

# Evaluating the Evidence for Brain-Based Biotypes of Psychiatric Vulnerability in the Acute Aftermath of Trauma

Ziv Ben-Zion, Ph.D., Tobias R. Spiller, M.D., Jakob N. Keynan, Ph.D., Roei Admon, Ph.D., Ifat Levy, Ph.D., Israel Liberzon, M.D., Ph.D., Arie Y. Shalev, M.D., Ph.D., Talma Hendler, M.D., Ph.D., Ilan Harpaz-Rotem, Ph.D.

**Objective:** The weak link between subjective symptom-based diagnostic methods for posttraumatic psychopathology and objectively measured neurobiological indices forms a barrier to the development of effective personalized treatments. To overcome this problem, recent studies have aimed to stratify psychiatric disorders by identifying consistent subgroups based on objective neural markers. Along these lines, a promising 2021 study by Stevens et al. identified distinct brain-based biotypes associated with different longitudinal patterns of posttraumatic symptoms. Here, the authors conducted a conceptual nonexact replication of that study using a comparable data set from a multimodal longitudinal study of recent trauma survivors.

**Methods:** A total of 130 participants (mean age, 33.61 years, SD=11.21; 48% women) admitted to a general hospital emergency department following trauma exposure underwent demographic, clinical, and neuroimaging assessments 1, 6, and 14 months after trauma. All analyses followed the pipeline outlined in the original study and were conducted in collaboration with its authors.

**Results:** Task-based functional MRI conducted 1 month posttrauma was used to identify four clusters of individuals based on profiles of neural activity reflecting threat and reward reactivity. These clusters were not identical to the previously identified brain-based biotypes and were not associated with prospective symptoms of posttraumatic psychopathology.

**Conclusions:** Overall, these findings suggest that the original brain-based biotypes of trauma resilience and psychopathology may not generalize to other populations. Thus, caution is warranted when attempting to define subtypes of psychiatric vulnerability using neural indices before treatment implications can be fully realized. Additional replication studies are needed to identify more stable and generalizable neuroimaging-based biotypes of posttraumatic psychopathology.

*AJP in Advance* (doi: 10.1176/appi.ajp.20220271)

Unlike in most fields of medicine, diagnosis in psychiatry remains restricted to subjective self-reports and observable symptoms (1). The biological mechanisms underlying psychiatric disorders are complex and still poorly understood. The weak link between established diagnostic methods and objectively measured biological indices (2) forms a barrier to the development of effective personalized treatments (3). To overcome this problem, many studies have aimed to stratify psychiatric disorders in an attempt to identify consistent subgroups based on objective biological markers (4, 5).

Accordingly, a recent innovative study by Stevens et al. (6) aimed to discover brain-based biotypes of trauma resilience and psychopathology in the acute aftermath of trauma, using data from the AURORA longitudinal study of trauma survivors (7). Using functional MRI (fMRI) data obtained during performance of simple and widely used tasks that probe

threat reactivity (8), reward reactivity (9), and inhibitory engagement (10), the authors identified four different clusters of individuals in a discovery cohort (N=69). In an internal replication cohort (N=77), three clusters were replicated: “reactive-disinhibited,” “low reward-high threat,” and “inhibited” clusters (see Figure 2 in the original paper [6]). Those replicated clusters were associated with different longitudinal trajectories of posttraumatic stress disorder (PTSD) and anxiety symptoms. Interestingly, the cluster of individuals showing heightened reactivity to both threat and reward was associated with the highest levels of subsequent PTSD and anxiety symptoms (6).

Multimodal longitudinal studies of the posttraumatic stress response are rare because they present formidable technical and conceptual challenges (11). These challenges include obtaining a large enough sample of recent trauma

survivors; optimal timing of assessments and sufficient follow-up duration to capture critical stages in the longitudinal development of PTSD (12); minimizing subjects' burden; and contacting, enrolling, evaluating, and retaining sensitive clinical populations. To the best of our knowledge, the only existing data set comparable to the AURORA study comes from the Neurobehavioral Moderators of Posttraumatic Disease Trajectories study (11) (ClinicalTrials.gov identifier: NCT03756545).

