ORIGINAL ARTICLE

WILEY

Design and methods of the Longitudinal Eating Disorders Assessment Project research consortium for veterans

Kelsie T. Forbush[1] | Trevor J. Swanson[1] | Melinda Gaddy[2] | Mary Oehlert[2] | Alesha Doan[1] | Robert W. Morgan[1] | Colin O'Brien[1] | Yiyang Chen[1] | Kylie Christian[1] | Q. Chelsea Song[3] | David Watson[4] | Joanna Wiese[5]

[1]University of Kansas, Lawrence, Kansas, USA

[2]VA Eastern Kansas Health Care System, Leavenworth, Kansas, USA

[3]Purdue University, West Lafayette, Indiana, USA

[4]University of Notre Dame, Notre Dame, Indiana, USA

[5]20th Medical Group, Shaw Air Force Base, South Carolina, USA

Correspondence
Kelsie T. Forbush, Department of Psychology, University of Kansas, 1415 Jayhawk Blvd., Lawrence, KS 66045, USA.
Email: kforbush@ku.edu

Abstract

Introduction: Military service members must maintain a certain body mass index and body fat percentage. Due to weight-loss pressures, some service members may resort to unhealthy behaviors that place them at risk for the development of an eating disorder (ED).

Objectives: To understand the scope and impact of EDs in military service members and veterans, we formed the Longitudinal Eating Disorders Assessment Project (LEAP) Consortium. LEAP aims to develop novel screening, assessment, classification, and treatment tools for veterans and military members with a focus on EDs and internalizing psychopathology.

Methods: We recruited two independent nationally representative samples of post-9/11 veterans who were separated from service within the past year. Study 1 was a four-wave longitudinal survey and Study 2 was a mixed-methods study that included surveys, structured-clinical interviews, and qualitative interviews.

Results: Recruitment samples were representative of the full population of recently separated veterans. Sample weights were created to adjust for sources of non-response bias to the baseline survey. Attrition was low relative to past studies of this population, with only (younger) age predicting attrition at 1-week follow-up.

Conclusions: We expect that the LEAP Consortium data will contribute to improved information about EDs in veterans, a serious and understudied problem.

KEYWORDS
assessment, department of defense, disordered eating, eating disorders, longitudinal, mental health, nationally representative, psychiatric disorders, VADIR, veterans

Eating disorders (EDs) are serious psychiatric disorders that negatively impact physical and mental health (Brown & Mehler, 2013; Forney et al., 2016), emotional wellbeing, work productivity, and social relationships (Streatfeild et al., 2021). EDs have a combined prevalence of 13.1%–15% among young women (Allen et al., 2013; Stice et al., 2013) and 3% among young men in the general population

(Allen et al., 2013), which is similar to the prevalence of depression or anxiety. EDs affect people of all ethnicities and socioeconomic levels, and occur across the lifespan (Mitchison et al., 2014; van Hoeken & Hoek, 2020). EDs have one of the highest mortality rates among psychiatric disorders (Fichter & Quadflieg, 2016; Iwajomo et al., 2021), with standardized mortality ratios that are two times higher than mortality due to heavy smoking (Chesney et al., 2014). Moreover, EDs are associated with suicide-related mortality in the USA, with suicide rates similar to (or higher than) depression, opioid addiction, or schizophrenia (Chesney et al., 2014). Due to the effects of EDs on general physical health, civilians with EDs incur additional healthcare costs associated with medical events (e.g., myocardial infarction) or due to comorbid health conditions, such as obesity, which is more common among people with EDs than the general population (Striegel Weissman & Rosselli, 2017). Similar to estimates in civilian populations, veteran patients with an ED incurred 1-year adjusted total healthcare costs that were $18,152 higher than matched veterans without an ED (Striegel-Moore et al., 1999), suggesting that there is a critical need to identify and treat EDs in veterans.

Veteran men and women appear to be at risk for the development of an ED (Cuthbert et al., 2020; Masheb et al., 2021). For example, veterans have elevated rates of disordered eating, overweight, and obesity compared to non-veteran civilians (Cuthbert et al., 2020). Although there is a wide prevalence range, recent studies suggest that 14.1%–32.8% of veteran women and 4.1%–18.8% of veteran men have a current ED (Ardittte Hall et al., 2017; Forman-Hoffman et al., 2012; Masheb et al., 2021). Military personnel must adhere to body mass index (BMI) and body fat percentage requirements, as well as meet physical performance standards that are assessed every six months. Failure to maintain these standards can result in referral to weight-loss programs and, eventually, discharge from service. The emphasis on body weight within the military may promote unhealthy attempts at weight loss and the development of eating disorders (EDs). For example, being on a weight loss diet while in the military is associated with an increased risk for ED symptoms (Mitchell et al., 2021). If left untreated, EDs result in significantly greater Veterans Affairs (VA) healthcare utilization and physical and psychiatric morbidity and mortality.

Despite the public-health importance of addressing EDs in active-duty and veteran populations, few nationally representative studies of veterans have included ED and assessments of related comorbid psychopathology. To address the need to better understand the scope and impact of EDs in military service members and veterans, the Department of Defense's Congressionally Directed Medical Research Program (CDMRP) included EDs as an Area of Encouragement for the first time in 2017. Since 2017, the CDMRP has continued to fund studies on EDs, which led to the development of the Longitudinal Eating Disorders Assessment Project (LEAP) research consortium. LEAP brings together researchers and clinicians across state lines to develop novel screening, assessment, classification, and treatment tools for veterans and military members with a focus on EDs and other forms of internalizing psychopathology. Our

consortium also includes researchers with expertise in qualitative methods and public policy, with a focus on identifying formal and informal policies and practices that can potentially inform ED treatment implementation efforts in military-relevant populations.

In this paper, we report the design, protocol, and methods of the first two datasets from the LEAP research consortium. We also report initial attrition analyses from the first wave of our longitudinal dataset (Study 1). As we describe in the paper, other data waves from Study 1 and Study 2 were ongoing at the time of this writing. Our primary objective was to develop and validate a military-specific transdiagnostic screening tool for eating and related disorders that can eventually be implemented in healthcare settings. The scientific premise that guided our research was that an integrated transdiagnostic framework for assessing and classifying EDs could lead to improvements in the ability to identify veterans who are most at-risk for poor post-discharge psychosocial adjustment and who need a referral to treatment services. Our primary aim was to validate/develop a transdiagnostic ED screening tool. Secondary aims were to (1) evaluate longitudinal relationships among ED symptoms, internalizing, and externalizing psychopathology, and (2) identify barriers to obtaining ED treatment during military service.

We recruited two independent nationally representative samples of veterans from the VA/DoD Identity Repository (VADIR), a VA office with access to DoD records and the only comprehensive registry of military personnel and veterans in the USA. Study 1 was a four-wave longitudinal survey-based study that assessed veterans at baseline, 1-week, 3-months, and 6-months. Study 2 was a cross-sectional study that included surveys, diagnostic interviews, and qualitative interviews. Although the first two studies from the LEAP consortium were primarily focused on EDs and internalizing disorders, the datasets include a wide range of demographic, military-service related, and other psychopathology constructs that may be of interest to general psychiatric researchers. Given the DoD's focus on publicly available data and our own commitment to open-science practices, de-identified data (with personal information removed) will be released to the public when it becomes available in compliance with the General Data Protection Regulation. Thus, the goal of this paper is to provide information on the design, protocol, methods for LEAP studies and initial attrition analyses for the first wave of the Study 1 LEAP dataset.

1 | METHODS

1.1 | Ethics approval

Ethical approval for the project was provided by the University of Kansas, VA Eastern Kansas Health Care System (VAEKHCS), and the Department of Defense Human Research Protections Office. All participants provided informed consent prior to engaging in any study-related procedures. The University of Kansas received a HIPAA waiver to recruit veterans who separated from service within the past year from VADIR.

1.2 | Study design overview

Study 1 is a four-wave longitudinal study that included online and paper-based surveys. Participants were assessed at baseline, 1-week follow-up, 3-month follow-up, and 6-month follow-up. Study 2 is a cross-sectional study that includes self-report surveys, structured clinical interviews, and qualitative interviews. In Study 2, veterans will be recruited to participate in a 1-h structured clinical interview that assesses current and lifetime EDs, mood disorders, generalized anxiety disorder, and post-traumatic stress disorder. Veterans who were diagnosed with an ED during the structured clinical interview will be invited to participate in an additional qualitative interview. Veterans with EDs will be recruited until we achieve our target sample size for qualitative interviews ($N = 50$). Qualitative methods will include semi-structured, open-ended, in-depth interviews that range between 45 and 60 min. These exploratory interviews are designed to identify institutional barriers or informal practices in the military that created obstacles for participants seeking ED and trauma-related care during active duty. Qualitative interviews will be piloted with 10% of the total sample ($N = 5$) to assess the validity of the questions and modify them according to participants' feedback. Interviews will be recorded and transcribed verbatim. Transcripts will be coded in-vivo (i.e., using the words of participants) whenever possible, using an inductive, open coding format to catalog dominant and frequent domains (Charmaz, 2006). Taxonomic analysis, which entails developing a classification system to synthesize data, will be used to inventory each domain (Bradley et al., 2007).

1.3 | Enrollment status

Baseline and 1-week follow-up recruitment for Study 1 is complete, whereas the 3- and 6-month follow-ups for Study 1 are ongoing. Study 2 launched in March 2022 and recruitment is ongoing (current participant numbers are reported in Table 1).

