					
Configural	239 (60)	.989	.099			Configural	116 (57)	.996	.058		
Metric	355 (74)	.983	.112	.006	.013	Metric	223 (69)	.990	.086	.006	.028
Metricp ^a	309 (72)	.985	.104	.004	.005	Metricp	180 (67)	.993	.074	.003	.016
Scalarp	467 (118)	.987	.098	.002	.006	Scalarp	303 (111)	.988	.075	.005	.001
Strictp	585 (134)	.972	.105	.015	.007	Strictp	427 (127)	.981	.088	.007	.013
Gender						Gender					
Configural	177 (40)	.994	.087			Configural	91 (38)	.998	.056		
Metric	190 (47)	.994	.082	.000	.005	Metric	103 (44)	.998	.054	.000	.002
Scalar	203 (70)	.995	.064	.001	.018	Scalar	115 (66)	.998	.040	.000	.014
Strict	220 (78)	.994	.063	.001	.001	Strict	130 (74)	.998	.041	.000	.001
Language						Language					
Configural	171 (40)	.994	.087			Configural	92 (38)	.998	.058		
Metric	235 (47)	.991	.096	.003	.009	Metric	142 (44)	.996	.072	.002	.014
Metricp ^a	222 (46)	.992	.094	.002	.007	Metricp ^a	126 (43)	.996	.067	.002	.005
Scalarp ^a	280 (69)	.990	.084	.002	.010	Scalarp ^a	181 (65)	.995	.064	.001	.003
Strictp ^a	348 (77)	.988	.090	.002	.006	Strictp ^a	248 (73)	.992	.074	.003	.010

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; df = degrees of freedom, free. Metricp = partial metric invariance, Scalarp = partial scalar invariance, Strictp = partial strict invariance.
^aHere we set the factor loadings of item 8 free.

feeling these emotions more than men in contexts that evoke self-stereotyping (Else-Quest et al., 2012).

In terms of the associations with external constructs, the GSQ-8 yielded good results. As expected, both subscales of the GSQ-8 correlated positively with PTSD symptoms (Shi et al., 2021), depressive symptoms (Kim et al., 2011), and self-blame (Duncan & Cacciatore, 2015), and negatively with life satisfaction (Bugay & Demir, 2011) and mental health-related quality of life (Owens et al., 2009; Persons et al., 2010). Importantly, both subscales further predicted PTSD symptoms, depression, life satisfaction, and mental health-related life quality over and above cognitions of self-blame. This is relevant because self-blame is an important factor in explaining variation in PTSD (Beck, Coffey et al., 2004; Foa et al., 1999). Further studies are needed to elucidate more nuanced and distinct contributions of guilt and shame measured with the GSQ-8.

The GSQ-8 has advantages over other guilt and shame scales. While being more parsimonious than most other scales, the GSQ-8 yields at least as good psychometric properties as well-established scales such as the PFQ-2, ASGS, ESS, SSGS and SSGS-8 (Lear et al., 2022). With its focus on experienced

guilt and shame over the past four weeks, the GSQ-8 proves useful in assessing these emotions as they are experienced over a longer period of time, which might be more useful when examining them in relation to psychopathology. As such, the use of the GSQ-8 can be recommended in both clinical and non-clinical contexts.

5.1. Limitations

The sample sizes of the included samples were somewhat small. The limited number of English-speaking participants precluded an examination of measurement invariance across all three languages. Consequently, measurement invariance across other languages than Dutch and German remains unknown and has to be formally tested in future research. Moreover, all three involved countries are Western countries precluding generalizations to non-Western populations. Similarly, samples were on average highly educated which may also undermine generalizability. Furthermore, the present study only assessed cross-sectional data. Thus, the test-retest reliability and longitudinal measurement invariance of the GSQ-8 remain to be investigated in future research. Also,

Table 5. Correlations of the GSQ-8 with validators and among the subscales.