Here, using a similar data set from recent trauma survivors and closely matched analytic processes, we aimed to perform a conceptual nonexact replication of the Stevens et al. study (6). Independent replications are needed as they represent a fundamental part of science and lead to greater confidence in previously reported findings (13, 14). These replications are particularly relevant to neuroimaging studies because of their large analytical variability (15) and to the field of psychiatry, which suffers from a well-known heterogeneity problem (16). Nevertheless, engaging in replication studies is often undervalued, can be difficult to publish, and has few direct incentives for researchers (17).

The main objective of the present study was to examine whether the previously identified brain-based clusters would generalize to our independent sample of trauma survivors. If these clusters were replicated, we aimed to test whether a cluster characterized by heightened reactivity to both threat and reward would be similarly associated with increased subsequent symptoms of PTSD and anxiety, which would demonstrate the stability of a biological phenomenon across similar measures of psychopathology.

## METHODS

The study was approved by the ethics committee in the local medical center (reference number 0207/14). All participants provided written informed consent in accordance with the Declaration of Helsinki and received financial remuneration at each assessment (1, 6, and 14 months posttrauma).

Data included in this study were collected between 2015 and 2020 as part of the NIMH-funded Neurobehavioral Moderators of Posttraumatic Disease Trajectories study. The study's design and detailed methodologies have been previously published (11), and aspects relevant to the present study are summarized below. Overall, we conducted all analyses as closely as possible to the published analytic pipeline of the original study (6), with the aim of replication, and included similar preprocessing and analysis of the neuroimaging data using fMRIprep (18) and SPM-12 (19), the same anatomical regions of interest (ROIs), and an identical clustering analysis. We used an adapted version of the R code applied in the original study, which was kindly provided by the corresponding author. The methods section here is organized in a similar way to the original publication to further facilitate the comparison between the two studies.

### Participants

Potential subjects for this study were adult civilians 18–65 years old who were consecutively admitted to a general

hospital emergency department after one of the following events: motor vehicle accident, bicycle accident, physical assault, robbery, hostilities, electric shock, fire, drowning, work accident, terror attack, or large-scale disaster. Individuals who had sustained head injuries, were unconscious on admission to the emergency department, or were not able to provide informed consent or comprehend the study's procedures were excluded from the study. Participants with conditions precluding MRI scanning (e.g., pacemaker, metal implants, and large tattoos) and those with current substance use disorder, current suicidal ideation, or lifetime psychotic disorder were also excluded. In contrast to the AURORA study, this study also excluded individuals with a prior diagnosis of PTSD. All participants provided oral consent to the study's screening telephone interview and written informed consent upon attending a subsequent diagnostic and eligibility ascertainment clinical interview.

