



Measuring Moral Injury: Further Validation of the MIES-C and EMIS-C in a Civilian Population

Margaux Morriss¹ · David Berle^{1,2}

Accepted: 28 June 2023 / Published online: 19 July 2023
© The Author(s) 2023

Abstract

Moral injury among civilians has garnered increased recognition in recent years. However, most moral injury measures have focused on military experiences, with language being military specific. This study aimed to further validate two self-report measures of civilian moral injury, the Moral Injury Events Scale - Civilian (MIES-C) and Expressions of Moral Injury Scale - Civilian (EMIS-C). Participants were 312 adults (99 males, 212 females, *Mage* = 37.24) who completed the MIES-C, EMIS-C, and a battery of associated measures on two occasions. For the MIES-C, four Confirmatory Factor Analysis (CFA) models were examined. A three-factor model, with covaried items, provided the best fit to the data. For the EMIS-C, three CFAs were examined. A two-factor model, with covaried items, and the bi-factor model provided the best fit to the data, with the two-factor model being preferred for reasons of parsimony and its absence of weak and negative item loadings. Both measures demonstrated acceptable test-retest reliability and convergent validity with associated measures of psychological distress. This study provides psychometrically sound tools for clinicians and researchers of civilian moral injury.

Keywords Moral injury · Trauma · Civilian · General population · Measurement · Psychometrics

Exposure to traumatic events can lead to debilitating psychological consequences. Our understanding of adjustment to traumas where one's life or physical safety is at stake has informed the diagnosis and treatment of Posttraumatic Stress Disorder (PTSD). However, there is an increasing acknowledgement that individuals can experience significant stressors where the threat is not necessarily an external threat to one's physical safety, but instead, a symbolic threat to one's integrity and existential sense of themselves and the world (Litz et al., 2009). Such experiences have been described as “moral injuries”, which can be viewed as a unique form of psychological suffering that results from transgressions of deeply held ethical and moral beliefs (Nash et al., 2013). These transgressions can involve “perpetrating, failing to prevent, bearing witness to, or learning

about acts that transgress deeply held moral beliefs” (Litz et al., 2009, p. 700). Its primary differentiation from fear-based responses is the symbolic threat to one's integrity as opposed to the external threat to one's physical safety (Litz et al., 2009).

Moral injury was first conceptualised in a military context, in order to understand the profound and persisting harms experienced by many military personnel and veterans. The concept appears to date back to Ancient Greek tragedies; however, the clinical term was first adopted by Shay (1994) following his work with U.S. veterans. Given the nature of combat and military-related experiences, military personnel can be exposed to traumatic events that violate their moral beliefs (Griffin et al., 2019). For example, killing enemy combatants, failing to prevent the suffering of fellow personnel, or betrayal by a trusted authority (Griffin et al., 2019; Shay, 2014). Exposure to these events can increase one's likelihood of developing moral injury, though it does not guarantee adverse consequences (Griffin et al., 2019).

The injury associated with these transgressions can manifest through feelings of guilt, shame, betrayal, anger, frustration, and sadness (Litz et al., 2009). These experiences

✉ David Berle
david.berle@anu.edu.au

¹ Graduate School of Health, University of Technology
Sydney, Sydney, Australia

² School of Medicine and Psychology, The Australian National
University, Canberra, Australia

can lead to depression, anxiety, social withdrawal, demoralisation, distrust, existential conflict, religious/spiritual distress, negative views of the self, sleep difficulties, substance use, self-harm, and suicidality (Barnes et al., 2019; Griffin et al., 2019; Jamieson et al., 2020). Further, moral injury is often associated with PTSD due to its basis in trauma exposure. Moral injury increases the likelihood of developing PTSD and is associated with increased PTSD symptom severity (Griffin et al., 2019). While the two concepts are related, there are key conceptual differences. The emotions associated with moral injury are typically those developed after the traumatic event; however, the emotions associated with PTSD are typically those experienced during the event (Barnes et al., 2019). The experience of shame and guilt, as opposed to fear, is central to moral injury. To ameliorate this injury, clinicians and researchers must first be able to quantify and measure the construct.

The instruments used to measure moral injury have almost exclusively been designed for military personnel. For instance, the language of many measures refers to ‘*the military experience*’ (Currier et al., 2018; Nash et al., 2013). This has allowed for valid and reliable measurement of moral injury among military-related populations. Examples include the Moral Injury Events Scale (MIES; Nash et al., 2013) and the Expressions of Moral Injury Scale – Military (EMIS-M; Currier et al., 2018). The MIES assesses for the occurrence of causes and associated symptoms, while the EMIS-M focuses on the possible outcomes that may arise following moral injury. Both the MIES and EMIS-M were developed using a rational, iterative process whereby experts generated pools of items which were then refined following review of empirical, clinical and theoretical sources in consultation with subject matter experts (i.e., EMIS-M; Currier et al., 2017) or via consensus (i.e., MIES; Nash et al., 2013). The MIES has demonstrated strong internal consistency (Cronbach’s $\alpha = 0.90$), temporal stability, construct validity, and discriminant validity (Bryan et al., 2016). Similarly, the EMIS-M has demonstrated strong internal consistency ($\alpha = 0.94$), test-retest reliability ($r = 0.80$), and convergent validity (Koenig et al., 2019). Both measures are positively correlated with measures of psychological distress and negatively correlated with measures of psychological wellbeing (Koenig et al., 2019). Both measures are popular among studies of military populations due to their brevity, strong psychometric properties, and ability to screen for possible treatment targets (Koenig et al., 2019).

Expanding to Non-Military Settings

While the moral injury experience resonates strongly with many military personnel, there is increasing acknowledgment that this suffering can be experienced by civilians (Griffin et al., 2019). So far, researchers have identified the presence of moral injury in various civilian populations. These include first responders (Lentz et al., 2021; Papazoglou & Chopko, 2017), correctional workers (Carleton et al., 2019), journalists (Feinstein et al., 2018), educators (Sugrue, 2019), veterinarians (Crane et al., 2015), healthcare professionals (Cartolovni et al., 2021; Mantri et al., 2020), and refugees (Hoffman et al., 2018; Nickerson et al., 2015). The associated symptoms appear to be similar to military populations, including feelings of guilt, shame, betrayal, distrust, and social withdrawal. Experiences include refugees leaving family members in war-torn countries, and police officers adhering to departmental policies which may be incongruent with personal beliefs. These findings suggest that the potential for moral injury is universal.

In non-military contexts, the term moral injury has at times been used almost interchangeably with terms such as “burnout” (e.g., Kopacz et al., 2019). However, moral injury is distinct from burnout in that it does not only arise in occupational context and is focused more on what happens to a person rather than the person’s coping resources (Dean et al., 2019). As the existing measures refer to the military in their wording, providing these measures to other populations would lead to difficulties with the reliability, validity, and interpretation of these results. To resolve this difficulty, researchers have used ad-hoc approaches to measuring civilian moral injury. Some researchers have created their own measures, for example the Moral Injury Symptom Scale – Healthcare Professionals (Mantri et al., 2020). These appear to remain setting-specific; thus, reducing their utility in other contexts. Other researchers have modified prominent military-specific measures, either by excluding certain items or altering the wording. For example, Feinstein et al. (2018) excluded the last three items in the MIES. While this was helpful in capturing moral injury among journalists, the tool was not validated. The attempts to overcome this barrier in civilian moral injury measurement remain context-specific and lack comprehensive validation efforts. This highlights the need for a validated measure of moral injury that is appropriate for all individuals regardless of setting.