1.3.1 | VADIR data extraction

Based on power calculations, anticipated attrition, and expected response rates, our Study 1 target sample size was $N = 1000$ veterans at baseline. For Study 2, our target sample size was $N = 400$. To achieve our targeted N's, we recruited two separate, non-overlapping nationally representative samples of veterans from the VADIR who separated from service within the past year. Our rationale for recruiting veterans who discharged within the past year was to collect information as temporally close to their military service as possible to help minimize retrospective recall biases about military-relevant experiences. VADIR has been used in past studies of mental health to obtain nationally representative samples of veterans

TABLE 1 Study 1 baseline sample characteristics

| | | Study 1 | | | | Study 2 | | | |
| | | Females | | Males | | Females | | Males | |
		VADIR	Sample	VADIR	Sample	VADIR	Sample	VADIR	Sample
Branch	Army	11938 (48.02%)	3654 (50.93%)	54868 (47.77%)	3648 (49.8%)	12862 (58.53%)	2293 (62.62%)	52482 (53.88%)	2054 (56.21%)
	Air Force	4203 (16.91%)	1236 (17.23%)	14583 (12.7%)	922 (12.59%)	3521 (16.02%)	607 (16.58%)	12161 (12.48%)	454 (12.42%)
	Marines	1790 (7.2%)	528 (7.36%)	22105 (19.25%)	1461 (19.95%)	1231 (5.6%)	229 (6.25%)	17211 (17.67%)	702 (19.21%)
	Navy	6931 (27.88%)	1757 (24.49%)	23300 (20.29%)	1294 (17.67%)	4361 (19.85%)	533 (14.55%)	15551 (15.97%)	444 (12.15%)
Personnel Category	Active Duty	9052 (36.41%)	2308 (32.17%)	34438 (29.98%)	1978 (27%)	8554 (38.93%)	1268 (34.63%)	37282 (38.28%)	1279 (35%)
	National Guard	3208 (12.9%)	970 (13.52%)	14301 (12.45%)	938 (12.81%)	4368 (19.88%)	690 (18.84%)	15285 (15.69%)	553 (15.13%)
	Reserve	12602 (50.69%)	3897 (54.31%)	66117 (57.57%)	4409 (60.19%)	9053 (41.2%)	1704 (46.53%)	44838 (46.03%)	1822 (49.86%)
Rank/Pay Category	Enlisted	16657 (67%)	4773 (66.52%)	74266 (64.66%)	4778 (65.23%)	17754 (80.79%)	2878 (78.59%)	76242 (78.27%)	2804 (76.74%)
	Non-Commissioned Officer	6657 (26.78%)	1952 (27.21%)	35660 (31.05%)	2258 (30.83%)	3213 (14.62%)	598 (16.33%)	17925 (18.4%)	730 (19.98%)
	Commissioned Officer	1530 (6.15%)	447 (6.23%)	4745 (4.13%)	280 (3.82%)	1001 (4.56%)	185 (5.05%)	3149 (3.23%)	118 (3.23%)
	Warrant Officer	18 (0.07%)	3 (0.04%)	185 (0.16%)	9 (0.12%)	7 (0.03%)	1 (0.03%)	89 (0.09%)	2 (0.05%)
Race	White	15629 (62.86%)	4591 (63.99%)	87929 (76.56%)	5715 (78.02%)	14249 (64.84%)	2461 (67.2%)	74759 (76.75%)	2898 (79.31%)
	Native Hawaiian/Pacific Islander	150 (0.6%)	33 (0.46%)	495 (0.43%)	19 (0.26%)	125 (0.57%)	16 (0.44%)	427 (0.44%)	9 (0.25%)
	Asian	1127 (4.53%)	315 (4.39%)	4852 (4.22%)	290 (3.96%)	910 (4.14%)	142 (3.88%)	3872 (3.98%)	129 (3.53%)
	Black or African American	5957 (23.96%)	1711 (23.85%)	15347 (13.36%)	960 (13.11%)	5275 (24%)	854 (23.32%)	13798 (14.17%)	497 (13.6%)
	American Indian or Alaskan Native	414 (1.67%)	105 (1.46%)	1437 (1.25%)	60 (0.82%)	314 (1.43%)	41 (1.12%)	1104 (1.13%)	27 (0.74%)
	Multiracial or Unknown Race	1585 (6.38%)	420 (5.85%)	4796 (4.18%)	281 (3.84%)	1102 (5.01%)	148 (4.04%)	3445 (3.54%)	94 (2.57%)
Ethnicity	Hispanic	3876 (15.59%)	1071 (14.93%)	15396 (13.4%)	946 (12.91%)	3258 (14.83%)	515 (14.06%)	12805 (13.15%)	446 (12.21%)
	Non-Hispanic	20986 (84.41%)	6104 (85.07%)	99460 (86.6%)	6379 (87.09%)	18717 (85.17%)	3147 (85.94%)	84600 (86.85%)	3208 (87.79%)
Deployment Status	Non-Deployed	17217 (69.25%)	4954 (69.05%)	61524 (53.57%)	3885 (53.04%)	18407 (83.76%)	2972 (81.16%)	69781 (71.64%)	2495 (68.28%)
	Deployed	7645 (30.75%)	2221 (30.95%)	53332 (46.43%)	3440 (46.96%)	3568 (16.24%)	690 (18.84%)	27624 (28.36%)	1159 (31.72%)
Age at Discharge	1st Quartile	6817 (27.42%)	1760 (24.53%)	34257 (29.83%)	2062 (28.15%)	6814 (31.01%)	853 (23.29%)	25358 (26.03%)	755 (20.66%)
	2nd Quartile	7618 (30.64%)	2289 (31.9%)	30550 (26.6%)	2041 (27.86%)	4245 (19.32%)	698 (19.06%)	34301 (35.21%)	1348 (36.89%)
	3rd Quartile	4849 (19.5%)	1428 (19.9%)	24697 (21.5%)	1578 (21.54%)	6026 (27.42%)	1163 (31.76%)	14407 (14.79%)	582 (15.93%)
	4th Quartile	5578 (22.44%)	1698 (23.67%)	25352 (22.07%)	1644 (22.44%)	4890 (22.25%)	948 (25.89%)	23339 (23.96%)	969 (26.52%)
Total Size		24862 (17.79%)	7175 (49.48%)	114856 (82.21%)	7325 (50.52%)	21975 (18.41%)	3662 (50.05%)	97405 (81.59%)	3654 (49.95%)

Note: This table displays the frequencies and percentages associated with the levels of each demographic variable used in the stratification procedure. For each study, stratification was performed separately for males and females to sample a 1:1 ratio. The meaning of the quartiles for Age at Discharge differed depending on the group. In Study 1, the quartiles for females included ages: (1) 17-24, (2) 25-27, (3) 28-30, (4) 31-66, whereas the quartiles for males included ages: (1) 17-25, (2) 26-27, (3) 28-30, (4) 31-68. In Study 2, the quartiles for females included ages: (1) 17-20, (2) 21-24, (3) 25-28, (4) 29-66, whereas the quartiles for males included ages: (1) 17-21, (2) 22-26, (3) 27-28, (4) 29-79. Note that age-at-discharge quartiles are not exactly 25% due to rounding.

(Aronson et al., 2019; Barth et al., 2014; Eber et al., 2013). This repository contains variables related to veterans' military service, demographic characteristics, and contact information. Demographic data included veterans' name, date of birth (used to calculate *age*), service start and end dates (to verify eligibility), address, telephone number, *race*, *ethnicity*, *sex*, whether the veteran was *deployed* post-9/11, *branch* of service (Army, Navy, etc.), *pay* category (enlisted, non-commissioned officer, officer, or warrant officer), *personnel* category (reserve, active duty, or national guard) and military *rank*. We requested two random samples of veterans (i.e., one sample for each study), to ensure that each sample was independent from the other sample. Date of birth and service end dates were used to calculate age at discharge.

1.3.2 | Data sharing and accessibility

De-identified and unclassified data collected from this study will be made publicly available through the Center for Open Science (OSF). Our procedures for sharing data will be similar to those used in other large-scale mental-health studies, such as the Minnesota Twin Family Study (Iacono et al., 2006). Investigators interested in using the data should contact the Principal Investigator (KTF) to provide a brief description of the intended use and proposed project timeline. These project descriptions will eventually be posted on OSF. Investigators will also be required to sign a data-use agreement. Such documentation will prevent duplication of research and provide a greater level of protection of intellectual property.

1.3.3 | Stratification procedures

Study 1
For Study 1, we obtained data for all 179,111 post-9/11 veterans separated from military service between October 1, 2018 and September 30, 2019. From this population, we selected a random sample of 4500 veterans stratified by *race*, *ethnicity*, *personnel category*, *rank* (which we operationalized through pay grade), *branch*, *age at discharge*, and *deployment status*. We oversampled women to recruit a 1:1 ratio of men to women to achieve an adequate sample of women. We achieved an equal sex ratio by dividing the full population into a group of women and a group of men, and then performed the stratification process separately for each group. To perform the stratification, we developed an R package called "sampleVADIR" (Swanson & Forbush, 2021). The stratification procedure implemented by the sampleVADIR function is described below.

First, we removed individuals from the dataset for whom ethnicity, address, or zip code was missing, as well as individuals whose rank was "ROTC." This left a total of 139,718 individuals in the population from which to sample (24,862 women, and 114,856 men). Before beginning the stratification process, we modified the *rank* and *age at discharge* variables into alternative formats to better suit the stratification algorithm. Given that there were 93 possible values for

the rank variable (which created too many possible groups to stratify), we converted the rank values into four categories corresponding to pay grades. Specifically, we created groups for enlisted individuals (E), non-commissioned officers (NCO), warrant officers (W), and commissioned officers (O). Given that *age at discharge* was defined as a continuous variable, we created four groups associated with quartiles observed in the population. The four quartiles were defined by the ages: (1) 17–25, (2) 25–27, (3) 28–30, and (4) 31–68. The first two quartiles differed slightly for men and women: for women, the upper limit of the first quartile was 24 and, for men, the lower limit of the second quartile was 26. A breakdown of the *age at discharge* quartiles for both studies is provided in Table 2.

