AMERICAN PSYCHOLOGICAL ASSOCIATION
TRAUMA PSYCHOLOGY
AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Trauma:
Theory, Research, Practice, and Policy

In the public domain
ISSN: 1942-9681

2024, Vol. 16, No. 2, 280–291
https://doi.org/10.1037/tra0001533

The Moral Injury and Distress Scale: Psychometric Evaluation and Initial Validation in Three High-Risk Populations

Sonya B. Norman[1, 2], Brandon J. Griffin[3, 4], Robert H. Pietrzak[5, 6], Carmen McLean[7, 8], Jessica L. Hamblen[1, 9], and Shira Maguen[10, 11]

[1] Executive Division, National Center for PTSD, White River Junction, Vermont, United States
[2] Psychiatry Department, University of California San Diego School of Medicine
[3] Center for Mental Healthcare & Outcomes Research, Health Services Research and Development Service, Little Rock, Arkansas, United States
[4] Center for Health Services Research, Department of Psychiatry, University of Arkansas for Medical Sciences
[5] Neurosciences Division, National Center for PTSD, West Haven, Connecticut, United States
[6] Psychiatry Department, Yale School of Medicine
[7] Dissemination and Training Division, National Center for PTSD, Palo Alto, California, United States
[8] Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University
[9] Psychiatry Department, Geisel School of Medicine at Dartmouth
[10] Mental Health Service, San Francisco VA Health Care System, San Francisco, California, United States
[11] Department of Psychiatry and Behavioral Sciences, University of California San Francisco School of Medicine

Objective: The concept of moral injury resonates with impacted populations, but research has been limited by existing measures, which have primarily focused on war veterans and asked about exposure to potentially morally injurious events (PMIEs) rather than PMIE exposure outcomes. Our goal was to develop and examine the psychometric properties of the Moral Injury and Distress Scale (MIDS), a new measure of the possible emotional, cognitive, behavioral, social, and/or spiritual sequelae of PMIE exposure. ***Method:*** The MIDS was validated by surveying three groups: military veterans, healthcare workers, and first responders ($N = 1{,}232$). ***Results:*** Most respondents (75.0%; $n = 924$) reported PMIE exposure. Analyses yielded 18 items that contributed to a single latent factor representing moral distress with fully or partially invariant configurations, loadings, and intercepts across occupational groups. The MIDS full-scale score demonstrated excellent internal consistency ($\alpha = .95$) and moderate 2-week stability ($r = .68$, $p < .001$, $n = 155$). For convergent validity, associations between the MIDS and PMIE exposure measures, as well as putative indicators of moral injury (e.g., guilt, shame), were positive and large ($r = .59$–$.69$, $p < .001$), as were correlations with posttraumatic stress, depressive, and insomnia symptoms ($r = .51$–$.67$, $p < .001$). The MIDS was a stronger predictor of functioning than PMIE exposure measures, explaining seven times greater unique variance (9% vs. 1%–1.3%). ***Conclusions:*** The MIDS is the first scale to assess moral injury symptoms indexed to a specific PMIE that is validated across several high-risk populations.

Clinical Impact Statement
The Moral Injury and Distress Scale is the first measure to assess moral injury symptoms, indexed to a specific potentially morally injurious event, that is validated across several high-risk populations, including veterans, healthcare workers, and first responders.

Keywords: moral injury, veterans, healthcare workers, first responders, trauma

Supplemental materials: https://doi.org/10.1037/tra0001533.supp

This article was published Online First June 22, 2023.

Shira Maguen https://orcid.org/0000-0002-1234-7234

Data and study materials are available to qualified investigators by contacting the authors. The authors have no conflicts of interest to disclose. We would like to thank Haley Mehlman for her assistance with this manuscript.

Robert H. Pietrzak served in a supporting role for conceptualization and writing–review and editing. Carmen McLean served in a supporting role for conceptualization and writing–review and editing. Jessica L. Hamblen served in a supporting role for conceptualization and writing–review and editing. Sonya B. Norman and Shira Maguen contributed equally to conceptualization and project administration. Sonya B. Norman, Brandon J. Griffin, and Shira Maguen contributed equally to writing–original draft, writing–review and editing, and methodology. Brandon J. Griffin contributed to formal analysis. Robert H. Pietrzak, Carmen McLean, and Jessica L. Hamblen served in a supporting role for conceptualization and writing–review and editing.

Correspondence concerning this article should be addressed to Shira Maguen, Mental Health Service, San Francisco VA Health Care System, 4150 Clement Street (116-P), San Francisco, CA 94121, United States. Email: Shira.Maguen@va.gov

In the most widely used and accepted definition of moral injury, it is defined as the lasting emotional, cognitive, behavioral, social, and/or spiritual impact of exposure to a potentially morally injurious event (PMIE) such as "perpetrating, failing to prevent, bearing witness to, or learning about acts that transgress deeply held moral beliefs or expectations (Litz et al., 2009, p. 696)." It is further conceptualized as the constellation of symptoms (e.g., guilt, disgust, inability to self-forgive) and related functional impairments (e.g., self-sabotaging behaviors, self-punishment) that can follow exposure to a PMIE (Jinkerson, 2016). Mixed methods empirical studies have supported this definition and conceptualization of moral injury (e.g., Griffin et al., 2019; Purcell et al., 2016). Other conceptualizations of moral injury are consistent with that of Litz and colleagues in regard to the core features and symptoms (Griffin et al., 2019), with the exception of Shay (2014), who proposed that betrayal by an authority or other in a position of power may also constitute a type of moral injury.

Research on moral injury has proliferated over the last decade, most commonly in studies of veterans exposed to war (e.g., Griffin et al., 2019; Maguen & Norman, 2022). Moral injury research has recently expanded from military veterans to include other populations, such as healthcare workers on the frontlines of the COVID-19 pandemic (Hines et al., 2021; Norman et al., 2021; Riedel et al., 2022). Research showing that moral injury is prevalent, distressing and impairing across populations, particularly among those working in high stress environments, has revealed a need for a psychometrically sound measure suitable for a broad range of populations. However, existing measures of moral injury have several limitations.

One limitation is that existing measures that purport to assess moral injury were developed and validated with veterans and query war-related situations (Bryan et al., 2016; Currier et al., 2020; Litz et al., 2022; Nash et al., 2013). There is a lack of measures validated with other impacted populations, including healthcare workers and first responders (emergency medical technicians, paramedics, firefighters, law enforcement, etc.), whose work routinely involves life-and-death situations, particularly during crises or disasters (Lentz et al., 2021). In some cases, scholars have adapted measures originally developed in veteran samples for use with other populations, although these are usually not validated for the new population (Khan et al., 2021; Koenig et al., 2019; Zhizhong et al., 2020). Moral injury assessments validated for use across populations and professions would allow for comparisons across impacted groups and enable the field to study the prevalence and impact of moral injury during disasters such as the COVID-19 pandemic (Williamson et al., 2020).

Another limitation is that many existing measures conflate exposure to a PMIE (e.g., witnessing the death of a civilian in war) with potential impacts of these exposures (e.g., the lasting psychological and spiritual distress) or measure only one or the other. A measure that clearly delineates exposure from the sequelae of the exposure would allow for a better understanding of whether certain types of exposure (e.g., perpetration vs. witnessing) lead to different moral injury responses. An ideal measure would assess a range of psychological and functional problems that characterize moral injury and link these impacts to a precipitating event or series of events, thus acknowledging that morally injurious exposures and outcomes are related but distinct constructs. When moral injury signs and symptoms are queried without indexing to a clearly defined PMIE, outcomes may reflect general distress, rather than moral injury specifically. For example, individuals may feel guilt or disgust about general life stressors (e.g., ruptured interpersonal relationships, unmet personal goals) that are not specific to a PMIE. In contrast, measuring PMIE exposure without the hallmark symptoms makes it challenging to fully understand symptoms and functional impairment related to these exposures. Relatedly, to accurately assess the efficacy of moral injury treatments, it is important to have measures that assess symptoms on a continuum from mild to severe and can identify changes in symptoms and functioning. Thus, a psychometrically sound scale is needed that comprehensively assesses PMIE exposure in addition to hallmark indicators of moral injury, including psychological, behavioral, social, and spiritual outcomes.