Recently, Thomas et al. (in press) adapted the MIES (Nash et al., 2013) and the EMIS-M (Currier et al., 2018) to become the Moral Injury Events Scale – Civilian (MIES-C) and the Expressions of Moral Injury Scale – Civilian (EMIS-C). The adaptations focused on incorporating language that is generalisable to all individuals and involved changes to item content for two of the nine MIES items and

to four of the 17 items of the EMIS-M as well as minor changes to the instructions of each. For example, changing “fellow service members” to “friends” (MIES-C; item 8). Likewise, generalised language was used in revising items of the EMIS-C, for example by changing “*My military experiences have taught me that it is only a matter of time before people will betray my trust.*” was replaced with “*My experiences have taught me that it is only a matter of time before people will betray my trust*” (EMIS-C; item 3). This work was a start to bridging the measurement gap in the civilian moral injury literature base. The researchers completed preliminary validation analyses, with results indicating sufficient convergent validity, divergent validity, and factor structures. Yet, some questions remain regarding the scales’ psychometric properties. Test-retest reliability is important when considering the clinical utility of a scale, as it helps to ensure that responses to items are likely to be consistent across time (Aldridge et al., 2017). In turn, it ensures that changes in scores reflect real changes in the individual’s moral injury experience. The ability to measure civilian moral injury will allow researchers to better understand the construct of moral injury, providing more confidence in the reliability and validity of their results. Further, increased understanding of this construct may enable future clinicians to make better informed treatment planning and intervention decisions.

The present study explores the measurement of moral injury within non-military related settings. It aims to further validate two brief measures of moral injury that can be used in civilian settings. It was hypothesised that factor structures in the current study will replicate those found in Thomas et al. (in press). Specifically, Confirmatory Factor Analyses (CFAs) will identify a three-factor model for the MIES-C, and a two-factor model for the EMIS-C. It was also hypothesised that factors will demonstrate convergent validity, evident in medium to strong associations with similar constructs. Further, it was hypothesised that scores on each of the factors will remain stable across a two-week interval.

Method

Participants

Participants were recruited through the online recruitment platform, Prolific Academic, chosen for its large participant pool and demographically diverse users (Palan & Schitter, 2018). Of the 318 participants who consented to the study, 312 (98.1%) completed demographic information and questionnaires. Imputation approaches were not used for the six participants who did not provide data, as: (1) data-points were not missing at random or completely at random,

and (2) the small amount of missing data (1.9% of participants) unlikely resulted in bias in our results. Of the 312 participants who comprised the initial data collection, 291 individuals returned to complete the follow-up assessment (93% retention rate). Participants were included if they were adults and spoke English. Based on a Monte Carlo simulation study (Wolf et al., 2013), a sample size of 200+ would provide sufficient statistical power for a two- or three-factor model with three indicator variables per factor, and with individual factor loadings ≥ 0.65 . Thus, the current sample was sufficiently powered.

Participants were offered financial compensation for their participation, provided they completed the study. They were paid GBP £6.00 per hour, up to a maximum of GBP £7.70.

Measures

Moral Injury Events Scale – Civilian (MIES-C)

The MIES-C (Thomas et al., in press) is a 9-item measure assessing possible causes and symptoms of moral injury among civilians. Items are rated on a 6-point Likert scale (1 = *Strongly disagree* to 6 = *Strongly agree*). Various instructions and items were modified to remove any reference to the military. For example, “I feel betrayed by fellow service members who I once trusted” was changed to “I feel betrayed by friends who I once trusted”. Factor validation showed three dimensions of moral injury; perceived transgressions by the self (Transgressions-Self), perceived transgressions by others (Transgressions-Other), and perceived betrayals by others (Betrayal). In the adaptation study, internal consistency across the subscales was fair to excellent; Transgressions-Self $\alpha = 0.89$, Transgressions-Other ($\alpha = 0.71$), and Betrayal ($\alpha = 0.82$). The study also demonstrated strong construct validity and positive associations with related constructs (Thomas et al., in press). In the current study, internal consistency was fair to excellent; Transgressions-Self ($\alpha = 0.92$), Transgressions-Other ($\alpha = 0.72$), and Betrayal ($\alpha = 0.81$). Internal consistency of the total score was good ($\alpha = 0.88$).

Expressions of Moral Injury Scale – Civilian (EMIS-C)

The EMIS-C (Thomas et al., in press) is a 17-item measure designed to assess moral injury-related outcomes for civilians. Items capture an individual’s beliefs, emotions, and behaviours arising from moral injury. Items are rated on a 5-point scale (1 = *Strongly disagree* to 5 = *Strongly agree*). Instructions and items were adapted to remove reference to the military. An example item adaptation is “I am ashamed of myself because of things that I have seen or done” instead of “I am ashamed of myself because of things that I did/

saw during my military service”. Factor analyses highlighted two dimensions: self-directed, and other-directed moral injury. It has demonstrated strong construct validity and good internal consistency; Self-Directed ($\alpha = 0.81$), and Other-Directed ($\alpha = 0.80$; Thomas et al., *in press*). Internal consistency in the current sample was found to range from good to excellent; Self-Directed ($\alpha = 0.90$), and Other-Directed ($\alpha = 0.89$). Internal consistency of the total score was excellent ($\alpha = 0.93$).

Life Events Checklist (LEC)

The LEC (Weathers, Blake et al., 2013) is a 17-item measure that assesses lifetime exposure to potentially traumatic events. Items are rated on a 6-point scale, with responses including *Happened to me*, *Witnessed it*, *Learned about it*, *Part of my job*, *Not sure*, and *Doesn't apply*. The scale has demonstrated good convergence with established trauma exposure measures and strong correlations with measures of psychological distress (Gray et al., 2004). The LEC was chosen to assist in characterising the sample.

PTSD Checklist for DSM-5 (PCL-5)

The PCL-5 (Weathers, Litz et al., 2013) is a 20-item measure designed to assess PTSD criteria, where higher scores indicate a greater presence of PTSD symptoms. Items are rated on a 5-point scale (0 = *Not at all* to 4 = *Extremely*). The scale has demonstrated strong internal consistency ($\alpha = 0.94$), test-retest reliability, and convergent and discriminant validity (Blevins et al., 2015; Ibrahim et al., 2018). Internal consistency in the current sample was excellent ($\alpha = 0.96$).

International Trauma Questionnaire (ITQ)

The ITQ (Cloitre et al., 2018) is a measure that assesses the presence and associated impairment of PTSD and complex PTSD symptoms, from the International Classification of Diseases-11 (World Health Organisation, 2019). The scale comprises of nine items assessing PTSD-specific symptoms, and nine items assessing CPTSD-specific symptoms. Items are rated on a 5-point scale (0 = *Not at all* to 4 = *Extremely*), where higher scores indicate a greater presence/impact. The questionnaire has shown sufficient internal consistency ($\alpha = 0.90$) and construct validity (Sele et al., 2020). Internal consistency in the current sample was excellent ($\alpha = 0.94$ for both the PTSD and CPTSD subscales).

Patient Health Questionnaire (PHQ-9)

The PHQ-9 (Kroenke et al., 2001) is a 9-item measure that assesses the severity of depressive symptoms. Items are

rated on a 4-point scale (0 = *Not at all* to 3 = *Nearly every day*), where higher scores indicate a greater presence of symptoms. The scale has been found to have strong internal consistency ($\alpha = 0.86$ to 0.90), test-retest reliability, construct validity, (Kroenke et al., 2001; Sun et al., 2020). Internal consistency in the current sample was excellent ($\alpha = 0.92$).