After preprocessing, the "sampleVADIR" package computed the number of unique strata defined by every combination of the seven variables used in the stratification. For women, there were 1305 unique strata. The proportion of each stratum within the population of women was then computed and multiplied by the number of women to be sampled; this value was then rounded to the nearest integer. This procedure resulted in a sample of 4500 individuals ($n = 2258$ women and $n = 2242$ men).

Unexpectedly, approximately 48% of our initial 4500 recruitment mailings were returned due to invalid addresses. Due to the high proportion of invalid addresses, we used the same stratification procedures to select an additional 14,750 cases from the original Study 1 VADIR data. To identify individuals with invalid addresses, we began mailing postcards to let veterans know they would receive an invitation letter. After implementing this procedure, 24% of 14,750 (3546) had invalid addresses. From the 11,204 valid addresses, we selected 10,000 by randomly removing members of overrepresented demographic groups, which mostly included randomly removing veterans who identified as White. Overrepresented groups were identified as those strata that had larger representation in the reduced sample, proportionally, than in the original sample of 14,750. Combining across all Study 1 initial mailings, our incorrect address rate was 38.56%. Of the remaining 8909 individuals that had valid addresses, 2.55% declined to participate and 87.16% did not respond to any mailings. Thus, our rate of enrollment from participants who received mail was 10.29%. Finally, 642 participants completed the 1-week survey, which was a 70.01% retention rate (or 29.99% attrition rate at 1-week follow-up). Demographic and military service characteristics of the individuals who were selected for recruitment are provided in Tables 1 and 2 (4500 from the initial sample and 10,000 from the additional sample).

Study 2
The VADIR variables requested for Study 2 were identical to those requested for Study 1 (see the previous section). For Study 2, we attained a representative sample of 7316 cases from the 119,380 individuals separated from military service between October 1, 2019 and September 30, 2020. Table 2 contains descriptive information regarding the demographic and military service characteristics associated with the total sample of 7316 individuals selected for

TABLE 2 Study 1 baseline sample characteristics

		VADIR Sample	Baseline	Weighted
Sex	Female	7175 (49.48%)	507 (62.13%)	404 (49.51%)
	Male	7325 (50.52%)	309 (37.87%)	412 (50.49%)
Branch	Army	7302 (50.36%)	437 (53.55%)	411 (50.37%)
	Air Force	2158 (14.88%)	122 (14.95%)	121 (14.83%)
	Marines	1989 (13.72%)	91 (11.15%)	112 (13.73%)
	Navy	3051 (21.04%)	166 (20.34%)	172 (21.08%)
Personnel Category	Active Duty	4286 (29.56%)	217 (26.59%)	241 (29.53%)
	National Guard	1908 (13.16%)	110 (13.48%)	107 (13.11%)
	Reserve	8306 (57.28%)	489 (59.93%)	467 (57.23%)
Rank/Pay Category	Enlisted	9551 (65.87%)	427 (52.33%)	537 (65.81%)
	Non-Commissioned Officer	4210 (29.03%)	309 (37.87%)	237 (29.04%)
	Commissioned Officer	727 (5.01%)	77 (9.44%)	41 (5.02%)
	Warrant Officer	12 (0.08%)	3 (0.37%)	1 (0.12%)
Race	White	10306 (71.08%)	629 (77.08%)	580 (71.08%)
	Native Hawaiian/Pacific Islander	52 (0.36%)	1 (0.12%)	3 (0.37%)
	Asian	605 (4.17%)	36 (4.41%)	34 (4.17%)
	Black or African American	2671 (18.42%)	98 (12.01%)	150 (18.38%)
	American Indian or Alaskan Native	165 (1.14%)	10 (1.23%)	9 (1.1%)
	Multiracial or Unknown Race	701 (4.83%)	42 (5.15%)	39 (4.78%)
Ethnicity	Hispanic	2017 (13.91%)	97 (11.89%)	114 (13.97%)
	Non-Hispanic	12483 (86.09%)	719 (88.11%)	702 (86.03%)
Deployment Status	Non-Deployed	8839 (60.96%)	452 (55.39%)	497 (60.91%)
	Deployed	5661 (39.04%)	364 (44.61%)	319 (39.09%)
Age at Discharge	1st Quartile	3822 (26.36%)	127 (15.56%)	215 (26.35%)
	2nd Quartile	4330 (29.86%)	237 (29.04%)	244 (29.9%)
	3rd Quartile	3006 (20.73%)	182 (22.3%)	169 (20.71%)
	4th Quartile	3342 (23.05%)	270 (33.09%)	188 (23.04%)
Total Size		14500	816	816

Note: Shows the frequencies and percentages associated with the levels of each demographic variable used in the stratification procedure. For each study, stratification was performed separately for males and females to sample the sexes with a 1:1 ratio. An implication of this was that the meaning of the quartiles for Age at Discharge differed depending on the group. In Study 1, the quartiles for females included ages: (1) 17-24, (2) 25-27, (3) 28-30, (4) 31-66, while the quartiles for males included ages: (1) 17-25, (2) 26-27, (3) 28-30, (4) 31-68. In Study 2, the quartiles for females included ages: (1) 17-20, (2) 21-24, (3) 25-28, (4) 29-66, while the quartiles for males included ages: (1) 17-21, (2) 22-26, (3) 27-28, (4) 29-79.

recruitment for Study 2. The same stratification strategy used to obtain the Study 1 recruitment sample was used to obtain the Study 2 recruitment sample.

1.4 | Description of recruitment procedures and materials

After receiving approval from our local and national VA privacy officers, VAEKHCS shared the stratified sample data with the University of Kansas through encrypted email. The stratified sample data were saved on a dual-authenticated research server at the University of Kansas, and veterans listed in the sample were sent recruitment mailings.

We used a modified version of Dillman's Tailored Design Method (Dillman, 2011; Hoddinott & Bass, 1986) to recruit participants. This three-stage procedure has been shown to maximize responses to survey-based research (Dillman et al., 2009). Across this three-stage procedure, a total of five mailings were sent to participants. Past studies that used the Tailored Design Method to recruit veterans from VADIR achieved response rates of 25%–44% (Coughlin et al., 2011). An initial letter (or postcard), brochure, and decline-to-participate card invited veterans to participate in an online Qualtrics® study (Stage 1). The initial letter also included a $1.00

incentive payment—regardless of whether the invited veteran chose to participate in the study—because this technique has been shown to increase survey response rates in previous research (Coughlin et al., 2011). We attempted to send non-responders a reminder letter two weeks after the initial mailing (Stage 2); however, due to printing delays and COVID-19 pandemic-related labor shortages at the U.S. postal service that impacted mail delivery speed, the time between each reminder varied (see Table S1). Next, non-responders were sent up to three mailings (Stage 3) that contained a reminder letter, two informed consent documents, decline postcard, study questionnaires, and a pre-paid return envelope. On August 12, 2021 the contents of reminder mailing packets were adjusted due to printing and postage costs. After this date, the first two reminder mailings contained a letter and decline postcard only. However, the final (third) reminder included a letter, two informed consent documents, decline postcard, study questionnaires, and a pre-paid return envelope.

1.5 | Measures

A list of measures is included in Table 3. Internal consistency for each measure is provided in Table 4. Unless otherwise noted, measures were administered in both Studies 1 and 2. Study measures are described below:

TABLE 3 Longitudinal Eating Disorders Assessment Project assessment schedule

Construct	Assessment(s)	Study 1 Baseline	One Week	Three Month	Six Month	Study 2
Demographics and military history	**Ad Hoc questionnaire adapted from NSV**: Demographics, Background questions	X				X
Hi-TIDE	**EPSI**: Body Dissatisfaction, Binge Eating, Cognitive Restraint, Excessive Exercise, Restricting, Purging, Muscle Building, and Negative Attitudes Toward Obesity	X	X	X	X	X
	IDAS-II: Depression (General Depression, Dysphoria, Suicidality, Lassitude, Insomnia, Appetite Loss, Appetite Gain), anxiety (Panic, Social Anxiety, Traumatic Intrusions, Claustrophobia, Traumatic Avoidance), positive mood (Well-being), obsessive-compulsive (Checking, Ordering, Cleaning), and bi-polar (Mania, Euphoria) symptoms	X	X	X	X	X
Current eating-disorder symptoms	**EDE-Q**: Restraint, Shape and Weight Concern, and Eating Concern	X				X
	SCOFF: Disordered eating					X
Military and non-military related trauma	**DRRI-2**: Predeployment Factors (Prior Stressors, Childhood Family Functioning), Deployment Factors (Difficult Living and Working Environment, Combat Experiences, Aftermath of Battle, Nuclear, Biological and Chemical Exposures, Perceived Threat, Preparedness, Deployment Support from Family and Friends, Unit Social Support, General Harassment, Sexual Harassment, Concerns about Life and Family, Disruptions, Family Stressors) Postdeployment factors (Postdeployment Stressors, Postdeployment Social Support, Postdeployment Family Functioning)	X				X
	Modified Military Sexual Trauma Screen: Unwanted sexual attention, forceful/threatened sexual contact	X				X
	PCL-5: Re-experiencing, avoidance, negative alterations in cognition and mood, hyper-arousal					X
Externalizing symptoms	**ESI-brief**: Callous aggression, substance use and general disinhibition	X		X	X	
	AUDIT: Alcohol use					X
Internalizing symptoms	**GAD-7**: Generalized anxiety					X
	PHQ-9: Major depressive disorder symptom severity and diagnosis					X
Psychiatric impairment	**WHODAS**: Understanding and communicating, Self-care, Getting along with people, Life activities, Participation in society	X		X	X	X
	CIA: Social, cognitive, personal	X		X	X	X
Healthcare utilization	**NSV**: Health Status, Health Care, Health Insurance, Employment, Income	X		X	X	X
Weight suppression, BMI, and weight history	**Weight control during service**: Skip meals, limit food or lose weight, supplement use, vomiting, binge, excessive exercise, uncomfortably full, feeling stuffed	X		X	X	X
	Self-report height & weight					
	Weight history: Highest weight as adult, lowest weight as adult, lost 10 pounds or more, gained 10 pounds or more					
Past disordered-eating behaviors to meet military requirements	**Modified EPSI**: Compensatory behaviors, restricting behaviors	X				X
Institutional and organizational barriers to accessing eating-disorder and trauma-related care	**Qualitative interview**: Institutional barriers to seeking and receiving eating-disorder and trauma-related care					X
DSM-5 disorders	**SCID-RV for *DSM-5* disorders**: Mood disorders, generalized anxiety disorder, eating disorders, and trauma-and stressor-related disorders					X
	EPSI-CRV: Body Dissatisfaction, Objective Binge Episodes, Binge Eating Disorder, Subjective Binge Eating, Restricting, Cognitive Restraint, Muscle Building, Purging, Excessive Exercise, Low Weight, Overvaluation Of Weight And Shape, Fear Of Weight Gain, Negative Attitudes Towards Obesity					X