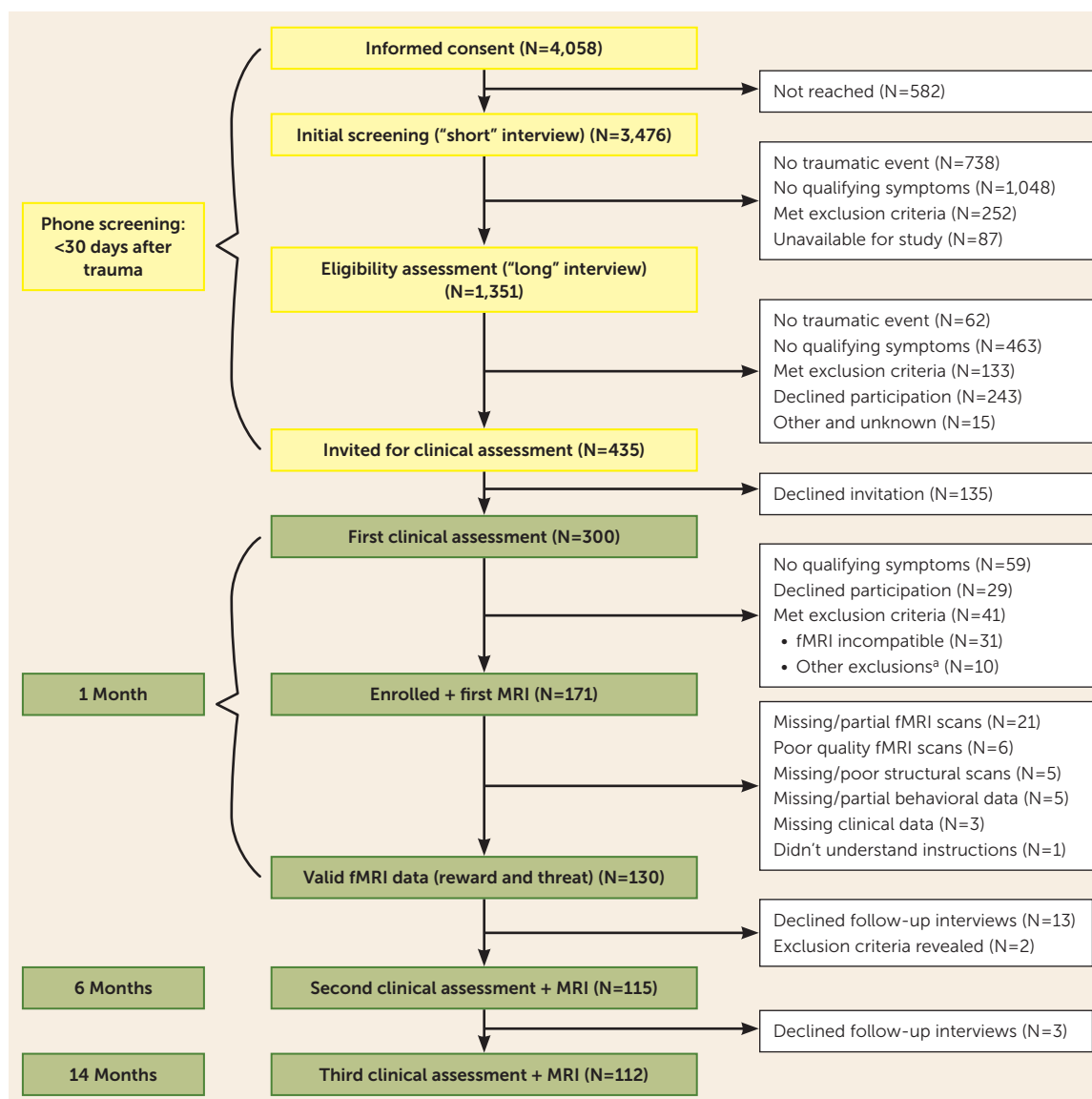
### Clinical Assessments

A comprehensive clinical interview was conducted by trained and certified clinical interviewers using the Clinician-Administered PTSD Scale (CAPS) (20, 21) to assess PTSD diagnosis and severity at each time point. To maintain continuity with decades of DSM-IV-based PTSD research in light of evidence of nonoverlapping groups of individuals diagnosed with PTSD based on DSM-IV or DSM-5 criteria and a consequent recommendation to use a broader PTSD definition for empirical research (22–24), we administered a combined clinical interview scoring both CAPS-4 (DSM-IV) and CAPS-5 (DSM-5) items (20, 21). A positive diagnosis of PTSD was inferred when a participant met either DSM-IV or DSM-5 diagnostic criteria or, in line with previous recommendations (25), had a total score  $\geq 40$  on the CAPS-4. As a secondary continuous measure of PTSD symptom severity, participants completed the PTSD Checklist (PCL) (26), a 17-item self-report questionnaire corresponding to the DSM symptom criteria for PTSD. As a continuous measure of anxiety symptom severity, participants completed the Beck Anxiety Inventory (BAI) (27), a 21-item self-report questionnaire measuring physical and cognitive anxiety symptoms.

### MRI

**Acquisition.** Whole-brain functional and anatomical images were acquired with a 3.0-T Siemens MRI system (MAGNETOM Prisma, Germany) with a 20-channel head coil at the Sagol Brain Institute, Wohl Institute for Advanced Imaging, Tel Aviv Sourasky Medical Center. Functional images were acquired in an interleaved order (anterior to posterior) with a T2\*-weighted gradient-echo planar imaging pulse sequence (TR=2,000 ms, TE=28 ms, flip angle=90°, voxel size=2.2 mm<sup>3</sup>, FOV=220×220 mm, slice thickness=3 mm, 36 slices per volume). A T1-weighted three-dimensional anatomical image was obtained with a magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR=2,400 ms, TE=2.29 ms, flip angle=8°, voxel size=0.7 mm<sup>3</sup>, FOV=224×224 mm) to enable optimal localization of the functional effects.