Generalised Anxiety Disorder Assessment (GAD-7)

The GAD-7 (Spitzer et al., 2006) is a 7-item measure of generalised anxiety disorder symptoms. Items are rated on a 4-point scale (0 = *Not at all* to 3 = *Nearly every day*). The GAD-7 has demonstrated strong construct validity, criterion validity, and internal consistency ($\alpha = 0.92$; Dhira et al., 2021; Spitzer et al., 2006). Internal consistency in the current sample was excellent ($\alpha = 0.93$).

Dimensions of Anger Reactions (DAR-5)

The DAR-5 (Forbes et al., 2014) is a 5-item measure designed to assess anger experiences. Items are rated on a 5-point scale (1 = *None or almost none of the time* to 5 = *All or almost all of the time*), where higher scores indicate greater levels of anger. The scale has demonstrated good internal consistency ($\alpha = 0.80$ to 0.90), convergent validity, and discriminant validity (Forbes et al., 2014; Goulart et al., 2021). Internal consistency in this sample was good ($\alpha = 0.88$).

Procedure

Ethics approval was received from the University of Technology Sydney Human Research Ethics Committee (ETH216541). Interested individuals were directed to Qualtrics where they read the study's Information and Consent Form. Consenting participants were able to revoke their consent at any time, with knowledge that their responses would not be used. The study was conducted across two time points: initial assessment and follow-up assessment. These time points occurred two weeks apart, to allow for investigation of test-retest reliability without risking excessive participant attrition. For the initial assessment, participants were asked to complete demographics questions, the MIES-C, EMIS-C, and all other self-report questionnaires. Participants from the initial study were then invited to complete the follow-up assessment via Prolific messaging. For the follow-up, participants were asked to complete the MIES-C and EMIS-C.

Data Analysis

Initial analyses were conducted using IBM SPSS Statistics Version 26. Data were merged, screened, and cleaned prior to any analyses. No missing values were detected. Next, CFAs were conducted to examine the factor structure of

Table 1 Demographic Characteristics (N=312)

Variable	<i>n</i>	%
Gender		
<i>Male</i>	99	31.7
<i>Female</i>	212	67.9
<i>Other</i>	1	0.3
Racial Background		
<i>Asian</i>	17	5.4
<i>African American</i>	17	5.4
<i>Caucasian</i>	239	76.6
<i>Multiracial/Biracial</i>	12	3.8
<i>Other</i>	27	8.7
Country of Residence		
<i>United Kingdom</i>	160	51.3
<i>United States of America</i>	18	5.8
<i>Canada</i>	34	10.9
<i>Australia</i>	25	8.0
<i>Ireland</i>	14	4.5
<i>Other</i>	61	19.6
Employment Status		
<i>Employed</i>	239	76.6
<i>Unemployed</i>	73	23.4
Relationship Status		
<i>In a relationship/Married/De Facto</i>	143	45.8
Variable	<i>n</i>	%
<i>Single/Widowed</i>	169	54.2
Lifetime Exposure to Trauma		
<i>Endorsement of 1 trauma exposure</i>	303	97.1
<i>Endorsement of > 1 trauma exposure</i>	291	93.3
<i>Endorsement of ≥ 3 trauma exposures</i>	282	90.4
Diagnosis of Mental Health Disorder		
<i>Yes</i>	96	30.8
<i>No</i>	213	68.3
<i>Prefer Not to Say</i>	3	1.0
Currently Engaged with Mental Health Support/Treatment		
<i>Yes</i>	55	17.6
<i>No</i>	257	82.4
Previously Engaged with Mental Health Support/Treatment		
<i>Yes</i>	122	39.1
<i>No</i>	190	60.9
Current Medication Use (Related to Mental Health)		
<i>Yes</i>	52	16.7
<i>No</i>	260	83.3
	<i>M</i>	<i>SD</i>
Age	37.24	13.98

Note. Data for age was only available for 310 participants

the MIES-C and EMIS-C. Factor models were examined in Mplus Version 8.3. All analyses in Mplus were based on polychoric correlations.

Four models were investigated for the MIES-C: (1) a single-factor model, to investigate the possibility of a unifying construct, (2) a two-factor model, consistent with the original MIES evaluation (Nash et al., 2013), (3) a three-factor model, consistent with Thomas et al. (in press), and (4) a bi-factor model, to investigate the possibility of a simultaneous uni- and multi-dimensional structure. For the three-factor model, Thomas et al. (in press) allowed item pairs 3/4 and 5/6 to covary, to improve model fit. This was replicated in the current study. Three models were investigated for the EMIS-C: (1) a single-factor model, (2) a two-factor model, consistent with Thomas et al. (in press) and the original EMIS-M evaluation (Currier et al., 2018), and (3) a bi-factor model. For the two-factor model, Thomas et al. (in press) allowed item pairs 1/7 and 8/14 to covary, which was replicated. It is noted that bi-factor models were not included in Thomas et al. (in press); however, they can provide clearer understanding of psychological constructs (Bornoalova et al., 2020).

A CFA approach was chosen due to the hypothesis-driven nature of the study and the established factor models found in the literature. Standard model fit criteria were used. Root Mean Square Error of Approximation (RMSEA) values ≤ 0.08 indicated acceptable fit, while values ≤ 0.06 indicated excellent fit (Hu & Bentler, 1999). Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values ≥ 0.90 indicated acceptable fit, while values ≥ 0.95 indicated excellent fit (Hu & Bentler, 1999). Internal consistency of each scale and subscale was assessed using Cronbach's alpha. Construct validity was assessed by examining correlations (Pearson's *r*). Test-retest reliability was assessed by examining correlations (Pearson's *r*) between initial and follow-up assessment scores. No post-hoc analyses were conducted. The analysis and reporting of this study adheres to COSMIN guidelines (Gagnier et al., 2021). See Supplementary Tables 1 for the COSMIN reporting checklist.

Results

Descriptive Statistics

The demographic and trauma-related characteristics of the sample are summarised in Table 1. Participants included 99 males (31.7%), 212 females (67.9%), and one individual who endorsed Other (0.3%). Participants were aged between 19 and 89 years ($M = 37.24$ years, $SD = 13.98$), and identified as Caucasian (76.6%), Asian (5.4%), African American (5.4%), Multiracial (3.8%), and Other (8.7%).

Confirmatory Factor Analyses

MIES-C

Goodness of fit indices suggest that the single-factor model did not fit the data ($\chi^2 = 722.99$, $df=27$, $p < 0.001$, RMSEA=0.29, CFI=0.91, TLI=0.88). Similarly, the two-factor model, consistent with Nash et al. (2013), did not fit the data ($\chi^2 = 420.64$, $df=26$, $p < 0.001$, RMSEA=0.22, CFI=0.95, TLI=0.93). The three-factor model, with covaried items 3/4 and 5/6, provided an excellent fit ($\chi^2 = 50.50$, $df=22$, $p < 0.001$, RMSEA=0.06, CFI=0.99, TLI=0.99). These results are consistent with Thomas et al. (in press).

Item factor loadings are presented in Fig. 1. The results from the bi-factor CFA indicated a *not positive definite matrix*; therefore, the results are not able to be presented. The bi-factor model did not fit the current data. The CFA results for each model are presented in Table 2.