Note: AUDIT=Alcohol Use Identification Test; CIA=Clinical Impairment Assessment; DRRI-2=Deployment Risk and Resilience Inventory-2; EDE-Q=Eating Disorder Examination Questionnaire; EPSI=Eating Pathology Symptoms Inventory; EPSI-CRV=Eating Pathology Symptoms Inventory – Clinician Rated Version; ESI-brief=Externalizing Symptoms Inventory-Brief; GAD-7=Generalized Anxiety Disorder Questionnaire 7; IDAS-II =Inventory for Depression and Anxiety Symptoms-II; NSV=National Survey of Veterans; SCID-RV=Structured Clinical Interview for DSM Disorders-Research Version; SCOFF (this is the full name of the measure and not an abbreviation); PCL-5=Post-Traumatic Stress Disorder Checklist; PHQ-9: Patient Health Questionnaire Depression Scale; and WHODAS=World Health Organization Disability Assessment Schedule.

1.5.1 | Alcohol Use Identification Test (AUDIT)

The AUDIT is a 10-item self-report screen for problematic alcohol use (Babor et al., 1992). The AUDIT has shown high internal consistency across a variety of samples and has a median sensitivity of 0.86 and median specificity of 0.89 (Reinert & Allen, 2002). The AUDIT was administered in Study 2.

1.5.2 | Clinical Impairment Assessment (CIA)

The CIA is a 16-item measure evaluating the severity of psychosocial impairment secondary to an ED (Bohn et al., 2008). This scale has

shown evidence for high internal consistency ($\alpha = 0.97$) and test-retest reliability ($r = 0.86$; Bohn et al., 2008).

1.5.3 | Demographics, body mass index, and weight history

We adapted the National Survey of Veterans (Department of Veterans Affairs, 2010) to collect basic and military-specific demographic information from veterans. Additional ad hoc demographic measures included: lifetime history of ED behaviors, body mass index ([BMI] kg/m^2), and participants' weight history. We also administered questions adapted from Garber et al. (2008) and the Eating Pathology

TABLE 4 Internal consistency reliability for selected study measures from Study 1

Measures	Study 1			
	Baseline	One-week	Three-month	Six-month
EPSI				
Body Dissatisfaction	0.9	0.9	0.91	0.9
Binge Eating	0.91	0.91	0.91	0.93
Cognitive Restraint	0.71	0.76	0.71	0.68
Excessive Exercise	0.87	0.87	0.87	0.84
Restricting	0.87	0.89	0.87	0.86
Purging	0.86	0.85	0.86	0.87
Muscle Building	0.82	0.79	0.79	0.79
Negative Attitudes Towards Obesity	0.9	0.92	0.93	0.93
IDAS-II				
General Depression	0.94	0.94	0.94	0.94
Dysphoria	0.93	0.93	0.93	0.93
Suicidality	0.87	0.87	0.84	0.89
Lassitude	0.86	0.85	0.86	0.87
Insomnia	0.87	0.87	0.9	0.89
Appetite Loss	0.85	0.86	0.84	0.84
Appetite Gain	0.83	0.79	0.82	0.82
Panic	0.89	0.9	0.9	0.89
Social Anxiety	0.89	0.89	0.9	0.91
Traumatic Intrusions	0.89	0.9	0.91	0.92
Claustrophobia	0.91	0.92	0.93	0.93
Traumatic Avoidance	0.91	0.93	0.93	0.92
Well-being	0.92	0.91	0.91	0.93
Checking	0.85	0.88	0.89	0.89
Ordering	0.85	0.85	0.87	0.85
Cleaning	0.89	0.92	0.91	0.92
Mania	0.87	0.87	0.89	0.9
Euphoria	0.73	0.73	0.76	0.77
WHODAS	0.97		0.97	0.97
CIA	0.97		0.97	0.97
EDE-Q	0.79			
DRRI-2				
Events	0.83			
Combat	0.93			
Warfare	0.94			
Others	0.89			

Note: These data are from the first 816 participants in Study 1. EDE-Q=Eating Disorder Examination Questionnaire; SCOFF (this is the full name of the measure and not an abbreviation); IDAS-II =Inventory for Depression and Anxiety Symptoms-II; DRRI-2=Deployment Risk and Resilience Inventory-2; ESI-brief=Externalizing Symptoms Inventory-Brief; WHODAS=World Health Organization Disability Assessment Schedule; CIA=Clinical Impairment Assessment; EPSI=Eating Pathology Symptoms Inventory.

Symptoms Inventory (Forbush et al., 2013) to assess past ED behaviors to meet military requirements.

1.5.4 | Deployment Risk and Resilience Inventory-2 (DRRI-2)

We used three scales from the original DRRI-2 inventory (Vogt et al., 2013): Pre-deployment Life Events, Combat Experiences, Relationships During Deployment. DRRI-2 scales have shown evidence for good internal consistency ($\alpha > 0.80$), criterion validity, and test-retest validity ($\alpha = 0.61$–0.94; Maoz et al., 2016; Vogt et al., 2013). The DRRI-2 scales showed high discriminant validity across scales (Vogt et al., 2013).

1.5.5 | Eating Disorder Examination Questionnaire (EDE-Q)

The EDE-Q (Fairburn & Beglin, 1994) is a 36-item questionnaire that assesses ED psychopathology. The EDE-Q has four subscales: Eating Concern, Weight Concern, Shape Concern, and Restraint. Subscales

can be averaged to form a Global Score. The EDE-Q has demonstrated evidence for high internal consistency, test-retest reliability, and criterion-related validity (Berg et al., 2012), although concerns have been raised about factor structure replicability and discriminant validity (Forbush et al., 2013; Thomas et al., 2014).

1.5.6 | Eating Pathology Symptoms Inventory (EPSI)

The EPSI (Forbush et al., 2013) is a 45-item questionnaire designed to measure ED behaviors and cognitions. Initial development and validation of EPSI scales demonstrated evidence for strong internal consistency ($\alpha = 0.84$–0.89) and test-retest reliability ($r = 0.73$). The EPSI has shown excellent convergent and discriminant validity (Forbush et al., 2013, 2014, 2020).

1.5.7 | EPSI-Clinician-Rated version (EPSI-CRV)

The EPSI-CRV (Forbush et al., 2020) is a semi-structured clinical interview that generates both dimensional scores and current DSM-5 ED diagnoses. The EPSI-CRV has shown high inter-rater reliability, discriminant and convergent validity, and criterion-related validity (Forbush et al., 2020). The EPSI-CRV was administered in Study 2.

1.5.8 | Externalizing Spectrum Inventory—Brief (ESI-Brief)

The ESI-Brief is a 145-item assessment tool that evaluates three higher-order dimensions of externalizing psychopathology, including: Callous Aggression, Substance Use, and General Disinhibition, which demonstrated high internal consistency (Patrick et al., 2013). The ESI-Brief was administered in Study 1.

1.5.9 | Generalized anxiety disorder-Q7

The GAD-7 (Spitzer et al., 2006) is a seven-item measure of generalized anxiety with strong support for its construct validity and reliability. Recent research found the GAD-7 had poor sensitivity and a high false-positive rate for identifying generalized anxiety disorder and other anxiety disorders (Beard & Björgvinsson, 2014). However, the GAD-7 performed well as a measure of anxiety symptom severity in a large psychiatric sample (Beard & Björgvinsson, 2014). The GAD-7 was administered in Study 2.

1.5.10 | Inventory of Depression and Anxiety Symptoms-II (IDAS-II)

The IDAS-II is a 99-item self-report questionnaire that contains 11 non-overlapping scales assessing depression, anxiety, mania, obsessive-compulsive, and trauma-related psychopathology. The IDAS-II has shown good-to-excellent internal consistency with median α's ranging from 0.84 to 0.87 across a variety of samples (Watson et al., 2012). IDAS-II scales have shown strong discriminant, convergent, and criterion validity (Watson et al., 2012).

1.5.11 | The Post-Traumatic Stress Disorder Checklist (PCL-5)

The PCL-5 (Blevins et al., 2015) is a 20-item screening measure. Studies show the PCL-5 has good stability, internal consistency, test-retest reliability, and convergent and discriminant validity (Bovin et al., 2016; Wortmann et al., 2016). The PCL-5 was administered in Study 2.