Another concern with existing measures is inconsistent factor structure, sometimes within the same measure (Bryan et al., 2016; Nash et al., 2013; Richardson et al., 2020), which obfuscates construct validity. One reason for these mixed findings may be that betrayal has been conceptualized in a number of different ways in relation to moral injury, and questions about whether it is a type of PMIE exposure, a separate or co-occurring construct, or a resulting symptom of a PMIE exposure remain (Norman et al., 2022). Early in the conceptualization of moral injury, Shay (2014) proposed that betrayal by leadership was a type of PMIE. Subsequently, some moral injury measures included betrayal from a variety of sources (e.g., others in the military, others outside the military; Nash et al., 2013) as types of PMIEs or symptoms of moral injury. These inconsistent uses of betrayal yielded inconsistent findings, likely because the items were often broad, went beyond the scope of PMIEs (feelings of betrayal were generally endorsed rather than tied to a specific PMIE event), and were not part of a preexisting theory (e.g., Litz et al., 2009; Shay, 2014). Some empirical studies have demonstrated that betrayal may be more associated with posttraumatic stress disorder (PTSD) than moral injury and may need to be considered separately from moral injury due to its unique signs and symptoms (Borges et al., 2021; Jordan et al., 2017; Maguen & Norman, 2022; Maguen et al., 2020). This uncertainty has led to inconsistencies, which have consequently limited the ability to accurately measure moral injury.

Additionally, some measures of moral injury utilize a bifactor model with a general factor representing overall distress and two specific factors representing distress directed toward either oneself or others (Currier et al., 2020). In practice, however, item responses on these and other measures of moral injury are typically summed to produce a total score without evidence demonstrating that the items are better measures of the overall general factor than the domain-specific factors (Chesnut et al., 2022; Koenig et al., 2018; Litz et al., 2022). Because unidimensionality is a critical assumption of most scoring models, a unidimensional scale to assess moral distress is needed.

To address the aforementioned limitations, our aim in the current study was to use prior moral injury research, expert opinion, and rigorous psychometric methods to develop and test the Moral Injury and Distress Scale (MIDS), a new measure of moral injury that could be used across PMIEs and populations and that separately measures PMIEs and reactions to these events. Based on Litz et al.'s (2009) definition of moral injury, this measure was designed to include information about index event(s) that met the threshold for exposure to PMIEs, as well as symptoms and related functional impacts of such exposures. We evaluated the psychometric properties and validity of

the MIDS in three populations at risk of moral injury, including combat veterans, healthcare workers, and first responders.

Method

Participants

We surveyed 1,232 combat veterans, healthcare workers, and first responders in the United States. For combat veterans, the sample was drawn from an existing panel at KnowledgePanel, a probability-based, online, nonvolunteer access survey panel of a nationally representative sample of over 50,000 households maintained by Ipsos, Inc. Panel members are recruited through national random samples by telephone and postal mail. KnowledgePanel recruitment includes both listed and unlisted telephone numbers, telephone and nontelephone households, cell-phone-only households, and households without Internet access. During recruitment, KnowledgePanel asked two screening questions to confirm combat veteran status ("Have you ever served on active duty in the U.S. Armed Forces, Military Reserves, or National Guard?" and "Did you ever serve in a combat or war zone (e.g., Korean War, Vietnam War, Gulf War, or any recent post-9/11 conflicts in the Middle East)?").

For healthcare workers and first responders, individuals were recruited through KnowledgePanel as opt-in samples from associated survey panels. To qualify for the study, these individuals needed to work in their field for at least 6 months and either still work in or have left the field no more than 5 years ago. For the minority who fit more than one group, veterans remained in the veteran group, and those who qualified as both first responders and healthcare workers were assigned to the first responder group. Consent to participate was obtained from all respondents upon panel enrollment, and all procedures involved in this work complied with the ethical standards of the relevant national and institutional guidelines for human research. More specifically, through a written agreement with Ipsos, Inc. (survey company), KnowledgePanel members are made aware that every survey and item is voluntary, and that data are collected and provided to clients anonymously. Ipsos does not require survey-specific written consent for KnowledgePanel members. We consulted with the VA San Diego Institutional Review Board, and they determined that no additional consent was required.

Most respondents (75.0%; $n = 924$) reported either participating in or witnessing a PMIE. The remainder reported no PMIE exposure (25.0%; $n = 308$). Those who endorsed PMIE exposure answered a follow-up question: "Please pick the event or series of events that is most troubling to you or that you think about the most. Write a few sentences to briefly describe the event(s)." Examples included, *I had to kill children that shot at us during a patrol* (veteran), *I was told to take patients off ventilators that I thought could be saved* (healthcare worker), and *I saw a fellow police officer use excessive force and did not intervene* (first responder). Two authors (Shira Maguen and Sonya B. Norman) reviewed the qualitative descriptions; 279 cases were excluded for providing insufficient information, causing it to be unclear whether they had experienced a PMIE or not (e.g., "I wish I knew the things I now know") or an example that was not a moral injury (e.g., "Personal goals that I did not achieve"). Since this is the first study that queried moral injury in a variety of populations resulting from a variety of experiences, not specific life-threatening combat or medical events, we believed this review was important to ensure we included examples that were probable PMIEs (rather than general life stressors). Only those who denied PMIE exposure or reported a valid PMIE exposure were included in the analyses ($N = 953$).

We report demographics for the full sample and each occupational subgroup in Table 1. Combat veterans ($n = 302$) reported serving in the Army (39.1%), Navy (25.2%), Air Force (18.5%), Marines

Table 1
Sociodemographics for the Full Sample and by Subgroup

Category	Full sample ($n = 953$)	Combat veterans ($n = 302$)	Healthcare workers ($n = 356$)	First responders ($n = 295$)
Gender				
Man	55.2	95.4	14.9	62.7
Woman	43.5	3.6	83.4	36.3
Other	0.5	0.6	0.6	0.7
Prefer not to answer	0.7	0.7	1.1	0.3
Age				
18–39 years	19.9	3.6	29.5	25.1
40–59 years	37.5	21.5	40.4	50.2
60+ years	42.6	74.8	30.1	24.7
Race/ethnicity				
White, non-Hispanic	75.4	82.5	71.3	73.2
Black, non-Hispanic	9.4	6.3	11.5	10.2
Other, non-Hispanic	2.2	1.7	3.9	0.7
Hispanic	9.8	7.3	10.4	11.5
Multiracial, non-Hispanic	3.1	2.3	2.8	4.4
Sexual orientation				
Heterosexual	93.7	96.7	92.1	92.5
LGBQ/other	6.2	3.3	7.6	7.5
Prefer not to answer	0.1	0.0	0.3	0.0
Employment status				
Full/part time	62.4	27.5	78.9	78.3
Unemployed	2.4	1.3	4.2	1.4
Disability (VA, SSI, SSDI)	2.7	5.0	1.4	2.0
Retired	30.1	65.2	11.8	16.3
Other	2.2	1.0	3.4	2.0
Prefer not to answer	0.1	0.0	0.3	0.0
Education status				
High school diploma/ GED	7.5	8.6	5.9	8.5
Some college/associate degree	37.6	39.1	34.8	39.3
Bachelor's degree	30.6	24.5	29.8	38.0
Master's degree or higher	24.2	27.8	29.5	14.2
Marital status				
Married	61.9	72.5	56.2	58.0
Divorced/separated/ widowed	23.1	21.5	23.6	24.1
Never married	15.0	6.0	20.2	18.0
Children under 18 years				
Yes	16.4	11.9	30.9	35.9
No	73.6	88.1	69.1	64.1
Annual household income				
49,999 or less	23.5	23.2	28.4	18.0
50,000–99,999	40.4	46.7	33.1	42.7
100,000 or more	36.1	30.1	38.5	39.3
Religious affiliation				
Catholic	24.2	26.8	23.0	23.1
Protestant	36.5	39.4	32.6	38.3
Other Christian	10.8	8.9	11.5	11.9
Other not Christian	5.5	3.3	7.0	5.8
Spiritual but not religious	9.4	7.9	11.2	8.8
Atheist/Agnostic	13.2	13.6	13.8	12.2
Prefer not to answer	0.3	0.0	0.8	0.0

Note. Values given as percentages. LGBQ = lesbian, gay, bisexual, and queer sexual orientation; GED = Genderal Education Development test or high school diploma equivalent; VA = Veterans Affairs; SSI = Supplemental Security Income; SSDI = Social Security Disability Insurance.