	Clinical (N = 209)			Traumatized (N = 566)			Non-clinical (N = 156)		
	Total	Shame	Guilt	Total	Shame	Guilt	Total	Shame	Guilt
Total	–	.92***	.92***	–	.91***	.93***	–	.89***	.92***
Shame	–	–	.69***	–	–	.70***	–	–	.64***
Guilt	–	–	–	–	–	–	–	–	–
PCL-5	.58***	.53***	.53***	.66***	.62***	.60***	–	–	–
SWLS	–	–	–	–.53***	–.49***	–.48***	–	–	–
PTCI – self-blame	.63***	.60***	.56***	.46***	.42***	.43***	–	–	–
BDI	.64***	.60***	.58***	–	–	–	.60***	.51***	.57***
DASS – depression	–	–	–	.67***	.62***	.61***	–	–	–
MHQoL	.55***	.51***	.50***	–	–	–	–	–	–
Male gender ^a	–1.32	–0.52	–0.80	–2.16***	–0.86***	–1.31***	–1.26	–0.98*	–0.29

Note. PCL-5 = Posttraumatic Stress Disorder Checklist for DMS-5; SWLS = Satisfaction with Life Scale; PTCI = The Posttraumatic Cognitions Inventory; BDI = Beck's Depression Inventory; DASS = Depression Anxiety Stress Scales; MHQoL = Mental health-related quality of life * $p < .05$, ** $p < .01$, *** $p < .001$.
^aEstimated regression weight.

Table 6. Multiple regression weights when adjusting for cognitions of self-blame.

	Clinical (N = 209)				Traumatized (N = 566)			
	β	SE	p	R ²	β	SE	p	R ²
PCL-5								
Shame	1.26	0.38	<.001	.33	1.84	0.24	<.001	.47
Guilt	1.29	0.36	.001		1.32	0.21	<.001	
PTCI – self-blame	0.24	0.21	.263		2.10	0.39	<.001	
SWLS								
Shame	–	–	–		–0.71	0.13	<.001	.30
Guilt	–	–	–		–0.50	0.11	<.001	
PTCI – self-blame	–	–	–		–0.79	0.21	<.001	
BDI-II								
Shame	0.87	0.21	<.001	.41	–	–	–	
Guilt	0.71	0.19	<.001		–	–	–	
PTCI – self-blame	0.25	0.12	.035		–	–	–	
MHQoL								
Shame	0.24	0.08	.001	.30	–	–	–	
Guilt	0.22	0.07	.002		–	–	–	
PTCI – self-blame	0.04	0.04	.354		–	–	–	
DASS -depression								
Shame	–	–	–		0.58	0.07	<.001	.45
Guilt	–	–	–		0.46	0.06	<.001	
PTCI – self-blame	–	–	–		0.37	0.12	.002	

Note. PCL-5 = Posttraumatic Stress Disorder Checklist for DMS-5; SWLS = Satisfaction with Life Scale; PTCI – self-blame = The Posttraumatic Cognitions Inventory – self-blame subscale; BDI = Beck Depression Inventory-II; DASS = Depression Anxiety Stress Scales; MHQoL = Mental health-related quality of life.

no conclusions on concurrent and incremental validity can be drawn as we did not include other instruments assessing the frequency of guilt and shame. Moreover, the mean endorsement of guilt and shame items was fairly low particularly for the non-clinical and traumatized samples. However, the GSQ-8 had similar psychometric properties in all three samples including the clinical sample with higher mean-levels endorsement indicating that low mean endorsement did not undermine the validity of the GSQ-8 as a measure of guilt and shame. Lastly, no feedback on items was gathered from participants. For further item refinement, it would be beneficial to include people with lived experience in the process of the item formulation. We believe these limitations provide fertile ground for further validation studies on the GSQ-8.

6. Conclusion

Understanding the involvement of guilt and shame in complex intrapersonal and interpersonal regulatory processes more thoroughly is of great relevance. The GSQ-8 presents an important and psychometrically sound addition to the available body of measurement tools. Its broad acceptability in terms of its parsimony and its utility in clinical and non-clinical samples as well as its availability in English, Dutch and German point to a useful scale for the assessment of guilt and shame across a broad range of contexts.

Open Scholarship



This article has earned the [Center for Open Science](#) badges for Open Data, Open Materials and Preregistered.

The data and materials are openly accessible at https://osf.io/t3dh7/?view_only=442bdfeda8ec4a429c07dee864d6e442.

Note

- Overall, $n = 46$ participants received the English version of the GSQ-8. This number was too small to be included in our measurement invariance analyses across languages.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The English, Dutch and German versions of the GSQ-8 and the analysis code to reproduce our results are available on the open science framework (https://osf.io/t3dh7/?view_only=442bdfeda8ec4a429c07dee864d6e442).