1.5.12 | Patient Health Questionnaire Depression Scale (PHQ-9)

The PHQ-9 (Kroenke et al., 2001) is a nine-item self-report measure of depression that is well-validated and reliable. The PHQ-9 has sensitivity and specificity of 0.88 for identifying major depression in primary care and obstetric-gynecology clinics (Kroenke et al., 2001). The PHQ-9 was administered in Study 2.

1.5.13 | SCOFF

The SCOFF (Morgan et al., 1999) is a five-item self-report screening measure for EDs. The SCOFF was administered in Study 2. Each letter in SCOFF corresponds to one of the items: making oneself **s**ick; losing **c**ontrol over eating; losing at least **o**ne stone in weight; feeling **f**at; and believing **f**ood has control over one's life. The SCOFF is a widely utilized instrument, with evidence for good specificity and sensitivity (see Kutz et al., 2020, for a review).

1.5.14 | The Structured Clinical Interview for *DSM-5* Disorders (SCID-5)

The SCID-5 (First et al., 2015) is a semi-structured clinical interview that generates current and lifetime psychiatric diagnoses. Studies have demonstrated the specificity/sensitivity, interrater reliability, and predictive validity of the interview (Osório et al., 2019). We administered modules to assess the following lifetime and current psychiatric diagnoses: mood disorders, generalized anxiety disorder, trauma- and stress-related disorders, and eating disorders. The SCID-5 was administered in Study 2.

1.5.15 | Qualitative interview

A qualitative interview will be administered to participants in Study 2. Participants will answer non-numerical questions concerning

experiences in the military. Specifically, they will be asked about institutional and organizational barriers related to seeking or receiving treatment for eating and trauma-related psychopathology.

1.5.16 | World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)

The WHODAS 2.0 (Üstün et al., 2010) is a 36-item measure of psychiatric impairment. Previous studies have found evidence for strong internal consistency ($\alpha = 0.86$) and test-retest reliability ($r = 0.98$) as well as concurrent and predictive validity (Gold, 2014; Üstün et al., 2010).

2 | RESULTS

2.1 | Sample representativeness

To assess sample representativeness, we computed correlations by creating a single variable to represent all unique strata within the full population of veteran men and women. Next, we correlated the frequencies of strata observed in the population with the strata we observed in our sample. In both Studies 1 and 2, the correlation of demographic characteristics between the stratified sample and the full population of veterans was 0.95 or higher for each sex; thus, our stratified samples were representative of the full populations from which they were drawn. Further details are provided in Table 4. As a supplementary analysis, we also tested representativeness of our sample across each of the seven categorical demographic variables by comparing the extracted samples to the full VADIR populations for each demographic variable using chi-square tests (see Supporting Information S1).

2.2 | Response bias and attrition analyses

For the Study 1 (longitudinal) sample, we conducted two analyses to investigate the extent to which demographic characteristics that were used to draw stratified samples from VADIR were associated with (1) non-response to our initial recruitment (response bias analysis), (2) retention from our initial recruitment at 1-week follow-up (attrition analysis), (3) missing data in the baseline survey for Study 1, and (4) invalid address. The objectives of these analyses were to assess whether certain demographic groups were significantly less likely to (1) participate (respond), (2) continue participation after completing the baseline survey for Study 1 (attrition), (3) have more than 10% of missing data, and (4) have invalid address. For all analyses, we used a logistic regression model with all seven demographic variables used for stratification, along with sex, as predictors. The first analysis was performed on the full recruitment group for Study 1 ($N = 14,500$), wherein we used the logistic regression model to predict participation in the baseline survey. For

the second analysis, we examined the baseline sample to identify whether any demographic characteristics were associated with participation in the 1-week follow-up survey. For the third analysis, we examined the baseline sample to identify whether any demographic characteristics were associated with missing values. For the last analysis, we examined whether demographic characteristics were associated with invalid addresses. Analyses included the following procedures: First, we fitted the saturated model (which included all predictors) and a reduced model, which included all terms except one predictor. Next, we conducted a likelihood-ratio test (LRT) comparing the saturated model to each reduced model to determine the significance of each variable in predicting participation at (1) baseline and (2) 1-week follow-up. We conducted post-hoc tests for variables with significant LRT results to assess specific levels of each variable that exhibited significant pairwise differences.

Results of the first analysis suggested that several demographic variables were significant predictors of response to the baseline survey, including gender, rank, race, ethnicity, age of discharge (see Figure 1). Men were significantly less likely to respond to our mailings than women (*Odds Ratio [OR]* = 0.54, 95% CI = [0.46, 0.62]). *Rank* (operationalized by pay grade) exhibited a statistically significant omnibus effect (Chi-squared = 26.10, *df* = 3, $p < 0.001$), such that enlisted individuals were less likely to participate than non-commissioned officers ($p < 0.001$), commissioned officers ($p < 0.001$), and warrant officers ($p = 0.02$). *Race* exhibited a statistically significant omnibus effect (Chi-squared = 36.26, *df* = 5, $p < 0.001$), such that people who identified as Black or African American were significantly less likely to participate than those who identified as White ($p < 0.001$), Asian ($p = 0.03$), and multiracial or unknown ($p = 0.004$). For *ethnicity*, people not of Hispanic origin were significantly more likely to participate in the baseline survey ($OR = 1.30$, 95% CI = [1.04, 1.63]). Finally, we found a statistically significant omnibus effect for *age of discharge* (Chi-squared = 22.92, *df* = 3, $p < 0.001$), such that the fourth quartile was significantly more likely to participate than the first ($p < 0.001$) and second ($p = 0.01$) quartiles. Moreover, the second and third quartiles were significantly more likely to participate than the first quartile ($p = 0.003$, $p < 0.001$, respectively).

For the second analysis, we found that only *age at discharge* was statistically significantly related to retention in the 1-week follow up (Chi-squared = 12.81, *df* = 3, $p = 0.005$) (see Figure 2). Individuals whose age at discharge was in the fourth quartile were significantly more likely to participate than the first ($p = 0.001$), second ($p = 0.02$) and third ($p = 0.01$) quartiles.

For the third analysis, we found that demographic characteristics were not related to missing data. Thus, there was no evidence of demographic differences between participants who had >10% of missing data and those who did not.

For the fourth analysis, several variables were significant predictors of invalid addresses, including branch, personnel category, age-at-discharge, and rank (see Figure 3). Branch exhibited a statistically significant omnibus effect (Chi-squared = 77.67, *df* = 3, $p < 0.001$), such that Navy veterans were more likely to have invalid
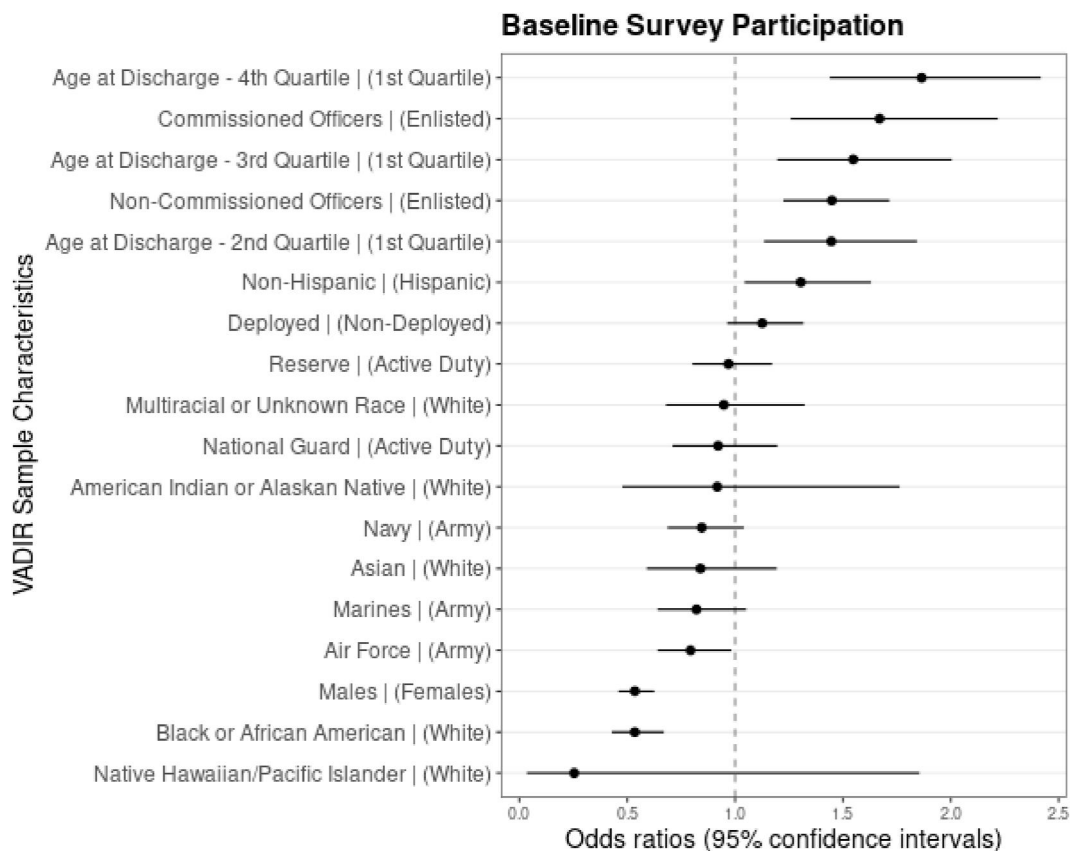
Baseline Survey Participation

FIGURE 1 Response bias analysis. Plot of odds ratios with 95% confidence intervals for logistic regression model predicting participation in the baseline survey from the group of potential recruits. For each predictor, the associated reference group is listed in parentheses. For instance, we see that the third coefficient from the bottom reflects the odds ratio for males (vs. females), and indicates that males were significantly less likely to participate in the study than females. One coefficient, for warrant officers, was omitted from the plot due to an extremely large confidence interval (resulting from low representation for that group; $N = 12$). For warrant officers (in comparison with those who were enlisted), the model returned: $OR = 4.97$, 95% $CI = [1.32, 18.80]$. The total sample size for the baseline survey was $N = 816$

addresses than Army ($p < 0.001$), Air Force ($p = 0.017$), and Marine veterans ($p < 0.001$). Personnel category exhibited a statistically significant omnibus effect (Chi-squared $= 13.86$, $df = 2$, $p < 0.001$), with Active Duty individuals being more likely to have invalid addresses than the National Guard ($p < 0.001$) and Reservists ($p = 0.047$). Age at discharge had a statistically significant effect (Chi-squared $= 41.09$, $df = 3$, $p < 0.001$), in which the 1st Quartile was more likely to have invalid addresses than the 2nd Quartile ($p < 0.001$); the 2nd Quartile was not significantly more likely to have invalid addresses than the 3rd Quartile ($p = 0.086$); and the 3rd Quartile was more likely to have invalid addresses than the 4th Quartile ($p = 0.016$). Rank had a significant effect (Chi-squared $= 9.53$, $df = 3$, $p = 0.023$)—Enlisted individuals were more likely to have invalid addresses than Non-Commissioned Officers ($p = 0.028$).