(7.6%), or other branch (9.6%, e.g., Coast Guard). Time in service ranged from 5 years or less (49.0%), between 6 and 19 years (19.9%), and 20 or more years (31.1%), and all reported at least one deployment to a warzone. Veterans reported "moderate" or greater levels of being exposed to and bothered by witnessing others' perceived transgressions (49.3%) and transgressing their own values by what they did (36.4%) or failed to do (27.5%).

Healthcare workers ($n = 356$) were nurses (43.8%; nurse practitioner, registered nurse, licensed practical nurse, etc.), allied health professionals (24.7%; pharmacist, psychologist, respiratory/physical/occupational/speech therapist, etc.), other nonclinical staff (19.1%; support staff, technician, administrator, volunteer), and physicians or physician assistants (7.0%), 5.3% did not provide details regarding their occupation. We were inclusive when conceptualizing healthcare workers, knowing that at certain times, such as during the COVID-19 pandemic, some may be asked to take on duties that are outside of their specialization or normal role. Time in occupation ranged from 5 years or less (21.3%), between 6 and 19 years (35.7%), 20 or more years (42.1%), and prefer not to answer (0.8%). Most healthcare workers (77.5%) were currently working in their role; 22.5% were no longer working in healthcare (e.g., retired). Healthcare workers reported "moderate" or greater levels of being exposed to and bothered by witnessing others' perceived transgressions (39.0%) and transgressing their own values by what they did (23.9%) or failed to do (22.0%).

First responders ($n = 295$) were law enforcement or corrections personnel (26.1%), emergency medical technicians or paramedics (18.3%), fire service or hazmat personnel (16.3%), and other first responders (39.3%; e.g., dispatcher, humanitarian/disaster worker, public works safety inspector). Time in occupation ranged from 5 years or less (24.4%), between 6 and 19 years (31.5%), 20 or more years (43.4%), and prefer not to answer (0.7%). Most (78.6%) were currently working in their role; 21.4% no longer worked as a first responder (e.g., retired). First responders reported "moderate" or greater levels of being exposed to and bothered by witnessing others' perceived transgressions (45.1%) and transgressing their own values by what they did (21.7%) or failed to do (16.3%).

Procedures

We used Boateng and colleagues' best practice guidelines for scale development and validation as a procedural framework (Boateng et al., 2018). Phase one involved generating (Step 1) and assessing the content validity (Step 2) of the candidate items. To do this, two authors (Shira Maguen and Sonya B. Norman) created a list of 41 face-valid items based on prior research and clinical experience to assess the sequelae of moral injury theorized by Litz et al. (2009). Items were developed to be rated on a continuous scale, recognizing that moral injury occurs on a continuum ranging from mild and transient moral distress, to more severe and lasting moral injury, which may be associated with significant functional impairment (Litz & Kerig, 2019). Four experts in psychological trauma reviewed the item list and suggested revisions. For example, they noted that several candidate items were indicators of mental health problems that commonly occur among those with moral injury as opposed to being indicators of moral injury per se (i.e., "I have had thoughts about hurting myself," "I put myself into risky situations," "I have harmed myself in some way," "I drink alcohol more," and "I use drugs more"). Similarly, five items appeared to assess nonspecific cynicism and disillusionment that were not necessarily associated with PMIEs (e.g., "I feel disillusioned by the world I live in," "I wonder what kind of world I live in") or not well aligned with Shay's (2014) definition positing betrayal by a "person who holds legitimate authority" (i.e., "I feel betrayed by people in my community"). Consistent with prior work in this area (Currier et al., 2018), we decided to eliminate these items prior to analysis because including them would artificially inflate correlations with mental health symptom measures and introduce a lack of clarity into our conceptualization of moral injury.

Phase two involved constructing the scale by pretesting questions (Step 3), administering the survey (Step 4), reducing the number of items (Step 5), and exploring the factor structure (Step 6). To pretest the questions, several of the authors (Shira Maguen, Sonya B. Norman, Robert H. Pietrzak, and Carmen McLean) conducted a series of interviews with veterans, healthcare workers, and first responders who provided feedback on the content validity and acceptability of the items. Veterans provided feedback in the context of a stakeholder engagement panel ($n = 4$), and healthcare workers ($n = 6$) and first responders ($n = 3$) were interviewed individually. The veteran panel meets regularly to give input to the National Center for PTSD investigators on research questions and educational products. Healthcare workers and first responders were colleagues who were interested in the topic of moral injury and offered to give input on the draft measure. We revised items as suggested by the various stakeholders. The items were then administered to respondents in two rounds. All respondents completed the baseline survey; a smaller subsample completed a follow-up survey 2 weeks later ($n = 155$). Finally, phase three involved testing the dimensionality (Step 7), reliability (Step 8), and validity (Step 9) of the newly developed scale. Data and study materials will be made available to qualified investigators by contacting the authors. A copy of the finalized scale is available in the online supplemental material 1.

Measures

Moral Injury and Distress Scale

The MIDS was created to provide a comprehensive assessment of morally injurious emotional, cognitive, spiritual, social, and/or behavioral sequelae in diverse populations, including combat veterans, healthcare workers, and first responders, ranging on a continuum from milder moral distress to more severe moral injury. Participants first identified an event or series of events in which they were bothered by something that they did, failed to do, or witnessed and that transgressed their moral beliefs (we used the Litz et al.'s [2009] definition of moral injury, which does not specify betrayal as a type of PMIE, but included feelings of betrayal by leadership/organization as a moral injury symptom, consistent with Shay [2014]). If participants did not identify such events, they were instructed to skip the rest of the scale and were considered to have no PMIE exposure. For those who identified an event, keeping the event(s) in mind, they indicated the extent to which a series of statements describing morally injurious sequelae was true of them in the last month using a 5-point response format (0 = *not at all*, 1 = *a little*, 2 = *moderately*, 3 = *quite a bit*, and 4 = *extremely*). Responses to each item were summed, such that higher scores indicated greater severity of moral distress.

Measures of Moral Injury and Putative Indicators of Moral Injury

We included existing moral injury measures as well as measures of constructs considered to be core components of moral injury. Given that existing moral injury measures were designed for use with veterans, some of the language was altered to make them applicable for a civilian population. The Morally Injurious Events Scale (MIES; Nash et al., 2013) is a nine-item self-report questionnaire used to assess exposure to and subjective distress from PMIEs. The scale assesses exposure by (a) witnessing, (b) perpetrating (through acts of commission or omission), or (c) being betrayed by leaders, peers, and others. Respondents indicated how much they agreed or disagreed with each statement (1 = *strongly disagree*, 6 = *strongly agree*). Item responses were averaged, such that higher scores on the MIES correspond to greater intensity of exposure and distress. Internal consistency of scores on the MIES was excellent in the current sample ($\alpha = .91$).

The Expressions of Moral Injury Scale—Military Version—Short Form (EMIS-M-SF; Currier et al., 2020) is a four-item self-report questionnaire based on the full-length Expressions of Moral Injury Scale-Military Version (EMIS; Currier et al., 2018). Two items capture self-directed moral injury, one asking about guilt and the other about shame in relation to events that happened during military service. The other two items ask about other-directed moral injury and query disgust and witnessing the moral failures of others. Respondents indicated how much they agreed or disagreed with each statement (1 = *strongly disagree*, 5 = *strongly agree*). Item responses were averaged to create a total score, with higher scores indicating greater levels of emotions and beliefs associated with moral injury. Internal consistency of participants' scores on the EMIS was good in the current sample ($\alpha = .87$).

Measures of constructs theorized to be core components of moral injury included the Trauma-Related Shame Inventory (TRSI), the Global Guilt scale of the Trauma-Related Guilt Inventory (TRGI), and selected subscales from the Religious and Spiritual Struggles Scale (RSSS). The Global Guilt scale of the TRGI (Kubany et al., 1996) consists of four items that measure trauma-related guilt (e.g., "I experience intense guilt related to what happened."). Respondents indicated how they felt about each statement (1 = *never/not at all true* to 5 = *always/extremely true*). Item responses were averaged, such that higher scores indicated more trauma-related guilt. Internal consistency of participants' scores on the TRGI Global Guilt scale was excellent in the current sample ($\alpha = .93$).