ORCID

Thole H. Hoppen <http://orcid.org/0000-0002-6050-8696>

References

- Asparouhov, T., & Muthén, B. (2010). *Weighted Least Squares Estimation with Missing Data*. Muthén & Muthén.
- Bannister, J. A., Colvonen, P. J., Angkaw, A. C., & Norman, S. B. (2018). Differential relationships of guilt and shame on posttraumatic stress disorder among veterans. *Psychological Trauma: Theory, Research, Practice and Policy*, 11(1), 35-42. <https://doi.org/10.1037/tra0000392>
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck depression inventory-II*. Psychological Corporation.
- Beck, J. G., Coffey, S. F., Palyo, S. A., Gudmundsdottir, B., Miller, L. M., & Colder, C. R. (2004). Psychometric

- properties of the Posttraumatic Cognitions Inventory (PTCI): a replication with motor vehicle accident survivors. *Psychological Assessment*, 16(3), 289–298. <https://doi.org/10.1037/1040-3590.16.3.289>
- Beck, J. G., Reich, C. M., Woodward, M. J., Olsen, S. A., Jones, J. M., & Patton, S. C. (2015). How do negative emotions relate to dysfunctional posttrauma cognitions? An examination of interpersonal trauma survivors. *Psychological Trauma: Theory, Research, Practice, and Policy*, 7(1), 3–10. <https://doi.org/10.1037/a0032716>
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a Clinician-Administered PTSD Scale. *Journal of Traumatic Stress*, 8(1), 75–90. <https://doi.org/10.1002/jts.2490080106>
- Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. *Journal of Traumatic Stress*, 28(6), 489–498. <https://doi.org/10.1002/jts.22059>
- Boeschoten, M. A., van der Aa, N., Bakker, A., Heide, F. J. J. t., Hoofwijk, M. C., Jongedijk, R. A., van Minnen, A., Elzinga, B. M., & Olf, M. (2018). Development and evaluation of the Dutch clinician-administered PTSD scale for DSM-5 (CAPS-5). *European Journal of Psychotraumatology*, 9(1), 1546085. <https://doi.org/10.1080/20008198.2018.1546085>
- Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2016). Psychometric properties of the PTSD checklist for diagnostic and statistical manual of mental disorders—fifth edition (PCL-5) in Veterans. *Psychological Assessment*, 28(11), 1379–1391. <https://doi.org/10.1037/pas0000254>
- Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the depression anxiety stress scales (DASS) in clinical samples. *Behaviour Research and Therapy*, 1(35), 79–89. [https://doi.org/10.1016/S0005-7967\(96\)00068-X](https://doi.org/10.1016/S0005-7967(96)00068-X)
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2), 230–258. <https://doi.org/10.1177/0049124192021002005>
- Bugay, A., & Demir, A. (2011). Psychometric properties of the Turkish version of trait shame and guilt scale. *Eurasian Journal of Educational Research*, 11(45), 19–32.
- Byrne, B. M., Shavelson, R. J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105(3), 456–466. <https://doi.org/10.1037/0033-2909.105.3.456>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Crawford, J. R., & Henry, J. D. (2003). The Depression Anxiety Stress Scales (DASS): normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology*, 42(2), 111–131. <https://doi.org/10.1348/014466503321903544>
- Cunningham, K. C. (2020). Shame and guilt in PTSD. In M. T. Tull & N. A. Kimbrel (Eds.), *Emotion in posttraumatic stress disorder* (pp. 145–171). Academic Press. <https://doi.org/10.1016/B978-0-12-816022-0.00006-5>