Correlations among the subscales were examined. The strongest correlation was observed between Transgressions-Other and Betrayal ($r=0.56$). A moderate correlation was found between Transgressions-Self and Transgressions-Other ($r=0.52$), and the weakest correlation was found between Transgressions-Self and Betrayal ($r=0.41$). All correlations were significant at the 0.01 level.

Fig. 1 Standardised Item Factor Loadings for the MIES-C Three-Factor Model

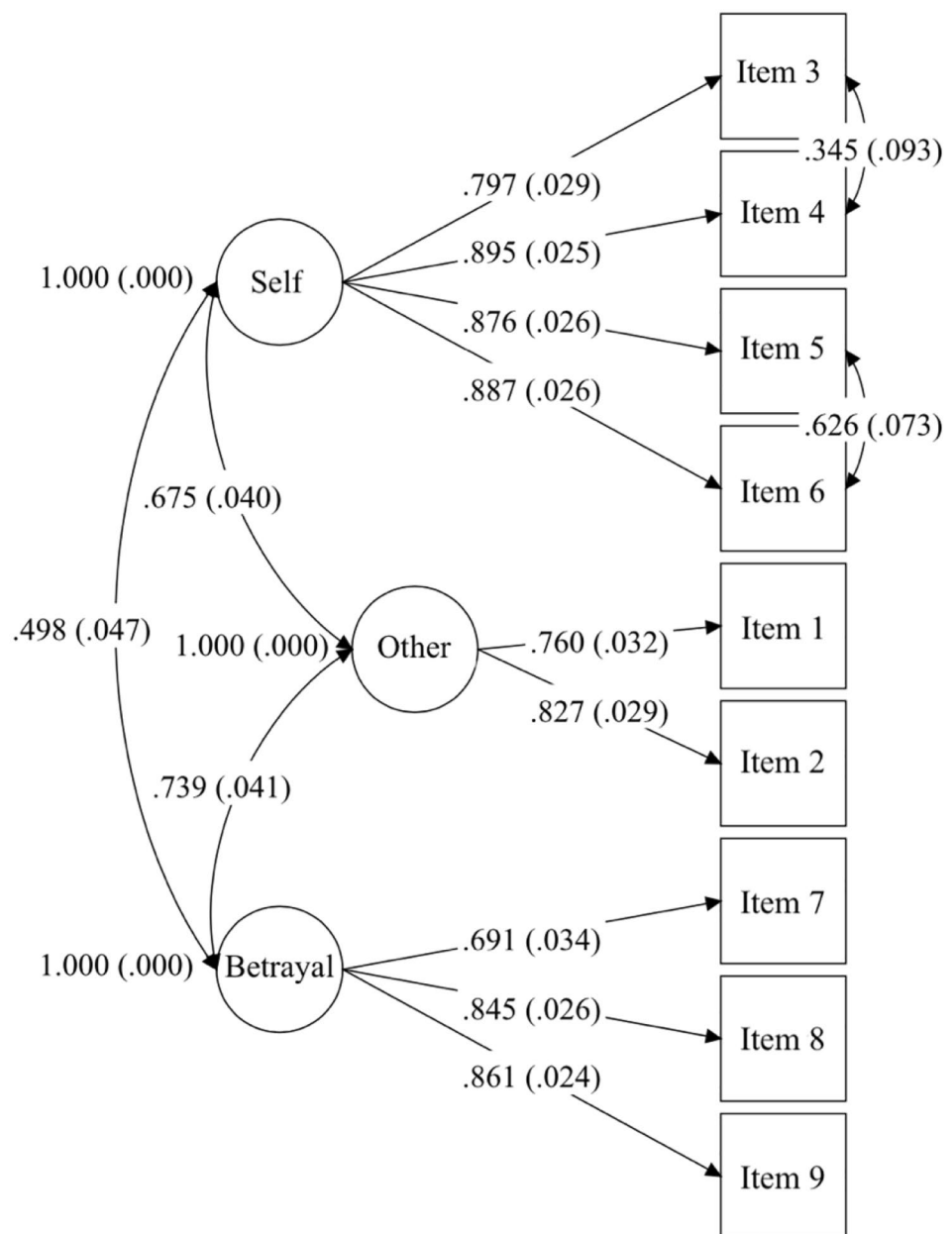


Table 2 Results of Confirmatory Factor Analyses for the MIES-C and EMIS-C

Model	χ^2	<i>df</i>	<i>p</i>	CFI	TLI	RMSEA 90% CI	
						Lower	Upper
MIES-C							
1. <i>Single-factor model</i>	722.99	27	<0.001	0.91	0.88	0.27	0.31
2. <i>Two-factor model</i>	420.64	26	<0.001	0.95	0.93	0.20	0.24
3. <i>Three-factor model</i>	50.50	22	<0.001	0.99	0.99	0.04	0.09
EMIS-C							
1. <i>Single-factor model</i>	784.76	119	<0.001	0.89	0.88	0.13	0.14
2. <i>Two-factor model</i>	330.88	116	<0.001	0.97	0.96	0.07	0.09
3. <i>Bi-factor model</i>	284.16	102	<0.001	0.97	0.97	0.07	0.09

Note. CFI, Comparative Fit Index; TLI, Tucker-Lewis Index; RMSEA, Root Mean Square Error of Approximation

EMIS-C

Goodness of fit indices suggest that the single-factor model did not fit the data ($\chi^2 = 784.76$, $df = 119$, $p < 0.001$, RMSEA = 0.13, CFI = 0.89, TLI = 0.88). However, the two-factor model, with covaried items 1/7 and 8/14, provided an acceptable fit ($\chi^2 = 330.88$, $df = 116$, $p < 0.001$, RMSEA = 0.08, CFI = 0.97, TLI = 0.96). These findings are consistent with Thomas et al. (in press). Item factor loadings are presented in Fig. 2. The bi-factor model was also found to be an acceptable fit ($\chi^2 = 284.16$, $df = 102$, $p < 0.001$, RMSEA = 0.08, CFI = 0.97, TLI = 0.97). Item factor loadings are presented in Fig. 3. Given that the two-factor (covaried) model and the bifactor model are not nested, chi-square difference tests could not be conducted. We therefore slightly favoured the two-factor model given its greater parsimony and the fact that a number of the item loadings for the bifactor model were low or negative in magnitude despite the favourable overall model fit. The CFA results for each model are presented in Table 2. Furthermore, a strong correlation was found between the two subscales of the EMIS-C ($r = 0.68$, $p < 0.001$).

Validity Analyses

The subscales for the MIES-C and EMIS-C all correlated significantly with each other. This provides support for construct validity. The strongest subscale correlation was observed between Betrayal (MIES-C) and Other-Directed (EMIS-C; $r = 0.66$). Other large correlations were found between Transgressions-Self (MIES-C) and Self-Directed (EMIS-C; $r = 0.55$), and between Transgressions-Other (MIES-C) and Other-Directed (EMIS-C; $r = 0.56$). The remaining subscale correlations were found to be moderate in magnitude: Transgressions-Self and Other-Directed ($r = 0.45$), Transgressions-Other and Self-Directed ($r = 0.40$), and Betrayal and Self-Directed ($r = 0.46$). All correlations were significant at the 0.01 level. Furthermore,

the correlation between the total scores was found to be large ($r = 0.69$, $p < 0.001$).