2.2.1 | Post-stratification weights

To control for non-response bias, we created post-stratification weights, which can be used in future analyses to adjust for non-

response bias. We applied a survey raking procedure (also known as "sample balancing") using the R package "anesrake" (Pasek, 2018). Raking is a method used to adjust survey data to ensure the distributions of sample characteristics closely mirror the distributions of population characteristics. In our case, the full population of interest was the group of veterans sampled from the VADIR data who were considered as candidates for participation (e.g., $N = 14,500$ in Study 1). Although Studies 1 and 2 drew samples from VADIR that represented the total population of recently separated veterans, we observed certain sources of non-response bias in which some demographic groups were more or less likely to respond to the surveys than other groups. Therefore, we needed to weight the data collected in both studies to match the characteristics of the recruitment groups to ensure that the final collected samples were representative of the population. Given that data collection for Study 2 is ongoing, we only applied this procedure to Study 1.

The raking algorithm implemented in the "anesrake" package (Pasek & Pasek, 2018) used the seven variables used for stratification (and sex) in the total population and evaluated which distributions deviated by 5% or more in the recruited samples. The algorithm iteratively adjusted the frequencies of stratified variables to produce
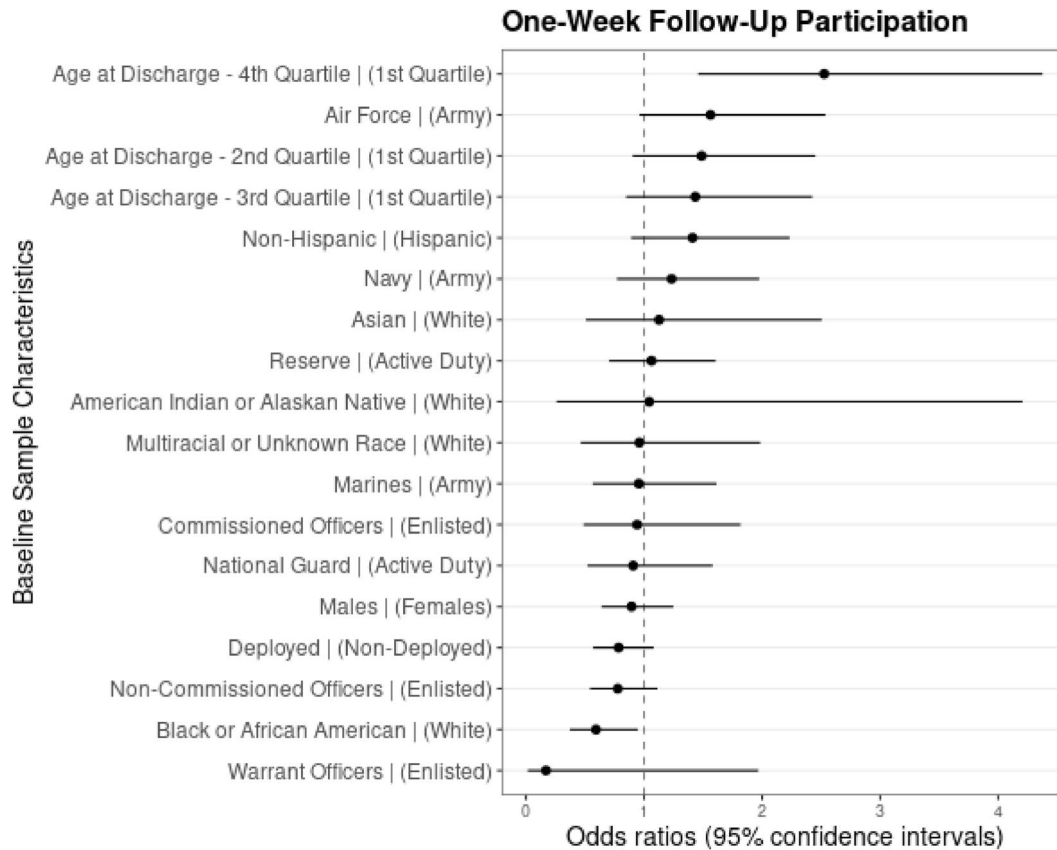
One-Week Follow-Up Participation

FIGURE 2 One-week follow-up attrition analysis. Plot of odds ratios with 95% confidence intervals for baseline demographic characteristics predicting participation in the 1-week follow-up survey. For each predictor, the associated reference group is listed in parentheses. For instance, the second coefficient from the bottom reflects the odds ratio for Black or African American (vs. White) individuals, and indicates that Black or African American participants were significantly less likely to participate in the 1-week follow-up than White participants. One case was removed from the analysis, as there was only one individual who identified their race as Native Hawaiian/Pacific Islander in the sample that completed the baseline survey. The total sample size at the 1-week follow-up was $N = 560$

a weight for each case in the sample until the distributions were aligned with the population distributions for each demographic characteristic. Weighted frequencies of demographic variables were then calculated for each level of each variable to assess the effectiveness of the raking procedure. These values are provided for Study 1 in Table 1. This procedure was highly successful, as indicated by an average error rate of less than 0.01%, with the maximum error for a given variable being 0.06%.

3 | DISCUSSION

The LEAP Consortium is focused on developing novel methods for screening, classifying, and treating veterans and military members with EDs and other internalizing disorders. The purpose of this paper was to provide information on the design, protocol, methods, response bias, and initial attrition analyses for the first two datasets from the LEAP Consortium. Study 1 is nearly complete and Study 2 began in March 2022. Study 1 is a four-wave longitudinal survey design and Study 2 is a cross-sectional study that will include both surveys, structured clinical interviews, and qualitative interviews

designed to identify barriers to seeking ED and trauma-related treatment while in the military. Results showed our recruitment samples were highly representative of the full population of recently separated veterans.

Despite the high correspondence between our recruitment sample of veterans invited to participate and the total population of recently separated veterans, not all invited veterans responded to our surveys. We identified several sources of response bias in our initial survey responses. For example, individuals who identified as men and enlisted personnel were less likely to participate in Study 1 compared to women and non-commissioned, commissioned, and warrant officers. Individuals in minoritized populations were also less likely to participate in Study 1 compared to veterans who identified as White. Finally, veterans who were 31–68 years of age at discharge were more likely to participate than veterans who were discharged at 17–25 or 25–27. Attrition analyses also revealed that veterans 31–68 years of age were the least likely to drop out of the study. No other demographic characteristics predicted attrition, suggesting that, overall, attrition largely was random rather than related to specific demographic features. To account for response bias, we created sample weights using a validated raking procedure to ensure
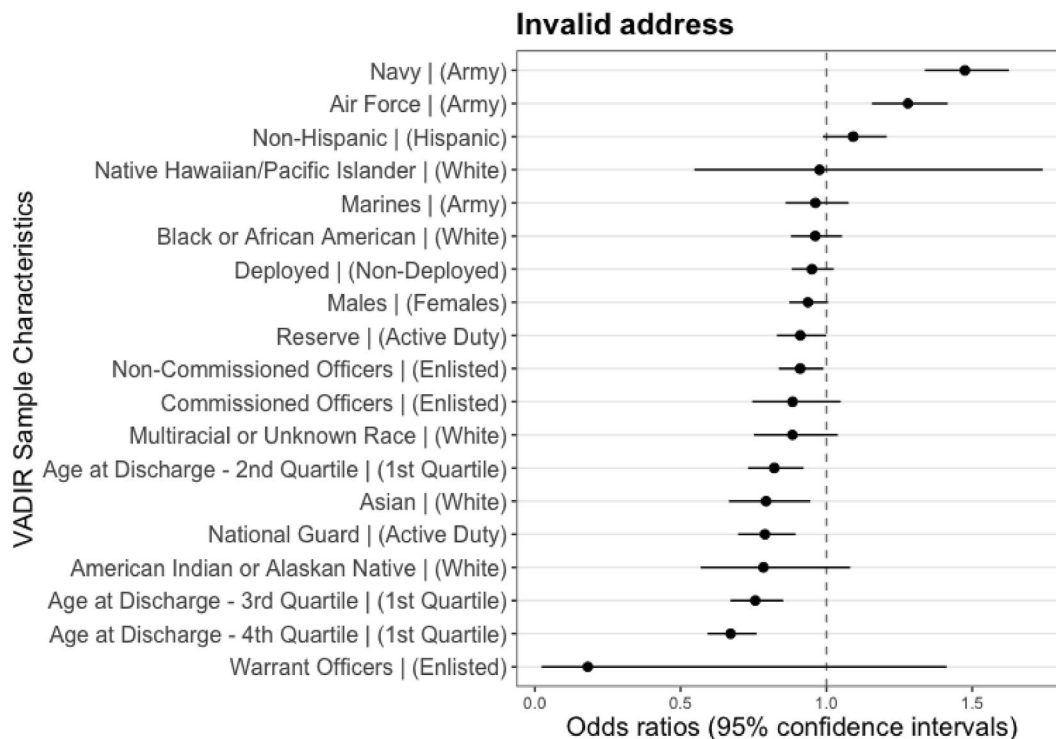
Invalid address

FIGURE 3 Response Bias Analysis due to Invalid Addresses. This figure provides a plot of odds ratios with 95% confidence intervals for VADIR sample characteristics predicting invalid address. For each predictor, the associated reference group is listed in parentheses like Figures 1 and 2

that researchers using LEAP data in the future will be able to generalize data back to the population from which they were drawn to ensure representative sampling.