The TRSI-24 (Øktedalen et al., 2014) is a 24-item self-report questionnaire used to measure trauma-related shame, conceptualized by four categories (internal condemnation, internal affective-behavioral, external condemnation, and external affective-behavioral). Respondents indicated how true a statement was for them (0 = *not true of me*, 4 = *completely true of me*). Item responses were averaged, such that higher scores on the TRSI indicated greater levels of trauma-related shame. Internal consistency of participants' scores on the TRSI was excellent in the current sample ($\alpha = .98$).

We utilized three subscales from the RSSS (Exline et al., 2014). The original scale is a 26-item self-report questionnaire that measures religious and spiritual struggles across six domains (divine, demonic, interpersonal, moral, doubt, and ultimate meaning). We used the divine (five items; "Felt as though God had abandoned me"), moral (four items; "Worried that my actions were morally or spiritually wrong"), and doubt (four items; "Struggled to figure out what I really believe about religion/spirituality) subscales. Respondents indicated the extent to which they have had each experience (1 = *not at all*, 5 = *a great deal*). Item responses were averaged, such that higher scores indicated greater levels of religious/spiritual distress. Internal consistency of participants' scores on the divine ($\alpha = .94$), moral ($\alpha = .91$), and doubt ($\alpha = .95$) subscales for the current sample was excellent.

Measures of Mental Health

We assessed psychological problems including depression, posttraumatic stress, and posttraumatic cognitions. The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001) is a nine-item self-report questionnaire that assesses depressive symptoms from the Diagnostic and Statistical Manual, Fourth Edition (DSM-IV) criteria. Respondents indicated how often they have been bothered by problems over the past 2 weeks (0 = *not at all* to 3 = *nearly every day*). Internal consistency of participants' scores on PHQ-9 items was excellent in the current sample ($\alpha = .91$).

The PTSD Checklist for DSM-5 (PCL-5; Bovin et al., 2016) is a 20-item self-report questionnaire that measures posttraumatic stress symptoms corresponding to the DSM-5 diagnostic criteria. Respondents indicated how often they have had each problem or symptom over the past month (0 = *not at all*, 4 = *extremely*), with higher scores representing greater levels of posttraumatic stress symptoms. This version does not ask participants to anchor responses to a specific trauma. Internal consistency of participants' scores on PCL-5 items was excellent in the current sample ($\alpha = .96$). Additionally, we administered the Posttraumatic Cognitions Inventory-Short Form (PTCI; Foa et al., 1999). The PTCI is a nine-item self-report questionnaire used to assess thoughts after a traumatic experience. Respondents indicated how much they agreed or disagreed with each statement (1 = *totally disagree*, 7 = *totally agree*). Internal consistency of participants' scores on the PTCI was good in the current sample ($\alpha = .85$).

We also queried about insomnia and alcohol use. The Insomnia Severity Index (ISI; Morin et al., 2011) is a seven-item self-report questionnaire that assesses the severity of daytime and nighttime features of insomnia. Participants indicated the extent to which each issue was a problem for them (0 = *none*, 4 = *very much*). Item responses were summed with higher scores indicating higher levels of insomnia. Internal consistency of participants' scores on the ISI was excellent in the current sample ($\alpha = .91$). The Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993) is a 10-item self-report questionnaire that assesses drinking behaviors in terms of hazardous use and dependence symptoms. Item responses were summed with higher scores indicating greater levels of problematic alcohol use. Internal consistency of participants' scores on the AUDIT was good in the current sample ($\alpha = .83$).

Other Convergent and Discriminant Validity Measures

The Brief Inventory of Psychosocial Functioning (B-IPF; Kleiman et al., 2020) is a seven-item self-report questionnaire used to measure impairments in psychosocial functioning related to stress. Respondents indicated their level of trouble over the past month (0 = *not at all*, 6 = *very much*, with an additional n/a option)

for domains including work/education, self-care, and social relationships. Item responses were summed, divided by the maximum possible domain scale score for the applicable items, and multiplied by 100. Higher scores indicated greater impairment. The B-IPF demonstrated good-to-excellent internal consistency in the current sample ($\alpha = .87$).

Conversely, the Connor-Davidson Resilience Scale (CD-RISC; Connor & Davidson, 2003) is a 25-item self-report questionnaire designed to assess adaptability or the propensity to "bounce-back" from a stressor. Respondents indicated how much each statement applied over the past month (0 = *not true at all*, 4 = *true nearly all the time*). Item responses were summed, with higher scores indicating greater resilience. Internal consistency of the CD-RISC was excellent in the current sample ($\alpha = .92$).

The Big 5 Inventory Short Form (Rammstedt & John, 2007) is an abbreviated version of the full-length Big 5 Inventory designed to measure five dimensions of personality, including neuroticism, openness to experience, conscientiousness, extraversion, and agreeableness. The short form is a 10-item (two items per dimension) self-report questionnaire on which respondents indicated the extent to which they agree that each statement describes them (1 = *disagree strongly*, 5 = *agree strongly*). Item responses were averaged, such that higher scores indicated a greater presence of each trait. Consistent with Eisinga et al.'s (2013) recommendation for assessing the reliability of two-item scales, we calculated the Spearman–Brown coefficient for each subscale. Findings for neuroticism were fair ($\rho = .58$; roughly similar to $\alpha = .57$). The remaining subscales were excluded due to low reliability.

Data Analysis

Analyses were conducted using IBM SPSS Statistics, Version 27 and Mplus, Version 8.4 (IBM Corp., 2020; Muthén & Muthén, 2017). Missing data diagnostics revealed that <1% of the data (0.4%) were missing for <5% of cases (4.7%); thus, bias associated with incomplete data was not considered to be problematic. Once data collection was complete, we used a random number generator to divide the baseline sample into two subsamples: an initial validation subsample composed of 60% of the cases ($n = 582$) and a cross-validation subsample composed of 40% of the cases ($n = 371$). We used the initial validation subsample to reduce the number of items and examine the initial factor structure. This involved creating an item discrimination index by summing all items and separating cases in the 25th percentile or below from those in the 75th percentile or above. We calculated the difference in the proportion of respondents who endorsed a moderate or greater level of distress for each item, and we compared rates of endorsement between those in the highest and lowest quartiles. Items were dropped if they failed to discriminate (difference in proportion <0.20) between those in the highest and lowest quartiles. Then, we calculated interitem correlations. When a dyad of items was highly correlated ($r = .70$), suggesting redundancy in item content, we retained the item with greater discriminatory ability as indicated by the item discrimination index and eliminated the item with less discriminatory ability.

Next, we examined the factor structure by conducting exploratory factor analysis (EFA) using principal axis factoring. This included a parallel analysis to compare eigenvalues from the existing dataset against eigenvalues from randomly generated datasets with similar dimensions (Hayton et al., 2004). To do this, we calculated the mean Eigenvalue of each factor from the randomly generated datasets and plotted them against the Eigenvalues for each factor from the observed MIDS data. We retained factors in the observed data with Eigenvalues that exceeded Eigenvalues of factors extracted from the random data. Then, we conducted a confirmatory factor analysis (CFA) in the cross-validation subsample to determine whether the observed factor structure replicated in an independent sample. Model fit was assessed using the χ^2 value and a three-index strategy (Fan & Sivo, 2005). Values of 0.90 or above for the comparative fit index (CFI; Tucker & Lewis, 1973), .06 or below for the root mean square error of approximation (RMSEA; Browne & Cudeck, 1992), and 0.08 or below for the standardized root mean square residual (SRMR; Kline, 2005) were interpreted as acceptable model fit.

To further test the dimensionality of the scale across occupational subgroups, we combined the initial validation and cross-validation subsamples into a full sample and specified a series of multigroup confirmatory factor analyses (MGCFA) to evaluate measurement invariance across occupational groups. This involved testing whether the number of factors (configural model), factor loadings (metric model), and intercepts (scalar model) were consistent across veterans, healthcare workers, and first responders. We selected the first responder group as the reference group, because the first responder group was more heterogeneous than the other groups in terms of participants' genders and ages (i.e., as expected, veterans had a higher representation of older men and healthcare workers were more likely to be women). If evidence initially supported full invariance, we compared latent means across the groups. Otherwise, we established partial scalar invariance by comparing nested MGCFA models in which we sequentially freed intercept constraints prior to comparing latent means. Nested models were evaluated using χ^2, RMSEA, and CFI difference tests. We interpreted nonsignificant χ^2 differences tests, RMSEA difference values less than or equal to .01, and CFI difference values less than or equal to 0.01 as evidence of equivalence (Chen, 2007; Cheung & Rensvold, 2002).