The total scores and subscale scores of each moral injury measure were found to correlate significantly with associated measures of psychological distress. The specific constructs included PTSD, complex PTSD, depression, anxiety, and anger. Regarding the MIES-C, the strongest correlation was observed with the PCL-5 and GAD-7 ($r = 0.49$, $p < 0.001$). For the EMIS-C, the strongest correlation was found with the ITQ ($r = 0.71$, $p < 0.001$ for both the PTSD and CPTSD subscales). Both findings indicate a strong association between moral injury and post-traumatic stress symptoms. The weakest correlations were observed with the DAR-5 ($r = 0.37$, $p < 0.001$ [MIES-C]; $r = 0.59$, $p < 0.001$ [EMIS-C]). All correlations are found in Table 3.

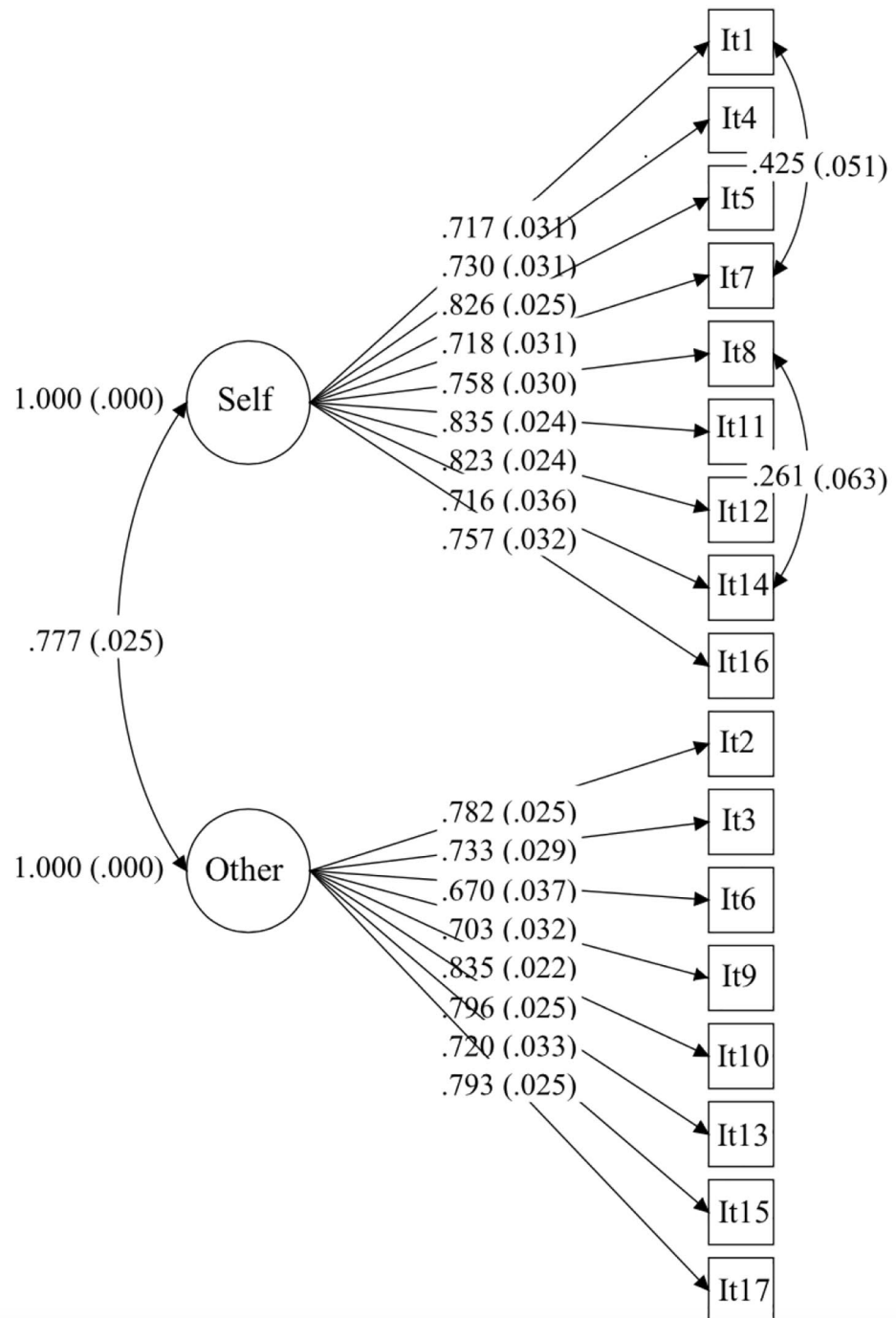
Test-Retest Reliability

To determine the criteria for test-retest reliability statistics, Cicchetti's (1994) classifications were used. Values between 0.40 and 0.59 are fair, 0.60 to 0.74 are good, and above 0.75 is excellent (Cicchetti, 1994). The test-retest reliability for the MIES-C subscales was found to be good; Transgressions-Self ($r = 0.70$), Transgressions-Other ($r = 0.60$), and Betrayal ($r = 0.68$). The test-retest reliability for the EMIS-C subscales was found to be excellent; Self-Directed ($r = 0.79$), and Other-Directed ($r = 0.79$). Significant correlations were found for all subscales.

Discussion

Moral injury was initially conceived with regard to military experiences. However, an increased recognition that the concept may apply in civilian contexts has highlighted the need for valid measurement in non-military populations. The current study aimed to further validate two brief measures of moral injury for use in the general population,

Fig. 2 Standardised Item Factor Loadings for the EMIS-C Two-Factor Model



the MIES-C and EMIS-C (Thomas et al., [in press](#)). Beyond replicating previous findings, the current study extended the initial validation study in two ways. First, in addition to the previously tested models, the study examined the fit of bifactor models for the MIES-C and EMIS-C. Second, the study examined the stability of the two measures across a two-week interval. Regarding factor analyses, a three-factor model was identified for the MIES-C and a two-factor model and bi-factor model was identified for the EMIS-C.

These models replicated optimum factor structures reported by Thomas et al. ([in press](#)). The measures also demonstrated sufficient convergent validity, with medium to strong correlations found with associated measures of psychological distress. Lastly, scores on each of the factors remained stable across a two-week interval. In sum, all hypotheses were supported.

The validity of the MIES-C was supported for the assessment of moral injury-related causes and responses. As