Certain limitations of the study should be noted. First, we experienced a higher proportion of returned mail than initially expected and encountered slow response time for mailed surveys due to U.S. postal service delays. These issues resulted in fewer baseline participants in Study 1 than we had initially expected. To address this concern, we recruited additional participants from VADIR and adjusted our protocol to send an initial postcard to check for inaccurate addresses prior to sending our first recruitment invitation letter and enrollment materials. These procedures were successful in ensuring we were able to recruit sufficient numbers of veterans to test our longitudinal research aims. Indeed, we observed 33% invalid addresses—versus 48%—after implementing initial postcards to check for invalid addresses. However, our procedural change resulted in some veterans participating later than the initial sample, often past the 1-year mark from when they separated from service. For example, the initial sample was sent their first recruitment letter on 9/15/2020 and the additional sample was sent their first recruitment letter on 4/23/2021. We also found demographic differences between individuals who had an invalid address and individuals who did not; however, the potential impact of this bias was minimized by post-stratification weights that correct for non-response bias. Second, although we sampled men and women veterans in a 1:1 ratio, women were more likely to enroll in Study 1. We attempted to

address this problem by creating post-stratification survey weights to adjust our Study 1 response data to ensure our final sample matched the full population of veterans recruited from VADIR on demographic characteristics. Third, due to pandemic-related study issues that prevented staff from in-person work and unexpected mailing delays, we were delayed in initiating Study 2. To manage our time constraints with budgetary limitations that prevented us from hiring additional staff to increase the rate of weekly interviews, we chose to focus diagnostic interviews on current EDs, mood disorders, generalized anxiety disorder, and post-traumatic stress disorder because these conditions were the focus of our newly developed self-reported screening tool. Due to our study design, it was not possible to conduct analyses on the prevalence or correlates of *DSM-5* defined externalizing disorders and certain internalizing disorders. However, researchers who are interested in testing hypotheses related to externalizing psychopathology or additional forms of internalizing psychopathology could use Study 1 or Study 2 self-reported data from the ESI-Brief (Patrick et al., 2013) or the IDAS-II, which comprehensively assess a range of externalizing and internalizing problems, respectively. Fourth, related to our previous concerns, we were unable to assess lifetime diagnoses for most mental-health disorders, given time and budget constraints. Fifth, Study 2 was ongoing and we were unable to identify (or correct for) non-response bias in this sample. Future research will focus on creating post-stratification weights for Study 2, when data collection is complete. Finally, it is possible that there were retrospective recall

TABLE 5 Global correlations between VADIR populations and samples

	Study 1		Study 2	
	Strata	r	Strata	r
Females	1305	0.98	965	0.95
Males	1756	0.99	1441	0.97

Note: This table shows the correlations between the frequencies associated with the strata observed in the VADIR databases and those observed in the stratified samples. In the columns labeled Strata, these values show the number of unique strata associated with each sex in the full populations. The strata are directly observed in the populations, and are based on crossing all levels of each of the seven categorical variables used in the stratification procedure. The correlations are then computed by taking the vector of frequencies associated with all strata in the populations and comparing them against the vector of frequencies associated with the same strata in the samples to obtain a global measure of similarity. Any strata that appeared in the population but not in the samples were still included in computing the correlations, but simply had a frequency of 0 within the samples.

biases that impacted the reporting of questions about barriers to accessing ED care while serving in the military. We attempted to mitigate this issue by recruiting veterans separated from service within the past year, to minimize recall biases. However, future studies are needed to assess access to ED services among active duty military service members (Table 5).

In conclusion, the current study offers a rich source of information about ED and internalizing symptoms and diagnoses, military-relevant variables, and self-reported externalizing psychopathology. Study 1 will represent one of the first longitudinal nationally representative samples of US veterans to include EDs and comorbid disorders, which will provide much-needed information about the temporal relationships among EDs, mood, anxiety, and trauma-related symptoms. Study 2 will provide prevalence estimates for current and lifetime EDs in veterans and identify barriers to seeking ED and trauma-informed care. We expect that the LEAP Consortium datasets will, therefore, contribute to improved information about EDs in veterans, an important and understudied problem.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

We have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

De-identified and unclassified data collected from this study will be made publicly available through the Center for Open Science (OSF).

Our procedures for sharing data will be similar to those used in other large-scale mental-health studies, such as the Minnesota Twin Family Study (Iacono et al., 2006). Investigators interested in using the data should contact the Principal Investigator (KTF) to provide a brief description of the intended use and proposed project timeline. These project descriptions will eventually be posted on OSF. Investigators will also be required to sign a data-use agreement. Such documentation will prevent duplication of research and provide a greater level of protection of intellectual property.

ORCID

Kelsie T. Forbush https://orcid.org/0000-0002-5900-4204
Robert W. Morgan https://orcid.org/0000-0002-0188-2709

REFERENCES

Allen, K. L., Byrne, S. M., Oddy, W. H., & Crosby, R. D. (2013). DSM–IV–TR and DSM-5 eating disorders in adolescents: Prevalence, stability, and psychosocial correlates in a population-based sample of male and female adolescents. *Journal of Abnormal Psychology*, 122(3), 720–732. https://doi.org/10.1037/a0034004

Arditte Hall, K. A., Bartlett, B. A., Iverson, K. M., & Mitchell, K. S. (2017). Military-related trauma is associated with eating disorder symptoms in male veterans. *International Journal of Eating Disorders*, 50(11), 1328–1331. https://doi.org/10.1002/eat.22782

Aronson, K. R., Perkins, D. F., Morgan, N. R., Bleser, J. A., Vogt, D., Copeland, L., Finley, E., & Gilman, C. (2019). Post-9/11 veteran transitions to civilian life: Predictors of the use of employment programs. *Journal of Veterans Studies*, 5(1), 14. https://doi.org/10.21061/jvs.v5i1.127

Babor, T. F., de la Fuente, J. R., Saunders, J., & Grant, J. (1992). *The alcohol use disorders identification test: Guidelines for use in primary health care* (WHO Publication No. 92.4). World Health Organization.

Barth, S. K., Dursa, E. K., Peterson, M. R., & Schneiderman, A. (2014). *Prevalence of respiratory diseases among veterans of operation enduring freedom and operation Iraqi freedom: Results from the national health study for a new generation of US veterans.* Oxford University Press.

Beard, C., & Björgvinsson, T. (2014). Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *Journal of Anxiety Disorders*, 28(6), 547–552. https://doi.org/10.1016/j.janxdis.2014.06.002

Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders*, 45(3), 428–438. https://doi.org/10.1002/eat.20931

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

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy*, 46(10), 1105–1110. https://doi.org/10.1016/j.brat.2008.06.012

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

Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: Developing taxonomy, themes, and

theory. *Health Services Research*, *42*(4), 1758–1772. https://doi.org/10.1111/j.1475-6773.2006.00684.x

Brown, C. A., & Mehler, P. S. (2013). Medical complications of self-induced vomiting. *Eating Disorders*, *21*(4), 287–294. https://doi.org/10.1080/10640266.2013.797317

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.

Chesney, E., Goodwin, G. M., & Fazel, S. (2014). Risks of all-cause and suicide mortality in mental disorders: A meta-review. *World Psychiatry*, *13*(2), 153–160. https://doi.org/10.1002/wps.20128

Coughlin, S., Aliaga, P., Barth, S., Eber, S., Maillard, J., Mahan, C., Kang, H., Schneiderman, A., DeBakey, S., Vanderwolf, P., & Williams, M. (2011). The effectiveness of a monetary incentive on response rates in a survey of recent U.S. veterans. *Survey Practice*, *4*, 1–8. https://doi.org/10.29115/sp-2011-0004

Cuthbert, K., Hardin, S., Zelkowitz, R., & Mitchell, K. (2020). Eating disorders and overweight/obesity in veterans: Prevalence, risk factors, and treatment considerations. *Current Obesity Reports*, *9*(2), 98–108. https://doi.org/10.1007/s13679-020-00374-1

Department of Veterans Affairs. (2010). *National survey of veterans, active duty service members, demobilized national guard and reserve members, family members, and surviving spouses*. Department of Veterans Affairs.

Dillman, D. A. (2011). *Mail and internet surveys: The tailored design method – 2007 update with new internet, visual, and mixed-mode guide*. John Wiley & Sons.