Next, we tested the reliability of scores by calculating Cronbach's alpha (α) for the full sample and within each occupational subgroup, and we tested the stability of scores using Pearson correlation on a subsample of respondents who completed a follow-up survey containing the MIDS items 2 weeks after the baseline survey. The validity of participants' scores on the MIDS was examined by calculating Pearson correlations with the MIDS and validation measures that assessed guilt/shame, religious/spiritual struggle, mental health symptoms, psychosocial functioning, resilience, and personality traits. We used Cohen's (1988) guidelines for interpretating effectsas small ($r = .10$–$.30$), medium ($r = .30$–50), and large ($r \geq .50$). Finally, we tested whether scores on the MIDS predicted variance in functional impairment above and beyond variance explained by individual differences (e.g., gender, age, and race) and two existing measures of PMIE exposure using hierarchical multiple regression.

Results

Item Reduction

Using the initial validation subsample ($n = 582$), we calculated the item discrimination index and eliminated two items that failed to distinguish between those who scored in the highest and lowest quartiles on a sum of all items in the initial pool (e.g., "My spiritual

life has changed for the worse"). Additionally, we dropped seven items that were highly correlated with at least one other item, suggesting that the items' contents were redundant. For example, we removed "My life feels less meaningful" and retained "My life feels like it has less purpose." When two items were highly correlated, we kept the item with greater discriminatory ability. In sum, 21 items met criteria for inclusion and were carried forward to factor extraction.

Factor Extraction

We conducted an EFA using principal axis factoring to identify the factor structure of the retained items in the initial validation subsample. Parallel analysis supported a single-factor solution, such that only one factor in the observed data had an Eigenvalue ($\lambda = 11.25$) that exceeded the mean Eigenvalue of the largest factor obtained from the randomly generated dataset options ($\lambda = 1.39$). The single factor explained 53.6% of the total item variance in the observed data. To further eliminate redundancy in items, we calculated inter-item partial correlations that controlled for the sum score of the selected items and eliminated three items with residuals that were moderate to strongly correlated with another retained item (Funk & Rogge, 2007). For instance, we eliminated "I don't take care of my basic needs as well as I used to" and kept "I don't take good care of myself." We ran the EFA again with the retained 18 items to estimate factor loadings, which ranged between 0.62 and 0.76 (Table 2).

Because we dropped items based on characteristics of only the initial validation subsample, replication of the factor structure in an independent sample was needed. For this reason, we conducted CFA using the maximum likelihood estimator with robust standard errors in the cross-validation subsample ($n = 371$). We specified a single factor with the 18 retained items as indicators. Model fit was acceptable: χ^2 (135, $N = 371$) = 337.46, $p < .001$, RMSEA = .064 (95% CI [0.055, 0.072], $p = .005$), CFI = 0.88, SRMR = 0.055. Inspection of the modification indices revealed that fit could be improved by allowing the residuals of "I don't feel like I deserve to be happy" and "I should not be forgiven" to covary. We allowed these residuals to covary because the items are both positively valanced, unlike the majority of negatively valanced items that comprise the MIDS. The revised model fit the data very well: χ^2 (134, $N = 371$) = 298.04, $p < .001$, RMSEA = .057 ([0.049, 0.066], $p = .080$), CFI = 0.91, SRMR = 0.052. Factor loadings ranged between 0.56 and 0.83 and are displayed with item-level descriptive statistics in Table 2. Model comparison using the Santorra Bentler scaled chi square difference test (TRd) revealed that the model with correlated residuals for the aforementioned items fit the data better than did the model with no residual covariances (TRd = 77.16, $p < .001$). Overall, these findings supported Litz et al.'s (2009) model of moral injury by extracting and replicating a single latent construct representing moral injury sequelae with psychological/behavioral, social, and spiritual/existential indicators.

Tests of Dimensionality and Measurement Invariance

To further test the dimensionality of the scale across occupational subgroups, after combining the initial validation and cross-validation subsamples into a full sample, we specified a series of MGCFA to evaluate measurement invariance across occupational groups (Table S2 in the online supplemental materials). The Veteran-First Responder configural model fit the data well, χ^2 (270, $N = 597$) = 611.67, $p < .001$, RMSEA = .065 (95% CI [0.058, 0.072], $p < .001$), CFI = 0.870, SRMR = 0.058, indicating that the number of factors was consistent across groups. Factor loadings were also equivalent; Chi-square, RMSEA, and CFI tests revealed no difference in fit

Table 2
Descriptive Statistics and Factor Loadings for MIDS Items by Subsample

	Initial validation subsample ($n = 582$)		Cross-validation subsample ($n = 371$)	
Item	Freq	Loading	Freq	Loading
I think about how I should have been able to do more.	19.4	0.74	20.2	0.78
I have withdrawn from others more often.	15.5	0.71	11.3	0.70
I feel guilty.	11.7	0.76	10.0	0.72
I doubt my own judgment.	7.4	0.75	8.6	0.80
I do not feel like I deserve to be happy.	5.3	0.76	5.1	0.72
I self-sabotage things in my life more often (relationships, things at work).	6.7	0.73	6.7	0.73
I feel helpless.	11.0	0.75	11.3	0.83
My life feels like it has less purpose.	7.7	0.73	8.6	0.84
I am worried that bad things will happen to me or my loved ones.	11.3	0.73	10.5	0.75
I have punished myself.	6.5	0.74	5.4	0.78
I feel disgusted.	10.0	0.71	10.2	0.69
I do not seek support because I feel like I do not deserve it.	5.0	0.72	4.6	0.73
I do not seek support because I worry others would not understand.	10.8	0.71	11.9	0.72
I feel betrayed by leaders or institutions.	25.6	0.62	22.4	0.62
I feel powerless.	11.3	0.70	10.2	0.80
I should not be forgiven.	6.4	0.66	7.8	0.56
My spirituality/faith is no longer a source of comfort.	7.4	0.65	8.1	0.64
I do not take good care of myself.	10.1	0.64	12.1	0.63

Note. Frequencies (Freq) are percentages of respondents who endorsed "moderately" or higher for each item. Frequency is reported as binary in this table to enhance interpretation but was analyzed as ordered categorical (0–4) in exploratory and confirmatory factor analyses. MIDS = Moral Injury and Distress Scale.

between the metric and configural models, $\Delta\chi^2$ (17) = 21.97, $p = .186$, ΔRMSEA = .002, ΔCFI = 0.000. Although there was no evidence of invariance based on the CFI and RMSEA difference tests, the Chi-square difference test indicated that the scalar model fit the data significantly worse than the metric model, $\Delta\chi^2$ (17) = 29.83, $p = .028$, ΔRMSEA = .002, ΔCFI = 0.006. When we freed the intercept for the item "I should not be forgiven" ($Int_{veteran}$ = 0.36, $Int_{first\ responder}$ = 0.22), we found evidence of partial scalar invariance, $\Delta\chi^2$ (16) = 22.16, $p = .138$, ΔRMSEA = .000, ΔCFI = 0.005. Thus, we compared latent means and found no significant differences between veterans and first responders on the MIDS overall factor score ($\Delta m = 0.05$, $p = .425$).

The Healthcare Worker-First Responder configural model fit the data well, χ^2 (266, $N = 651$) = 647.27, $p < .001$, RMSEA = .065 (95% CI [0.060, 0.073], $p < .001$), CFI = 0.876, SRMR = 0.055, indicating that the number of factors was consistent across groups. Again, factor loadings were equivalent between groups. Chi-square, RMSEA, and CFI tests demonstrated no differences in fit between the metric and configural models, $\Delta\chi^2$ (17) = 16.85, $p = .464$, ΔRMSEA = .002, ΔCFI = 0.001. Intercepts were also equivalent between groups; Chi-square, RMSEA, and CFI tests demonstrated no differences in fit between the scalar and metric models, $\Delta\chi^2$ (17) = 11.90, $p = .806$, ΔRMSEA = .001, ΔCFI = 0.004. Thus, we compared latent means and found no significant differences between veterans and first responders on the MIDS overall factor score ($\Delta m = 0.02$, $p = .670$).