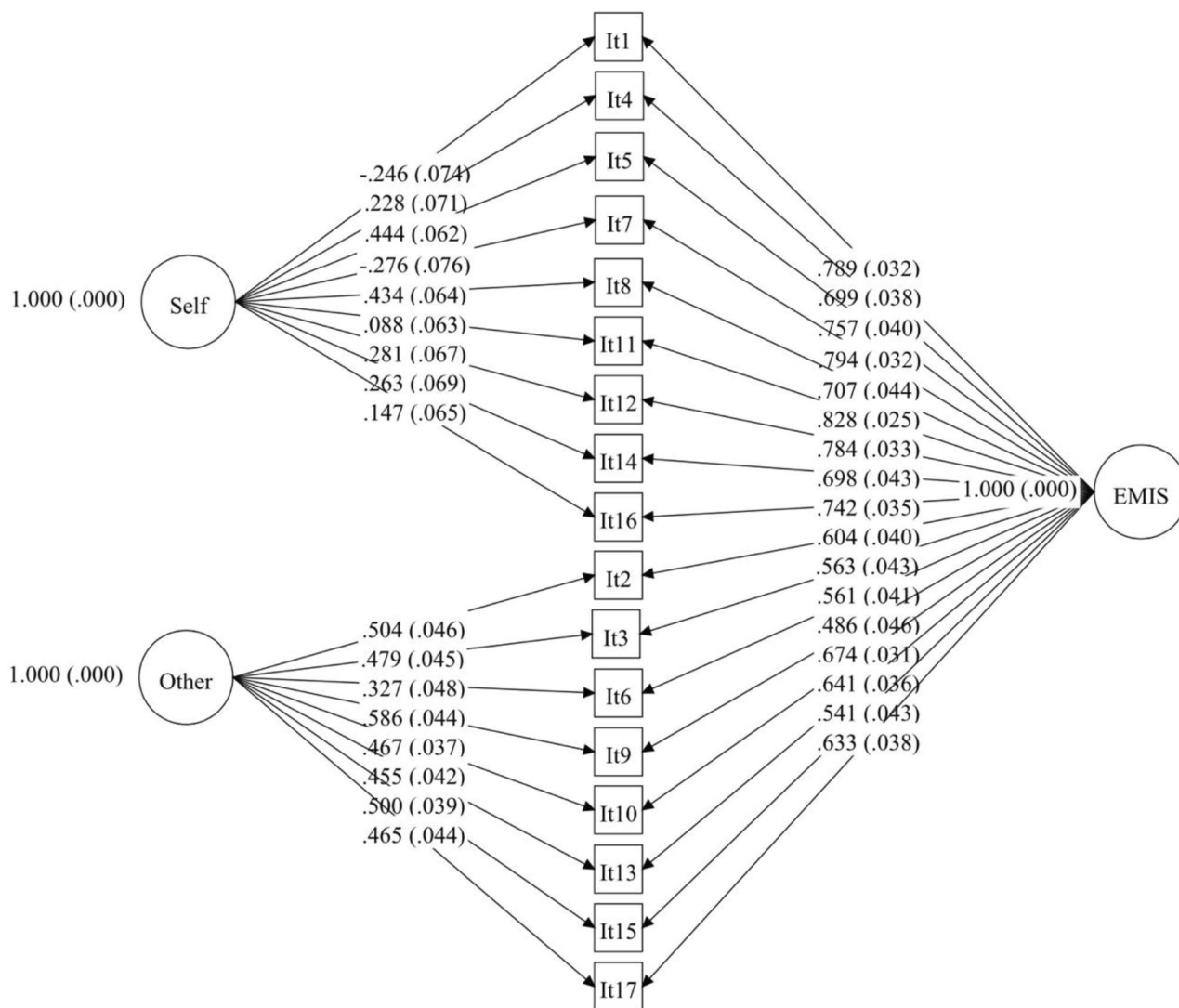


Fig. 3 Standardised Item Factor Loadings for the EMIS-C Bi-factor Model

Table 3 Correlations between Moral Injury Measures and Associated Constructs

	PCL-5	ITQ PTSD	ITQ CPTSD	PHQ-9	GAD-7	DAR-5
MIES-C Total	0.49	0.46	0.46	0.43	0.49	0.37
Trans-Self	0.36	0.36	0.36	0.30	0.36	0.32
Trans-Other	0.40	0.35	0.35	0.37	0.41	0.27
Betrayal	0.46	0.40	0.40	0.41	0.45	0.30
EMIS-C Total	0.70	0.71	0.71	0.60	0.66	0.59
Self	0.65	0.69	0.69	0.58	0.61	0.57
Other	0.62	0.61	0.61	0.52	0.60	0.51

Note. All correlations are significant at the 0.01 level. *PCL-5*, PTSD Checklist for DSM-5; *ITQ PTSD*, International Trauma Questionnaire PTSD Subscale; *ITQ CPTSD*, International Trauma Questionnaire CPTSD Subscale; *PHQ-9*, Patient Health Questionnaire; *GAD-7*, Generalised Anxiety Disorder Assessment; *DAR-5*, Dimensions of Anger Reactions

predicted, the three-factor model with covaried items 3/4 and 5/6 provided an acceptable fit. For this sample, moral injury was portrayed within three dimensions: perceived transgressions by the self, perceived transgressions by others, and perceived betrayals. The single-factor, two-factor, and bi-factor model were not supported; thus, responses are best interpreted at the subscale level. The strongest subscale correlation was seen between Transgressions-Other and Betrayal. It is possible that transgressions by others is more associated with feelings of betrayal, whereas transgressions by the self is more associated with feelings of shame and guilt (Kelley et al., 2019). Furthermore, the scale demonstrated sufficient convergent validity, by portraying medium-sized correlations with measures of PTSD, complex PTSD, depression, anxiety, and anger. The MIES-C was also found to have adequate test-retest reliability.

To evaluate the MIES-C, there are certain strengths and limitations to consider. Firstly, the measure is brief and publicly accessible. Its wording is non-specific, meaning that it can be applied to any individual regardless of their setting. It also captures the various transgressions that are outlined in Litz et al.'s (2009) widely used definition of moral injury: perpetrating, witnessing, and failing to prevent. It is noted however that some of the wording is vague. For example, numerous items begin with the phrase “I am troubled...” (items 2, 4, and 6). Individuals will likely have different interpretations of the meaning of ‘troubled’, thus impacting on the information gathered. Moreover, the brevity of the MIES-C means that few items assess each dimension. It is therefore possible that idiosyncratic experiences of moral injury may be overlooked. When administering this measure, researchers and clinicians will likely require a clinical interview to understand the associated emotions, target event, and functional impact of the injury.

The EMIS-C demonstrated utility in its assessment of the outcomes of moral injury. As predicted, the two-factor model with covaried items 1/7 and 8/14 provided an acceptable fit for the data. The findings also showed the bi-factor model to be an acceptable fit. As the bifactor and two-factor models are not nested, we were not able to determine if one provided a better relative fit than the other, but favoured the two-factor model on grounds of parsimony and because the bifactor model had a number of weak and negative item loadings. The EMIS-C also showed sufficient convergent validity, by demonstrating large-sized correlations with expected measures of psychological distress. Lastly, the EMIS-C was found to have strong test-retest reliability.

The strengths of the EMIS-C also include its brevity, non-specificity, and accessibility. Due to its focus on an individual's adjustment following moral injury, the EMIS-C is likely sensitive to detect improvement. Within clinical contexts, this information is valuable when evaluating

treatment progress (Barnes et al., 2019). Adjustment difficulties following a moral injury experience are not yet codified as a diagnosis; thus, the lack of clinical cut-off scores is understandable. However, it remains unclear to what extent the scores correspond to functional impairment. Further, the instructions do not specify a time period. This could lead to difficulties in interpreting whether an individual's response is considered pathological or reasonable, and recent or historical. It is possible that further evaluation will lead to clearer guidance in the interpretation of scores. Overall, the MIES-C and EMIS-C were found to be psychometrically sound tools.

The validation of these measures provides researchers with general measurement tools to continue to explore civilian moral injury. Further studies might investigate the measurement invariance of these scales across different civilian populations. This would enable an examination of construct validity in different groups and shed light on the ways in which moral injury might manifest differently among various civilian groups. Sensitivity to treatment is yet to be demonstrated for these measures. However, our finding that scores remain stable in the absence of interventions provides confidence that treatment-related changes in scores would reflect a real response to the respective intervention. Clinical validation may also aid therapists in selecting interventions (e.g., self-compassion), building trust, and assessing treatment duration (Williamson et al., 2021). In turn, this may also lead to the development of evidence-based treatment protocols for moral injury-related presentations.