- DeCou, C. R., Cole, T. T., Lynch, S. M., Wong, M. M., & Matthews, K. C. (2016). Assault-related shame mediates the association between negative social reactions to disclosure of sexual assault and psychological distress. *Psychological Trauma: Theory, Research, Practice and Policy*, 9(2), 166–172. <https://doi.org/10.1037/tra0000186>
- DeCou, C. R., Lynch, S. M., Weber, S., Richner, D., Mozafari, A., Huggins, H., & Perschon, B. (in print). On the association between trauma-related shame and symptoms of psychopathology: A meta-analysis. *Trauma, Violence, & Abuse*, 15248380211053617. <https://doi.org/10.1177/15248380211053617>
- Dempsey, H. L. (2017). A comparison of the social-adaptive perspective and functionalist perspective on guilt and shame. *Behavioral Sciences*, 7(4), 83. <https://doi.org/10.3390/bs7040083>
- Di Sarno, M., Di Pierro, R., & Madeddu, F. (2022). Shame and guilt-proneness in an Italian sample: Latent structure and gender invariance of the personal feelings questionnaire-2 (PFQ-2). *Current Psychology*, 41, 276–288. <https://doi.org/10.1007/s12144-019-00570-w>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13
- Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment*, 10(2), 83–89. <https://doi.org/10.1037/1040-3590.10.2.83>
- Duncan, C., & Cacciatore, J. (2015). A systematic review of the peer-reviewed literature on self-blame, guilt, and shame. *OMEGA - Journal of Death and Dying*, 71(4), 312–342. <https://doi.org/10.1177/0030222815572604>
- Eekers, D. B. P., Di Perri, D., Roelofs, E., Postma, A., Dijkstra, J., Ajithkumar, T., Alapetite, C., Blomstrand, M., G Burnet, N., & Calugaru, V. (2021). Update of the EPTN atlas for CT- and MR-based contouring in neuro-oncology. *Radiotherapy and Oncology*, 160, 259–265. <https://doi.org/10.1016/j.radonc.2021.05.013>
- Else-Quest, N. M., Higgins, A., Allison, C., & Morton, L. C. (2012). Gender differences in self-conscious emotional experience: A meta-analysis. *Psychological Bulletin*, 138(5), 947–981. <https://doi.org/10.1037/a0027930>
- Eterović, M., Medved, V., Bilić, V., Kozarić-Kovačić, D., & Žarković, N. (2020). Poor agreement between two commonly used measures of shame and guilt-proneness. *Journal of Personality Assessment*, 102(4), 499–507. <https://doi.org/10.1080/00223891.2019.1585361>
- Ferguson, T. J., & Crowley, S. L. (1997). Measure for measure: A multitrait-multimethod analysis of guilt and shame. *Journal of Personality Assessment*, 69(2), 425–441. https://doi.org/10.1207/s15327752jpa6902_12
- Fletcher, K. E. (2011). Understanding and assessing traumatic responses of guilt, shame, and anger among children, adolescents, and young adults. *Journal of Child & Adolescent Trauma*, 4(4), 339–360. <https://doi.org/10.1080/19361521.2011.623146>
- Foa, E. B., Ehlers, A., & Clark, D. M. (1999). The Posttraumatic Cognitions Inventory (PTCI): development and validation. *Psychological Assessment*, 11(3), 303–314. <https://doi.org/10.1037/1040-3590.11.3.303>
- Ginzburg, K., Butler, L. D., Giese-Davis, J., Cavanaugh, C. E., Neri, E., Koopman, C., Classen, C. C., & Spiegel, D. (2009). Shame, guilt, and posttraumatic stress disorder in adult survivors of childhood sexual abuse at risk for human immunodeficiency virus: Outcomes of a randomized clinical trial of group psychotherapy treatment. *Journal of Nervous and Mental Disease*, 197(7), 536–542. <https://doi.org/10.1097/NMD.0b013e3181ab2ebd>
- Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the Satisfaction With Life Scale