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., & Messer, B. L. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the internet. *Social Science Research*, *38*(1), 1–18. https://doi.org/10.1016/j.ssresearch.2008.03.007

Eber, S., Barth, S., Kang, H., Mahan, C., Dursa, E., & Schneiderman, A. (2013). The national health study for a new generation of United States veterans: Methods for a large-scale study on the health of recent veterans. *Military Medicine*, *178*(9), 966–969. https://doi.org/10.7205/milmed-d-13-00175

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363–370. https://doi.org/10.1002/1098-108x(199412)16:4<363::aid-eat2260160405>3.0.co;2-#

Fichter, M. M., & Quadflieg, N. (2016). Mortality in eating disorders-results of a large prospective clinical longitudinal study. *International Journal of Eating Disorders*, *49*(4), 391–401. https://doi.org/10.1002/eat.22501

First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2015). *Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV)*. American Psychiatric Association.

Forbush, K. T., Wildes, J. E., Bohrer, B. K., Hagan, K. E., Clark, K. E., Chapa, D. A. N., Perko, V., & Christensen, K. A. (2020). Development and initial validation of the eating Pathology symptoms inventory—Clinician rated version (EPSI-CRV). *Psychological Assessment*, *32*(10), 943–955. https://doi.org/10.1037/pas0000820

Forbush, K. T., Wildes, J. E., & Hunt, T. K. (2014). Gender norms, psychometric properties, and validity for the eating pathology symptoms inventory. *International Journal of Eating Disorders*, *47*(1), 85–91. https://doi.org/10.1002/eat.22180

Forbush, K. T., Wildes, J. E., Pollack, L. O., Dunbar, D., Luo, J., Patterson, K., Pollpeter, M., Stone, A., Bright, A., & Watson, D. (2013). Development and validation of the eating pathology symptoms inventory (EPSI). *Psychological Assessment*, *25*(3), 859–878. https://doi.org/10.1037/a0032639

Forman-Hoffman, V. L., Mengeling, M., Booth, B. M., Torner, J., & Sadler, A. G. (2012). Eating disorders, post-traumatic stress, and sexual trauma in women veterans. *Military Medicine*, *177*(10), 1161–1168. https://doi.org/10.7205/milmed-d-12-00041

Forney, K. J., Buchman-Schmitt, J. M., Keel, P. K., & Frank, G. K. (2016). The medical complications associated with purging. *International Journal of Eating Disorders*, *49*(3), 249–259. https://doi.org/10.1002/eat.22504

Garber, A. K., Boyer, C. B., Pollack, L. M., Chang, Y. J., & Shafer, M.-A. (2008). Body mass index and disordered eating behaviors are associated with weight dissatisfaction in adolescent and young adult female military recruits. *Military Medicine*, *173*(2), 138–145. https://doi.org/10.7205/milmed.173.2.138

Gold, L. H. (2014). DSM-5 and the assessment of functioning: The World Health Organization disability assessment schedule 2.0 (WHODAS 2.0). *Journal of the American Academy of Psychiatry & Law Online*, *42*, 173–181.

Hoddinott, S. N., & Bass, M. J. (1986). The Dillman total design survey method. *Canadian Family Physician*, *32*, 2366–2368.

Iacono, W. G., McGue, M., & Krueger, R. F. (2006). Minnesota center for twin and family research. *Twin Research & Human Genetics*, *9*(6), 978–984. https://doi.org/10.1375/twin.9.6.978

Iwajomo, T., Bondy, S. J., de Oliveira, C., Colton, P., Trottier, K., & Kurdyak, P. (2021). Excess mortality associated with eating disorders: Population-based cohort study. *British Journal of Psychiatry*, *219*(3), 487–493. https://doi.org/10.1192/bjp.2020.197

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

Kutz, A. M., Marsh, A. G., Gunderson, C. G., Maguen, S., & Masheb, R. M. (2020). Eating disorder screening: A systematic review and meta-analysis of diagnostic test characteristics of the SCOFF. *Journal of General Internal Medicine*, *35*(3), 885–893. https://doi.org/10.1007/s11606-019-05478-6

Maoz, H., Goldwin, Y., Lewis, Y. D., & Bloch, Y. (2016). Exploring reliability and validity of the deployment risk and resilience inventory-2 among a nonclinical sample of discharged soldiers following mandatory military service. *Journal of Traumatic Stress*, *29*(6), 556–562. https://doi.org/10.1002/jts.22135

Masheb, R. M., Ramsey, C. M., Marsh, A. G., Decker, S. E., Maguen, S., Brandt, C. A., & Haskell, S. G. (2021). DSM-5 eating disorder prevalence, gender differences, and mental health associations in United States military veterans. *International Journal of Eating Disorders*, *54*(7), 1171–1180. https://doi.org/10.1002/eat.23501

Mitchell, K. S., Scioli, E. R., Galovski, T., Belfer, P. L., & Cooper, Z. (2021). Post-traumatic stress disorder and eating disorders: Maintaining mechanisms and treatment targets. *Eating Disorders*, 1–15(3), 292–306. https://doi.org/10.1080/10640266.2020.1869369

Mitchison, D., Hay, P., Slewa-Younan, S., & Mond, J. (2014). The changing demographic profile of eating disorder behaviors in the community. *BMC Public Health*, *14*(1), 943. https://doi.org/10.1186/1471-2458-14-943

Morgan, J. F., Reid, F., & Lacey, J. H. (1999). The SCOFF questionnaire: Assessment of a new screening tool for eating disorders. *BMJ*, *319*(7223), 1467–1468. https://doi.org/10.1136/bmj.319.7223.1467

Osório, F. L., Loureiro, S. R., Hallak, J. E. C., Machado-de-Sousa, J. P., Ushirohira, J. M., Baes, C. V., Apolinario, T. D., Donadon, M. F., Bolsoni, L. M., Guimarães, T., Fracon, V. S., Silva-Rodrigues, A. P. C., Pizeta, F. A., Souza, R. M., Sanches, R. F., dos Santos, R. G., Martin-Santos, R., & Crippa, J. A. S. (2019). Clinical validity and intrarater and test–retest reliability of the structured clinical interview for DSM-5–Clinician version (SCID-5-CV). *Psychiatry & Clinical Neurosciences*, *73*(12), 754–760. https://doi.org/10.1111/pcn.12931

Pasek, J. (2018). anesrake: ANES raking implementation, R package version 0.80.

Pasek, J., & Pasek, M. (2018). ANES raking implementation and weighted statistics.

Patrick, C. J., Kramer, M. D., Krueger, R. F., & Markon, K. E. (2013). Optimizing efficiency of psychopathology assessment through quantitative modeling: Development of a brief form of the externalizing spectrum inventory. *Psychological Assessment*, *25*(4), 1332–1348. https://doi.org/10.1037/a0034864

Reinert, D. F., & Allen, J. P. (2002). The alcohol use disorders identification test (AUDIT): A review of recent research. *Alcoholism: Clinical & Experimental Research*, *26*(2), 272–279. https://doi.org/10.1111/j.1530-0277.2002.tb02534.x

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Stice, E., Marti, C. N., & Rohde, P. (2013). Prevalence, incidence, impairment, and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. *Journal of Abnormal Psychology*, *122*(2), 445–457. https://doi.org/10.1037/a0030679

Streatfeild, J., Hickson, J., Austin, S. B., Hutcheson, R., Kandel, J. S., Lampert, J. G., Myers, E. M., Richmond, T. K., Samnaliev, M., Velasquez, K., Weissman, R. S., & Pezzullo, L. (2021). Social and economic cost of eating disorders in the United States: Evidence to inform policy action. *International Journal of Eating Disorders*, *54*(5), 851–868. https://doi.org/10.1002/eat.23486

Striegel-Moore, R. H., Garvin, V., Dohm, F.-A., & Rosenheck, R. A. (1999). Eating disorders in a national sample of hospitalized female and male veterans: Detection rates and psychiatric comorbidity. *International Journal of Eating Disorders*, *25*(4), 405–414. https://doi.org/10.1002/(sici)1098-108x(199905)25:4<405::aid-eat5>3.0.co;2-f

Striegel Weissman, R., & Rosselli, F. (2017). Reducing the burden of suffering from eating disorders: Unmet treatment needs, cost of illness, and the quest for cost-effectiveness. *Behaviour Research & Therapy*, *88*, 49–64. https://doi.org/10.1016/j.brat.2016.09.006

Swanson, T., & Forbush, K. (2021). SampleVADIR: Draw stratified samples from the VADIR database. https://CRAN.R-project.org/package=sampleVADIR

Thomas, J. J., Roberto, C. A., & Berg, K. C. (2014). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *Advances in Eating Disorders: Theory, Research and Practice*, *2*(2), 190–203. https://doi.org/10.1080/21662630.2013.840119

Üstün, T. B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korff, M., & Pull, C. (2010). Developing the World Health Organization disability assessment schedule 2.0. *Bulletin of the World Health Organization*, *88*(11), 815–823. https://doi.org/10.2471/blt.09.067231

van Hoeken, D., & Hoek, H. W. (2020). Review of the burden of eating disorders: Mortality, disability, costs, quality of life, and family burden. *Current Opinion in Psychiatry*, *33*(6), 521–527. https://doi.org/10.1097/yco.0000000000000641

Vogt, D., Smith, B. N., King, L. A., King, D. W., Knight, J., & Vasterling, J. J. (2013). Deployment risk and resilience inventory-2 (DRRI-2): An updated tool for assessing psychosocial risk and resilience factors among service members and veterans. *Journal of Traumatic Stress*, *26*(6), 710–717. https://doi.org/10.1002/jts.21868

Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., Stasik, S. M., & Ruggero, C. J. (2012). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment*, *19*(4), 399–420. https://doi.org/10.1177/1073191112449857

Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., Foa, E. B., Young-McCaughan, S., Yarvis, J. S., Hembree, E. A., Mintz, J., Peterson, A. L., & Litz, B. T. (2016). Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychological Assessment*, *28*(11), 1392–1403. https://doi.org/10.1037/pas0000260

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.