Besides being the first to examine the test-retest reliability of these scales, an additional contribution of the current study was consideration of a bi-factor model. There is increasing acknowledgement that psychological constructs often comprise both uni- and multi-dimensional structures (Bornovalova et al., 2020); therefore, the study utilised this approach in assessing the construct validity of the MIES-C and EMIS-C.

The study also had limitations. While there was a high rate of participant retention for our determination of test-retest reliability (93%), this form of reliability was only examined across a two-week interval. It is likely that these measures will be used as pre- and post-intervention tools; thus, longer intervals would have provided valuable information (Desmet et al., 2021). Further, the findings rely solely on self-report information. Self-report methods have the potential for bias, as participants may misunderstand items or interpret concepts differently (Gayer-Anderson et al., 2020). It is also noted that the use of a non-clinical population limits the generalisability of the results to clinical settings. Another limitation relates to the recruitment method. The results derived from Prolific seemed to indicate particularly high rates of trauma exposure (see Table 1), compared to

rates identified in established epidemiological surveys (e.g., approximately 75% in Mills et al., 2011). Over-endorsement of trauma exposure suggests that endorsement of morally injurious events may have also been inflated. Interviewer-based assessments may have allowed additional verification of participant responses.

The injury associated with moral transgressions is likely to be seen within clinical contexts (Barnes et al., 2019). Therefore, future studies should seek to validate the scales with treatment seeking samples and among individuals who are most likely to be administered them. Considering the measures themselves, it may be worthwhile exploring the utility of two separate measures. One measure might screen for an individual's exposure to a morally injurious event, and the other might screen for the individual's response. This could address the limitation of unclear item meaning and provide clearer understanding of the experience; therefore, it is a suggestion for future exploration.

Our findings replicated Thomas et al.'s (in press) study with an independent sample and provided further examination into the psychometric properties of both measures. The three-factor model was identified as an acceptable fit for the MIES-C, while the two-factor and bi-factor model were identified as acceptable for the EMIS-C. Both measures demonstrated sufficient test-retest reliability and were found to correlate with associated measures of psychological distress. Considering the current ad-hoc approaches to measurement, universally applicable measures are needed to further understand the construct and its implications. Ultimately, the use of these validated measures will provide further insight into the experience of moral injury among civilians.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10862-023-10071-7>.

Acknowledgements The authors thank Victoria Thomas for kindly sharing the MIES-C and the EMIS-C.

Funding No funding was received for conducting this study.

Open Access funding enabled and organized by CAUL and its Member Institutions

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless

indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aldridge, V. K., Dovey, T. M., & Wade, A. (2017). Assessing test-retest reliability of psychological measures: Persistent methodological problems. *European Psychologist*, 22(4), 207–218. <https://doi.org/10.1027/1016-9040/a000298>.
- Barnes, H. A., Hurley, R. A., & Taber, K. H. (2019). Moral injury and PTSD: Often co-occurring yet mechanistically different. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 31(2), 4–103. <https://doi.org/10.1176/appi.neuropsych.19020036>.
- 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>.
- Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biological Psychiatry*, 88(1), 18–27. <https://doi.org/10.1016/j.biopsych.2020.01.013>.
- Bryan, C. J., Bryan, A. O., Anestis, M. D., Anestis, J. C., Green, B. A., Etienne, N., Morrow, C. E., & Ray-Sannerud, B. (2016). Measuring moral injury: Psychometric properties of the moral injury events scale in two military samples. *Assessment*, 23(5), 557–570. <https://doi.org/10.1177/1073191115590855>.
- Carleton, R. N., Affi, T. O., Taillieu, T., Turner, S., Krakauer, R., Anderson, G. S., MacPhee, R. S., Ricciardelli, R., Cramm, H. A., Groll, D., & McCreary, D. R. (2019). Exposures to potentially traumatic events among public safety personnel in Canada. *Canadian Journal of Behavioural Science*, 51(1), 37–52. <https://doi.org/10.1037/cbs0000115>.
- Cartolovni, A., Stolt, M., Scott, P. A., & Suhonen, R. (2021). Moral injury in healthcare professionals: A scoping review and discussion. *Nursing Ethics*, 28(5), 590–602. <https://doi.org/10.1177/0969733020966776>.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardised assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>.
- Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., Karatzias, T., & Hyland, P. (2018). The international trauma questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, 138(6), 536–546. <https://doi.org/10.1111/acps.12956>.
- Crane, M. F., Phillips, J. K., & Karin, E. (2015). Trait perfectionism strengthens the negative effects of moral stressors occurring in veterinary practice. *Australian Veterinary Journal*, 93(10), 354–360. <https://doi.org/10.1111/avj.12366>.
- Currier, J., Farnsworth, J. K., Drescher, K. D., McDermott, R. C., Sims, B. M., & Albright, D. L. (2018). Development and evaluation of the expressions of moral injury scale – military version. *Clinical Psychology & Psychotherapy*, 25(3), 474–488. <https://doi.org/10.1002/cpp.2170>.
- Dean, W., Talbot, S., & Dean, A. (2019). Reframing clinician distress: Moral injury not burnout. *Federal Practitioner*, 36(9), 400–402.