Tests of Reliability, Stability, and Descriptive Statistics

Next, we performed tests of reliability and stability. Internal consistency of MIDS scores was excellent for the full sample ($\alpha = .95$) and for veterans ($\alpha = .94$), healthcare workers ($\alpha = .95$), and first responders ($\alpha = .94$). Scores were moderately correlated across a 2-week interval among participants who were randomly selected to complete the MIDS follow-up survey ($r = .68$, $p < .001$, $n = 155$). We calculated descriptive statistics for the full sample ($M = 7.22$, $SD = 11.11$) and for veterans ($M = 7.76$, $SD = 11.00$), healthcare workers ($M = 7.04$, $SD = 11.56$), and first responders ($M = 6.88$, $SD = 10.67$). Independent samples t-tests revealed no evidence of crude mean differences in MIDS total scores between occupational subgroups (p's = .321–.860). Taken together, MIDS scores showed evidence of excellent internal consistency among relevant military and civilian populations and good stability across repeated administrations, while also being sensitive to potential fluctuations over time.

Tests of Convergent and Discriminant Validity

To evaluate convergent validity (Table 3), we examined associations between participants' scores on the MIDS and measures that purport to assess exposure to PMIEs. Scores on the MIDS were positively and strongly correlated with scores on the MIES ($r = .64$, $p < .001$) and EMIS ($r = .59$, $p < .001$). Associations between scores on the MIDS and measures of constructs theorized to be core components of moral injury, including trauma-related shame ($r = .68$, $p < .001$) and guilt ($r = .69$, $p < .001$), were positive and large. Similarly, scores on the MIDS were positively related to religious/spiritual struggles such as worry about moral wrongdoing ($r = .60$, $p < .001$), as well as feeling abandoned/punished by God/divine ($r = .35$, $p < .001$) and doubting one's beliefs ($r = .39$, $p < .001$) to a lesser extent.

Because prior research documents moderate and positive correlations between PMIE exposure and poor mental health (for a meta-analytic review, see McEwen et al., 2021), we expected to find an even stronger association between measures of morally injurious outcomes (i.e., the MIDS) and mental/behavioral health symptoms. Consistent with this theorizing, associations between scores on the MIDS and measures of depression ($r = .60$, $p < .001$) and posttraumatic stress ($r = .67$, $p < .001$) were positive and large. Associations between scores on the MIDS and behavioral health measures including insomnia ($r = .51$, $p < .001$) and hazardous alcohol use ($r = .18$, $p < .001$) were also positive and small to moderate in magnitude. Also, we observed a strong positive association between participants' scores on the MIDS and a multidomain composite of impaired psychosocial functioning ($r = .56$, $p < .001$) and a moderate negative association between scores on the MIDS and scores on a measure of general resilience ($r = -.40$, $p < .001$).

In terms of divergent validity, it is essential that moral injury be differentiated from other responses to traumatic events including PTSD. To that end, scores on the MIDS explained about 40% of the variation in scores on the aforementioned measure of posttraumatic stress (i.e., PCL-5; $r = .67$, $p < .001$) and about 25% of the variation on a measure of posttraumatic cognitions ($r = .53$, $p < .001$). These results suggest that moral injury and PTSD share some overlap, possibly to the extent that both occur in response to highly stressful events, though the two appear to result in unique cognitive interpretations and affective reactions. Furthermore, MIDS scores were only moderately associated with neuroticism ($r = .34$, $p < .001$), which suggests that the emotional and social sequelae of moral injury are not likely explained by a general tendency toward negative feelings or emotional instability.

Tests of Incremental Validity

Finally, to determine whether MIDS scores explained variance in outcomes beyond what was explained by existing measures of PMIE exposure, we conducted a hierarchical multiple regression with the severity of impairment in psychosocial functioning as the dependent variable. Independent variables included gender ($\beta = .05$, $p = .068$), age ($\beta = -.16$, $p < .001$), and minority race/ethnicity ($\beta = .03$, $p = .303$) entered in Step 1, EMIS ($\beta = .05$, $p = .209$) and MIES ($\beta = .17$, $p < .001$) total scores entered in Step 2, and MIDS ($\beta = .41$, $p < .001$) total scores entered in Step 3. The full model predicted 36.6% of the variance in functioning scores, $F(6, 911) = 87.69$, $p < .001$, $R^2 = .37$). After adjusting for gender, age, and race, adding EMIS and MIES scores explained 22.8% of the variance in functional impairment, $\Delta F(2, 912) = 143.62$, $p < .001$, $\Delta R^2 = .23$. MIDS scores explained an additional 9.0% of the variation in functioning, after accounting for variance explained by demographics and existing moral injury measures, $\Delta F(1, 911) = 129.58$, $p < .001$, $\Delta R^2 = .09$. When all independent variables were included in the model, MIDS scores uniquely explained 9.0% of variation in impairment versus the MIES and EMIS, which respectively explained 1.3% and <1.0% of the variance in impairment.

Discussion

The goal of this study was to develop and evaluate the psychometric properties and validity of the MIDS, the first measure designed to

Table 3
Bivariate Associations Among Study Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. MIDS	—															
2. MIES	.64*	—														
3. EMIS	.59*	.66*	—													
4. TRSI	.68*	.49*	.47*	—												
5. TRGI	.69*	.52*	.52*	.57*	—											
6. RSS divine	.35*	.27*	.25*	.40*	.31*	—										
7. RSS moral	.60*	.52*	.48*	.54*	.51*	.38*	—									
8. RSS doubt	.39*	.31*	.29*	.40*	.32*	.64*	.46*	—								
9. PHQ-9	.60*	.45*	.42*	.59*	.46*	.37*	.44*	.36*	—							
10. PCL-5	.67*	.54*	.50*	.62*	.55*	.40*	.48*	.39*	.74*	—						
11. PTCI	.53*	.47*	.45*	.55*	.47*	.33*	.46*	.38*	.52*	.50*	—					
12. ISI	.51*	.43*	.40*	.48*	.41*	.35*	.41*	.35*	.75*	.65*	.44*	—				
13. AUDIT	.18*	.18*	.18*	.19*	.15*	.40*	.19*	.16*	.16*	.17*	.15*	.15*	—			
14. CDRISC	−.40*	−.25*	−.32*	−.36*	−.36*	−.26*	−.30*	−.27*	−.43*	−.39*	−.34*	−.40*	−.03	—		
15. B-IPF	.56*	.46*	.42*	.54*	.43*	.37*	.50*	.40*	.69*	.65*	.54*	.58*	.19*	−.39*	—	
16. BIG 5-N	.34*	.25*	.31*	.32*	.34*	.25*	.28*	.26*	.37*	.39*	.35*	.34*	.11*	−.49*	.38*	—

Note. MIDS = Moral Injury and Distress Scale; MIES = Moral Injury Events Scale; EMIS = Expressions of Moral Injury Scale; TRSI = Trauma-Related Shame Inventory; TRGI = Trauma-Related Guilt Inventory; RSS Divine, Moral, and Doubt subscales = Religious & Spiritual Struggles Scale; PHQ-9 = Patient Health Questionnaire-9; PCL-5 = PTSD Symptom Checklist-5; PTSD = posttraumatic stress disorder; PTCI = Posttraumatic Cognitions Inventory; ISI = Insomnia Severity Index; AUDIT = Alcohol Use Disorders Identification Test; CD-RISC = Connor-Davidson Resilience Scale; B-IPF = Brief Inventory of Psychosocial Functioning; N = Big 5 Personality Dimensions of Neuroticism.
* $p < .001$.

assess both PMIE exposure and a comprehensive set of sequelae and symptoms and to measure moral injury exposures and outcomes across professions and populations. As hypothesized, the MIDS showed excellent internal consistency and good stability across repeated administrations. Associations of scores on the MIDS with other measures of moral injury and putative indicators of moral injury (e.g., guilt, shame) were positive and large, as were correlations with measures of the most closely related posttraumatic psychopathology, including PTSD symptoms, depression, insomnia, and psychosocial impairment. The MIDS also operated as expected with measures of discriminant validity (i.e., neuroticism). Examination of incremental validity showed that the MIDS was a stronger predictor of impaired functioning than the EMIS or MIES, uniquely explaining 9% of the variation in functioning, suggesting it has additional, unique explanatory utility over other commonly used moral injury measures.