- (SWLS): psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment*, 27(2), 127–132. <https://doi.org/10.1027/1015-5759/a000058>
- Guenole, N., & Brown, A. (2014). The consequences of ignoring measurement invariance for path coefficients in structural equation models. *Frontiers in Psychology*, 5, 980. <https://doi.org/10.3389/fpsyg.2014.00980>
- Harder, D. W., Cutler, L., & Rockart, L. (1992). Assessment of shame and guilt and their relationships to psychopathology. *Journal of Personality Assessment*, 59(3), 584–604. https://doi.org/10.1207/s15327752jpa5903_12
- Harder, D. W., & Greenwald, D. F. (1999). Further validation of the shame and guilt scales of the Harder Personal Feelings Questionnaire-2. *Psychological Reports*, 85(1), 271–281. <https://doi.org/10.2466/pr0.1999.85.1.271>
- Harder, D. W., & Lewis, S. J. (1987). The assessment of shame and guilt. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in Personality Assessment*, Vol. 6, (pp. 89–114). Lawrence Erlbaum Associates, Inc.
- Harder, D. W., & Zalma, A. (1990). Two promising shame and guilt scales: A construct validity comparison. *Journal of Personality Assessment*, 55(3-4), 729–745. <https://doi.org/10.1080/00223891.1990.9674108>
- Held, P., Owens, G. P., & Anderson, S. E. (2015). The interrelationships among trauma-related guilt and shame, disengagement coping, and PTSD in a sample of treatment-seeking substance users. *Traumatology*, 21(4), 285–292. <https://doi.org/10.1037/trm0000050>
- Hoblitzelle, W. (1982). *Developing a measure of shame and guilt and the role of shame in depression* [Unpublished doctoral dissertation]. Yale University, New Haven, CT.
- Hoppen, T. H., & Morina, N. (2021). I'd do much better, if only ... counterfactual comparisons related to traumatic life events. *Cognition and Emotion*, 35(2), 409–416. <https://doi.org/10.1080/02699931.2020.1832967>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kim, S., Thibodeau, R., & Jorgensen, R. S. (2011). Shame, guilt, and depressive symptoms: A meta-analytic review. *Psychological Bulletin*, 137(1), 68–96. <https://doi.org/10.1037/a0021466>
- Krüger-Gottschalk, A., Knaevelsrud, C., Rau, H., Dyer, A., Schäfer, I., Schellong, J., & Ehring, T. (2017). The German version of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): psychometric properties and diagnostic utility. *BMC Psychiatry*, 17(1), 1–9. <https://doi.org/10.1186/s12888-017-1541-6>
- Leach, C. W., & Cidam, A. (2015). When is shame linked to constructive approach orientation? A meta-analysis. *Journal of Personality and Social Psychology*, 109(6), 983–1002. <https://doi.org/10.1037/pspa0000037>
- Lear, M. K., Lee, E. B., Smith, S. M., & Luoma, J. B. (2022). A systematic review of self-report measures of generalized shame. *Journal of Clinical Psychology*, 78(7), 1288–1330. <https://doi.org/10.1002/jclp.23311>
- Lee, D. A., Scragg, P., & Turner, S. (2001). The role of shame and guilt in traumatic events: A clinical model of shame-based and guilt-based PTSD. *British Journal of Medical Psychology*, 74(4), 451–466. <https://doi.org/10.1348/0007112011161109>
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>
- Muris, P., & Meesters, C. (2014). Small or big in the eyes of the other: On the developmental psychopathology of self-conscious emotions as shame, guilt, and pride. *Clinical Child and Family Psychology Review*, 1(17), 19–40. <https://doi.org/10.1007/s10567-013-0137-z>
- Müller, J., Wessa, M., Rabe, S., Dörfel, D., Knaevelsrud, C., Flor, H., Maercker, A., & Karl, A. (2010). Psychometric properties of the posttraumatic cognitions inventory (PTCI) in a German sample of individuals with a history of trauma. *Psychological Trauma: Theory, Research, Practice, and Policy*, 2(2), 116–125. <https://doi.org/10.1037/a0018603>
- Owens, G. P., Steger, M. F., Whitesell, A. A., & Herrera, C. J. (2009). Posttraumatic stress disorder, guilt, depression, and meaning in life among military veterans. *Journal of Traumatic Stress*, 22(6), 654–657. <https://doi.org/10.1002/jts.20460>
- Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment*, 5(2), 164–172. <https://doi.org/10.1037/1040-3590.5.2.164>
- Persons, E., Kershaw, T., Sikkema, K. J., & Hansen, N. B. (2010). The impact of shame on health-related quality of life among HIV-positive adults with a history of childhood sexual abuse. *AIDS Patient Care and STDs*, 24(9), 571–580. <https://doi.org/10.1089/apc.2009.0209>
- Pugh, L. R., Taylor, P. J., & Berry, K. (2015). The role of guilt in the development of post-traumatic stress disorder: A systematic review. *Journal of Affective Disorders*, 182, 138–150. <https://doi.org/10.1016/j.jad.2015.04.026>
- R Core Team. (2013). *R: A Language and Environment for Statistical Computing* [Computer Software]. R Foundation for Statistical Computing.
- Rice, S. M., Kealy, D., Treeby, M. S., Ferlatte, O., Oliffe, J. L., & Ogrodniczuk, J. S. (2018). Male guilt—and shame-proneness: The Personal Feelings Questionnaire (PFQ-2 brief). *Archives of Psychiatry and Psychotherapy*, 2(2), 46–54. <https://doi.org/10.12740/APP/86488>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. Version 0.5–12 (BETA). *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Shi, C., Ren, Z., Zhao, C., Zhang, T., & Chan, SH-W. (2021). Shame, guilt, and posttraumatic stress symptoms: A three-level meta-analysis. *Journal of Anxiety Disorders*, 82, 102443. <https://doi.org/10.1016/j.janxdis.2021.102443>
- Street, A. E., & Arias, I. (2001). Psychological abuse and posttraumatic stress disorder in battered women: Examining the roles of shame and guilt. *Violence and Victims*, 16(1), 65–78. <https://doi.org/10.1891/0886-6708.16.1.65>
- Tangney, J. P., & Dearing, R. L. (2003). *Shame and Guilt*. Guilford Press.
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58(1), 345–372. <https://doi.org/10.1146/annurev.psych.56.091103.070145>
- Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology*, 101(3), 469–478. <https://doi.org/10.1037/0021-843X.101.3.469>