- Desmet, M., Van Nieuwenhove, K., De Smet, M., Meganck, R., Deeren, B., Van Huele, I., Decock, E., Raemdonck, E., Cornelis, S., Truijens, F., Zeuthen, K., & Schiepek, G. (2021). What too strict a method obscures about the validity of outcome measures. *Psychotherapy Research*, 31(7), 882–894. <https://doi.org/10.1080/10503307.2020.1865584>.
- Dhira, T. A., Rahman, M. A., Sarker, A. R., & Mehareen, J. (2021). Validity and reliability of the generalised anxiety disorder-7 (GAD-7) among university students of Bangladesh. *PLoS One*, 16(12), e0261590. <https://doi.org/10.1371/journal.pone.0261590>.
- Feinstein, A., Pavisian, B., & Storm, H. (2018). Journalists covering the refugee and migration crisis are affected by moral injury not PTSD. *JRSM Open*, 9(3), 1–10. <https://doi.org/10.1177/2054270418759010>.
- Forbes, D., Alkemade, N., Mitchell, D., Elhai, J. D., McHugh, T., Bates, G., Novaco, R. W., Bryant, R., & Lewis, V. (2014). Utility of the dimensions of anger reactions – 5 (DAR-5) scale as a brief anger measure. *Depression and Anxiety*, 31(2), 166–173. <https://doi.org/10.1002/da.22148>.
- Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of Life Research*, 30(8), 2197–2218. <https://doi.org/10.1007/s11136-021-02822-4>.
- Gayer-Anderson, C., Reininghaus, U., Paetzold, I., Hubbard, K., Beards, S., Mondelli, V., Di Forti, M., Murray, R. M., Pariante, C. M., Dazzan, P., Craig, T. J., Fisher, H. L., & Morgan, C. (2020). A comparison between self-report and interviewer-rated retrospective reports of childhood abuse among individuals with first-episode psychosis and population-based controls. *Journal of Psychiatric Research*, 123(2), 145–150. <https://doi.org/10.1016/j.jpsychires.2020.02.002>.
- Goulart, A. C., Bismarchi, D., Rienzo, M., Syllos, D. H., & Wang, Y. P. (2021). Dimensions of anger reactions (DAR-5): A useful screening tool for anger in the general population. *International Journal of Psychiatry in Clinical Practice*, 25(4), 421–429. <https://doi.org/10.1080/13651501.2020.1821893>.
- Gray, M. J., Litz, B., Hsu, J. L., & Lombardo, T. W. (2004). Psychometrics properties of the life events checklist. *Assessment*, 11(4), 330–341. <https://doi.org/10.1177/1073191104269954>.
- Griffin, B. J., Purcell, N., Burkman, K., Litz, B. T., Bryan, C. J., Schmitz, M., Villierme, C., Walsh, J., & Maguen, S. (2019). Moral injury: An integrative review. *Journal of Traumatic Stress*, 32(3), 350–362. <https://doi.org/10.1002/jts.22362>.
- Hoffman, J., Liddell, B., Bryant, R. A., & Nickerson, A. (2018). The relationship between moral injury appraisals, trauma exposure, and mental health in refugees. *Depression and Anxiety*, 35(11), 1030–1039. <https://doi.org/10.1002/da.22787>.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Ibrahim, H., Ertl, V., Catani, C., Ismail, A. A., & Neuner, F. (2018). The validity of posttraumatic stress disorder checklist for DSM-5 (PCL-5) as screening instrument with kurdish and arab displaced populations living in the Kurdistan region of Iraq. *Bmc Psychiatry*, 18(259), 1–8. <https://doi.org/10.1186/s12888-018-1839-z>.
- Jamieson, N., Maple, M., Ratnarajah, D., & Usher, K. (2020). Military moral injury: A concept analysis. *International Journal of Mental Health Nursing*, 29(6), 1049–1066. <https://doi.org/10.1111/inm.12792>.
- Kelley, M. L., Bravo, A. J., Davies, R. L., Hamrick, H. C., Vinci, C., & Redman, J. C. (2019). Moral injury and suicidality among combat-wounded veterans: The moderating effects of social connectedness and self-compassion. *Psychological Trauma: Theory Research Practice and Policy*, 11(6), 621–629. <https://doi.org/10.1037/tra0000447>.
- Koenig, H. G., Youssef, N. A., & Pearce, M. (2019). Assessment of moral injury in veterans and active duty military personnel with PTSD: A review. *Frontiers in Psychiatry*, 10(443), 1–15. <https://doi.org/10.3389/fpsy.2019.00443>.
- Kopacz, M. S., Ames, D., & Koenig, H. G. (2019). It's time to talk about physician burnout and moral injury. *The Lancet Psychiatry*, 6(11), E28. [https://doi.org/10.1016/S2215-0366\(19\)30385-2](https://doi.org/10.1016/S2215-0366(19)30385-2).
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>.
- Lentz, L. M., Smith-MacDonald, L., Malloy, D., Carleton, R. N., & Bremault-Phillips, S. (2021). Compromised conscience: A scoping review of moral injury among firefighters, paramedics, and police officers. *Frontiers in Psychology*, 12(639781), 1–15. <https://doi.org/10.3389/fpsyg.2021.639781>.
- Litz, B. T., Stein, N., Delaney, E., Lebowitz, L., Nash, W. P., Silva, C., & Maguen, S. (2009). Moral injury and moral repair in war veterans: A preliminary model and intervention strategy. *Clinical Psychology Review*, 29(8), 695–706. <https://doi.org/10.1016/j.cpr.2009.07.003>.
- Mantri, S., Lawson, J. M., Wang, Z., & Koenig, H. G. (2020). Identifying moral injury in healthcare professionals: The moral injury symptom scale-HP. *Journal of Religion and Health*, 59(5), 2323–2340. <https://doi.org/10.1007/s10943-020-01065-w>.
- Mills, K. L., McFarlane, A. C., Slade, T., Creamer, M., Silove, D., Teesson, M., & Bryant, R. (2011). Assessing the prevalence of trauma exposure in epidemiological surveys. *Australian and New Zealand Journal of Psychiatry*, 45(5), 407–415. <https://doi.org/10.3109/00048674.2010.543654>.
- Nash, W. P., Carper, M., Mills, T. L., Au, M. A., T., Goldsmith, A., & Litz, B. T. (2013). Psychometric evaluation of the moral injury events scale. *Military Medicine*, 178(6), 646–652. <https://doi.org/10.7205/MILMED-D-13-00017>.
- Nickerson, A., Schnyder, U., Bryant, R. A., Schick, M., Mueller, J., & Morina, N. (2015). Moral injury in traumatised refugees. *Psychotherapy and Psychosomatics*, 84(2), 122–123. <https://doi.org/10.1159/000369353>.
- Palan, S., & Schitter, C. (2018). Prolific.ac – a subject pool for online experiments. *Journal of Behavioural and Experimental Finance*, 17(4), 22–27. <https://doi.org/10.1016/j.jbef.2017.12.004>.
- Papazoglou, K., & Chopko, B. (2017). The role of moral suffering (moral distress and moral injury) in police compassion fatigue and PTSD: An unexplored topic. *Frontiers in Psychology*, 8(1999), 1–5. <https://doi.org/10.3389/fpsyg.2017.01999>.
- Sele, P., Hoffart, A., Bækkelund, H., & Øktedalen, T. (2020). Psychometric properties of the international trauma questionnaire (ITQ) examined in a norwegian trauma-exposed clinical sample. *European Journal of Psychotraumatology*, 11(1), 1796187. <https://doi.org/10.1080/20008198.2020.1796187>.
- Shay, J. (1994). *Achilles in Vietnam: Combat trauma and the undoing of character*. Scribner.
- Shay, J. (2014). Moral injury. *Psychoanalytic Psychology*, 31(2), 182–191. <https://doi.org/10.1037/a0036090>.
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalised anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>.
- Sugrue, E. P. (2019). Moral injury among professionals in K-12 education. *American Educational Research Journal*, 57(1), 43–68. <https://doi.org/10.3102/0002831219848690>.
- Sun, Y., Fu, Z., Bo, Q., Mao, Z., Ma, X., & Wang, C. (2020). The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. *Bmc Psychiatry*, 20(474), 1–7. <https://doi.org/10.1186/s12888-020-02885-6>.

- Thomas, V., Bizumic, B., Cruwys, T., & Walsh, E. (in press). Measuring civilian moral injury: Adaptation and validation of the moral injury events scale (civilian) and expressions of moral injury scale (civilian). *Psychological Trauma: Theory Research Practice and Policy*. <https://doi.org/10.1037/tra0001490>.
- Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013a). The Life Events Checklist for DSM-5 (LEC-5) [Measurement instrument]. Available from: <https://www.ptsd.va.gov>.
- Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013b). The PTSD Checklist for DSM-5 (PCL-5) – Standard [Measurement instrument]. Available from: <https://www.ptsd.va.gov>.
- Williamson, V., Murphy, D., Phelps, A., Forbes, D., & Greenberg, N. (2021). Moral injury: The effect on mental health and implications for treatment. *The Lancet Psychiatry*, 8(6), 453–455. [https://doi.org/10.1016/S2215-0366\(21\)00113-9](https://doi.org/10.1016/S2215-0366(21)00113-9).
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 76(6), 913–934. <https://doi.org/10.1177/0013164413495237>.
- World Health Organisation. (2019). *International statistical classification of diseases and related health problems* (11th ed.). World Health Organisation.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.