While interest in moral injury has increased over the past decade, research has been hindered by limitations in our ability to effectively measure the construct. For example, because of the lack of appropriate moral injury measures, treatment studies have examined related outcomes such as changes in PTSD symptoms, guilt, or functioning (e.g., Litz et al., 2021; Norman et al., 2022), which obscure whether interventions are truly alleviating moral injury. The MIDS was designed to address these limitations. Based on the most widely accepted conceptual definition (Litz et al., 2009), the MIDS queries exposure to PMIEs, the extent to which one is bothered by those events, as well as emotional, cognitive, behavioral, and/or spiritual sequelae proposed to be common moral injury reactions. This structure allows for a clear delineation of events, symptoms, and other sequelae. It also allows clinicians and investigators to know whether reactions are linked to exposure to a PMIE or are associated with distress stemming from other types of trauma or stressors. The MIDS is also the first measure validated across military/civilian populations and types of PMIEs. Benefits of having such a measure include the ability to compare the prevalence and severity of moral injury across populations and PMIE types and being prepared to measure moral injury under a wide variety of circumstances. For example, at the start of the pandemic, there was much speculation about moral injury among healthcare providers, but no validated measures to assess prevalence. The MIDS allows us to quickly measure moral injury when events involving new populations or PMIEs occur.

Although moral injury has many facets (e.g., emotional, cognitive, behavioral), we found an underlying single factor that explained the majority of the total item variance. This is not surprising because all these facets relate to reactions associated with a specific PMIE; they are all parts of a unified response and thus manifest as a single factor. Perhaps if the scale were longer (e.g., ten items each of spiritual aspects, emotional aspects, etc.), we would have found a multidimensional factor structure similar to Koenig et al.'s (2018) 45-item Moral Injury Symptom Scale, but given the purposeful brevity of the MIDS, the single factor is not surprising. While our intention in developing the MIDS was to briefly and broadly assess the potential impacts of moral injury across domains identified by Litz et al. (2009), the impact of PMIE exposure on specific domains (e.g., religiousness/spirituality) may be better assessed using more tailored instrumentation.

Notably, an ongoing controversy in the literature has been about how betrayal relates to moral injury; specifically, whether betrayal is a PMIE, a common correlate of PMIE exposure, or a resultant symptom of moral injury following a PMIE (Griffin et al., 2019; Maguen & Norman, 2022). Carefully weighing past literature, consultation with other experts, and guided by Litz et al.'s (2009) and Shay's (2014) conceptual models, we included feelings of betrayal by authority as a moral injury symptom that can result from a PMIE. For example, a service member may see an authority figure make decisions that the service member believes puts others in harm's way (a "witnessing" PMIE) and as a result feel betrayed. We decided not to include more all-encompassing types of betrayal

included in some other measures because it is often not clear if those types of items are related to PMIEs or are feelings that result from other types of general stressors of negative life events (e.g., I feel betrayed by others outside of the military). Further highlighting the complicated role of betrayal in moral injury, findings reported in Table 2 show that participants were more likely to endorse "feeling betrayed by leaders or institutions" to a moderate or greater degree than any other item on the MIDS; however, the item was overall less strongly associated with the underlying latent factor representing moral injury in comparison to other items on the scale. Thus, betrayal may be a more ubiquitous experience and context in which PMIEs are likely to occur, though feeling betrayed in and of itself appears to be a poor indicator of the overall severity of morally injurious outcomes. Whether all betrayal from trauma should be considered part of the moral injury construct remains an important theoretical and empirical question.

Limitations

Given that moral injury is a relatively new and evolving construct, there may be alterations to this construct as additional research emerges. Consequently, the MIDS may need to be modified over time. Also, our sample was community-dwelling rather than treatment-seeking. While this optimizes the sample's generalizability to the targeted populations, because the measure is intended to be used in community and clinical settings, further research is needed to evaluate the psychometric properties and validity of the MIDS in treatment-seeking samples. While the MIDS is intended to be used to measure change (e.g., pre- to posttreatment), we were not able to collect data on change in this first study; thus, the MIDS' psychometric properties in measuring change are not yet known. More longitudinal research is needed using the MIDS, especially given that our repeated assessments occurred across a brief two-week time frame. Additionally, some healthcare workers were no longer working in the field. Unfortunately, we do not know the reason for this (e.g., retirement, burnout, etc.). While we included exclusively combat veterans, the majority were older and had served during the Vietnam era; a future goal would be to include a larger group of more recent combat veterans. Because individuals often internalize their moral beliefs from valued religious/spiritual groups and traditions, it is also critical that researchers examine how moral injury is impacted by religiousness/spirituality. Unfortunately, individuals who identified with religious/spiritual minority groups within the United States comprised very small percentages of our sample. While we developed the MIDS to be comprehensive in assessing a broad array of indicators considered hallmark symptoms of moral injury, the items may not adequately capture the multifactorial symptoms and impairments attributed to moral injury among every individual, group, or culture. Future research is therefore needed to continue refining and examining individuals' lived experiences with moral injury. Relatedly, the MIDS was validated with U.S. veterans, healthcare workers, and first responders; future research is needed to understand how moral injury may vary in other cultures.

Future Directions

Results of this study show that the MIDS has excellent internal consistency, good test–retest stability, excellent convergent and discriminant validity, and high incremental validity, suggesting that it is a psychometrically sound measure of moral injury that deserves further study. A psychometrically sound measure of morally injurious outcomes like the MIDS equips researchers to understand the severity and course of individuals' moral distress, with the goal of identifying those whose distress is likely to cause clinically significant functional deficits and therefore benefit from treatment. Next steps in evaluating the psychometric properties of the MIDS are to evaluate whether it can measure symptom change. This will require data on treatment-seeking samples, before and after treatment and ideally with posttreatment follow-up timepoints. In addition, the MIDS should be evaluated in other populations such as those who work in child protective services, border patrol agents, medical students, displaced populations, teachers, and those exposed to mass violence events such as school shootings. Similarly, the MIDS will need to be validated with international and diverse samples to ensure it is appropriate to use across cultures. Future work is also needed to identify if there are differences by race, ethnicity, or gender and other potentially contributing variables. Epidemiologic work can inform the prevalence of moral injury in the general population. The MIDS can be included in a wide variety of studies to accelerate our understanding of the prevalence and course of moral injury.

References

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Borges, L. M., Holliday, R., Barnes, S. M., Bahraini, N. H., Kinney, A., Forster, J. E., & Brenner, L. A. (2021). A longitudinal analysis of the role of potentially morally injurious events on COVID-19-related psychosocial functioning among healthcare providers. *PLoS ONE*, *16*(11), Article e0260033. https://doi.org/10.1371/journal.pone.0260033

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

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

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

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chesnut, R. P., Richardson, C. B., Morgan, N. R., Bleser, J. A., Mccarthy, K. J., & Perkins, D. F. (2022). The moral injury symptoms scale-military version-short form: Further scale validation in a U.S. veteran sample. *Journal of Religion and Health*, *61*(4), 3384–3401. https://doi.org/10.1007/s10943-022-01606-5

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor–Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, *18*(2), 76–82. https://doi.org/10.1002/da.10113

Currier, J. M., 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 and Psychotherapy*, *25*(3), 474–488. https://doi.org/10.1002/cpp.2170

Currier, J. M., Isaak, S. L., & McDermott, R. C. (2020). Validation of the expressions of Moral Injury Scale-Military version-short form. *Clinical Psychology and Psychotherapy*, *27*(1), 61–68. https://doi.org/10.1002/cpp.2407

Eisinga, R., Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman–Brown? *International Journal of Public Health*, *58*(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

Exline, J. J., Pargament, K. I., Grubbs, J. B., & Yali, A. M. (2014). The Religious and Spiritual Struggles Scale: Development and initial validation. *Psychology of Religion and Spirituality*, *6*(3), 208–222. https://doi.org/10.1037/a0036465

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling*, *12*(3), 343–367. https://doi.org/10.1207/s15328007sem1203_1

Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. F., & Orsillo, S. 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

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, *21*(4), 572–583. https://doi.org/https://doi.org/10.1037/0893-3200.21.4.572

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

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191–205. https://doi.org/10.1177/1094428104263675

Hines, S. E., Chin, K. H., Glick, D. R., & Wickwire, E. M. (2021). Trends in moral injury, distress, and resilience factors among healthcare workers at the beginning of the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, *18*(2), Article 488. https://doi.org/10.3390/ijerph18020488

IBM Corp. (2020). *IBM SPSS statistics for windows, Version 27.0*.