- Tignor, S. M., & Colvin, C. R. (2016). The interpersonal adaptiveness of dispositional guilt and shame: A meta-analytic investigation. *Journal of Personality, 85*(3), 341–363. <https://doi.org/10.1111/jopy.12244>
- Tracy, J. L., & Robins, R. W. (2006). Appraisal antecedents of shame and guilt: Support for a theoretical model. *Personality & Social Psychology Bulletin, 32*(10), 1339–1351. <https://doi.org/10.1177/0146167206290212>
- van Emmerik, A. A., Schoorl, M., Emmelkamp, P. M., & Kamphuis, J. H. (2006). Psychometric evaluation of the Dutch version of the posttraumatic cognitions inventory (PTCI). *Behaviour Research and Therapy, 44*(7), 1053–1065. <https://doi.org/10.1016/j.brat.2005.07.002>
- van Krugten, F. C., Busschbach, J. J., Versteegh, M. M., Hakkaart-van Roijen, L., & Brouwer, W. B. (2022). The Mental Health Quality of Life Questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Quality of Life Research, 31*(2), 633–643. <https://doi.org/10.1007/s11136-021-02935-w>
- Vizin, G., Urbán, R., & Unoka, Z. (2016). Shame, trauma, temperament and psychopathology: Construct validity of the Experience of Shame Scale. *Psychiatry Research, 246*, 62–69. <https://doi.org/10.1016/j.psychres.2016.09.017>
- Wang, Y-P, & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Brazilian Journal of Psychiatry, 35*(4), 416–431. <https://doi.org/10.1590/1516-4446-2012-1048>
- Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD Checklist for DSM-5 (PCL-5). National Center for PTSD: www.ptsd.va.gov.
- Wibbelink, C. J. M., Lee, C. W., Bachrach, N., Dominguez, S. K., Ehring, T., van Es, S. M., Fassbinder, E., Köhne, S., Mascini, M., Meewisse, M-L, Menninga, S., Morina, N., Rameckers, S. A., Thomaes, K., Walton, C. J., Wigard, I. G., & Arntz, A. (2021). The effect of twice-weekly versus once-weekly sessions of either imagery rescripting or eye movement desensitization and reprocessing for adults with PTSD from childhood trauma (IREM-Freq): a study protocol for an international randomized clinical trial. *Trials, 22*(1), <https://doi.org/10.1186/s13063-021-05712-9>
- Wilson, J. P., Droždek, B., & Turkovic, S. (2006). Posttraumatic shame and guilt. *Trauma, Violence, & Abuse, 7*(2), 122–141. <https://doi.org/10.1177/1524838005285914>
- Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., Foa, E. B., Young-McCaughan, S., Yarvis, J. S., & Hembree, E. A. (2016). Psychometric analysis of the PTSD checklist-5 (PCL-5) among treatment-seeking military service members. *Psychological Assessment, 28*(11), 1392–1403. <https://doi.org/10.1037/pas0000260>