Jinkerson, J. D. (2016). Defining and assessing moral injury: A syndrome perspective. *Traumatology*, *22*(2), 122–130. https://doi.org/10.1037/trm0000069

Jordan, A. H., Eisen, E., Bolton, E., Nash, W. P., & Litz, B. T. (2017). Distinguishing war-related PTSD resulting from perpetration- and betrayal-based morally injurious events. *Psychological Trauma: Theory, Research, Practice, and Policy*, *9*(6), 627–634. https://doi.org/10.1037/tra0000249

Khan, A. J., Nishimi, K., Tripp, P., Maven, D., Jiha, A., Woodward, E., Inslicht, S., Richards, A., Neylan, T. C., Maguen, S., & O'Donovan, A. (2021). COVID-19 related moral injury: Associations with pandemic-related perceived threat and risky and protective behaviors. *Journal of Psychiatric Research*, *142*, 80–88. https://doi.org/10.1016/j.jpsychires.2021.07.037

Kleiman, S. E., Bovin, M. J., Black, S. K., Rodriguez, P., Brown, L. G., Brown, M. E., Lunney, C. A., Weathers, F. W., Schnurr, P. P., Spira, J., Keane, T. M., & Marx, B. P. (2020). Psychometric properties of a brief measure of posttraumatic stress disorder-related impairment: The Brief Inventory of Psychosocial Functioning. *Psychological Services*, *17*(2), 187–194. https://doi.org/10.1037/ser0000306

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. Guilford Press.

Koenig, H. G., Ames, D., Youssef, N. A., Oliver, J. P., Volk, F., Teng, E. J., Haynes, K., Erickson, Z. D., Arnold, I., O'Garo, K., & Pearce, M. (2018). The Moral Injury Symptom Scale—Military version. *Journal of Religion and Health*, *57*(1), 249–265. https://doi.org/10.1007/s10943-017-0531-9

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*, Article 443. https://doi.org/10.3389/fpsyt.2019.00443

Kroenke, K., Spitzer, R. L., & Williams, J. B. (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

Kubany, E. S., Haynes, S. N., Abueg, F. R., Manke, F. P., Brennan, J. M., & Stahura, C. (1996). Development and validation of the Trauma-Related Guilt Inventory (TRGI). *Psychological Assessment*, *8*(4), 428–444. https://doi.org/10.1037/1040-3590.8.4.428

Lentz, L. M., Smith-MacDonald, L., Malloy, D., Carleton, R. N., & Brémault-Phillips, S. (2021). Compromised conscience: A scoping review of moral injury among firefighters, paramedics, and police officers. *Frontiers in Psychology*, *12*, Article 639781. https://doi.org/10.3389/fpsyg.2021.639781

Litz, B. T., & Kerig, P. K. (2019). Introduction to the special issue on moral injury: Conceptual challenges, methodological issues, and clinical applications. *Journal of Traumatic Stress*, *32*(3), 341–349. https://doi.org/10.1002/jts.22405

Litz, B. T., Plouffe, R. A., Nazarov, A., Murphy, D., Phelps, A., Coady, A., Houle, S. A., Dell, L., Frankfurt, S., Zerach, G., Levi-Belz, Y., Ashbaugh, A., Yeterian, J., Berke, D., Carney, J., Cummings, M., Chartoff, R., Ellickson-Larew, S., … Murray, K. (2022). Defining and assessing the syndrome of moral injury: Initial findings of the Moral Injury Outcome Scale Consortium. *Frontiers in Psychiatry*, *13*, Article 923928. https://doi.org/10.3389/fpsyt.2022.923928

Litz, B. T., Rusowicz-Orazem, L., Doros, G., Grunthal, B., Gray, M., Nash, W., & Lang, A. J. (2021). Adaptive disclosure, a combat-specific PTSD treatment, versus cognitive-processing therapy, in deployed marines and sailors: A randomized controlled non-inferiority trial. *Psychiatry Research*, *297*, Article 113761. https://doi.org/10.1016/j.psychres.2021.113761

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

Maguen, S., Griffin, B. J., Copeland, L. A., Perkins, D. F., Finley, E. P., & Vogt, D. (2020). Gender differences in prevalence and outcomes of exposure to potentially morally injurious events among post-9/11 veterans. *Journal of Psychiatric Research*, *130*, 97–103. https://doi.org/https://doi.org/10.1016/j.jpsychires.2020.06.020

Maguen, S., & Norman, S. B. (2022). Moral injury. *PTSD Research Quarterly*, *33*(1). https://www.ptsd.va.gov/publications/rq_docs/V33N1.pdf

McEwen, C., Alisic, E., & Jobson, L. (2021). Moral injury and mental health: A systematic review and meta-analysis. *Traumatology*, *27*(3), 303–315. https://doi.org/10.1037/trm0000287

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: Psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*, *34*(5), 601–608. https://doi.org/10.1093/sleep/34.5.601

Muthén, B. O., & Muthén, L. K. (2017). *Mplus user's guide* (Version 8.1) [Computer software]. https://www.statmodel.com/

Nash, W. P., Marino Carper, T. L., Mills, M. A., Au, 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

Norman, S. B., Feingold, J. H., Kaye-Kauderer, H., Kaplan, C. A., Hurtado, A., Kachadourian, L., Feder, A., Murrough, J. W., Charney, D.,

Southwick, S. M., Ripp, J., Peccoralo, L., & Pietrzak, R. H. (2021). Moral distress in frontline healthcare workers in the initial epicenter of the COVID-19 pandemic in the United States: Relationship to PTSD symptoms, burnout, and psychosocial functioning. *Depression and Anxiety*, *38*(10), 1007–1017. https://doi.org/10.1002/da.23205

Norman, S. B., Nichter, B., Maguen, S., Na, P. J., Schnurr, P. P., & Pietrzak, R. H. (2022). Moral injury among U.S. Combat veterans with and without PTSD and depression. *Journal of Psychiatric Research*, *154*, 190–197. https://doi.org/10.1016/j.jpsychires.2022.07.033

Øktedalen, T., Hagtvet, K. A., Hoffart, A., Langkaas, T. F., & Smucker, M. (2014). The trauma related shame inventory: Measuring trauma-related shame among patients with PTSD. *Journal of Psychopathology and Behavioral Assessment*, *36*(4), 600–615. https://doi.org/10.1007/s10862-014-9422-5

Purcell, N., Koenig, C. J., Bosch, J., & Maguen, S. (2016). Veterans' perspectives on the psychosocial impact of killing in war. *The Counseling Psychologist*, *44*(7), 1062–1099. https://doi.org/10.1177/0011000016666156

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Richardson, C. B., Chesnut, R. P., Morgan, N. R., Bleser, J. A., Perkins, D. F., Vogt, D., Copeland, L. A., & Finley, E. (2020). Examining the factor structure of the moral injury events scale in a veteran sample. *Military Medicine*, *185*(1-2), e75–e83. https://doi.org/https://doi.org/10.1093/milmed/usz129

Riedel, P.-L., Kreh, A., Kulcar, V., Lieber, A., & Juen, B. (2022). A scoping review of moral stressors, moral distress and moral injury in healthcare workers during COVID-19. *International Journal of Environmental Research and Public Health*, *19*(3), Article 1666. https://doi.org/10.3390/ijerph19031666

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption—II. *Addiction*, *88*(6), 791–804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x

Shay, J. (2014). Moral injury. *Psychoanalytic Psychology*, *31*(2), 182–191. https://doi.org/10.1037/a0036090

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/BF02291170

Williamson, V., Murphy, D., & Greenberg, N. (2020). COVID-19 and experiences of moral injury in front-line key workers. *Occupational Medicine (Oxford, England)*, *70*(5), 317–319. https://doi.org/10.1093/occmed/kqaa052

Zhizhong, W., Koenig, H. G., Yan, T., Jing, W., Mu, S., Hongyu, L., & Guangtian, L. (2020). Psychometric properties of the moral injury symptom scale among Chinese health professionals during the COVID-19 pandemic. *BMC Psychiatry*, *20*(1), Article 556. https://doi.org/10.1186/s12888-020-02954-w

Received September 10, 2022
Revision received April 3, 2023
Accepted April 21, 2023 ■