FIGURE 1. CONSORT flowchart for the “Neurobehavioral Moderators of Posttraumatic Disease Trajectories” study



<sup>a</sup> Other exclusions (N=10) included serious medical condition requiring clinical attention (N=5), chronic PTSD before current event (N=2), current substance use disorder (N=1), head injury (N=1), and no traumatic event (N=1).

*fMRI tasks.* A face-matching task (28) similar to that used by Stevens et al. (6) was used to probe threat reactivity. The safe or risky domino choice paradigm (29) was used to measure reward reactivity, which was not the same as the simpler reward task used by Stevens et al.; nevertheless, the two reward reactivity tasks showed similar neural activations in key regions associated with response to reward in the brain (see results in the online supplement). Unlike the original study, we did not have an fMRI task measuring response inhibition (for a full description of the tasks, see the online supplement).

*fMRI data preprocessing.* Functional images were pre-processed with fMRIPrep, version 1.5.8 (18). Functional imaging scans were coregistered to the anatomical T1-

weighted images, corrected for motion, spatially realigned, slice-time corrected, normalized to the 2009 ICBM-152 template, and smoothed with a 6-mm kernel (for full details, see methods in the online supplement).

*fMRI data analysis.* Similar to the original study (6), the analysis was performed with SPM, version 12 (19). The final sample was restricted to participants with good-quality data across all fMRI tasks (N=130; see below). The ROIs were the same regions used in the original study, which were kindly provided by the corresponding author. They were anatomically defined and included the left and right amygdala, insula, subgenual and dorsal anterior cingulate cortex (sgACC and dACC, respectively), nucleus accumbens (NAcc), and orbitofrontal

**TABLE 1. Demographic and clinical characteristics of the study participants**

Measure	Participants With Trauma Exposure (N=130)	
	Mean	SD
Age (years)	36.61	11.21
Education (years) <sup>a</sup>	14.25	2.74
Clinician-Administered PTSD Scale for the DSM-5 (total score)	24.58	11.41
Clinician-Administered PTSD Scale for the DSM-IV (total score)	51.30	22.45
PTSD Checklist (total score)	45.77	14.39
Beck Anxiety Inventory (total score)	19.91	12.63
	N	%
Gender		
Female	62	48
Male	68	52
Marital status <sup>b</sup>		
Single	89	71
Married	24	19
Divorced	12	10
Trauma type		
Motor vehicle accident	115	88
Assault	10	8
Other	5	4

<sup>a</sup> Eleven participants did not report their education level in years.

<sup>b</sup> Five participants did not provide information about their marital status.

cortex (OFC) (for full details, see methods in the online supplement).

### Procedure

A total of 4,058 consecutive trauma survivors admitted to the emergency department were contacted by telephone within 10–14 days after trauma exposure, were given information about the study, and provided informed assent (Figure 1). Of those, 3,476 individuals underwent initial screening (i.e., “short” interview) and confirmed the occurrence of a psychologically traumatic event and related symptoms; 1,351 individuals subsequently underwent eligibility assessment (i.e., “long” interview), which further assessed acute stress symptoms that were suggestive or indicative of chronic PTSD risk (30). A total of 435 individuals met the inclusion criteria for this study, did not meet any of the exclusion criteria, and were subsequently invited for an in-person clinical interview. Of these individuals, 300 attended the interviews, and 171 also underwent fMRI assessment, both within 1 month after the trauma. Of these, 41 individuals were excluded for the following reasons: missing (N=16) or partial (N=5) functional scans of the threat or reward reactivity tasks; poor-quality functional scans (e.g., excessive movement or artifacts) (N=6); missing or poor-quality structural scans (N=5); missing or partial behavioral data from the tasks (N=5); failure to properly understand the instructions (N=1); or missing clinical data (N=3). A final sample of 130 individuals with valid anatomical and functional brain imaging data from both fMRI tasks were included in this report (Table 1).

### Clustering Analysis

As noted above, we used R code that was adapted from the code applied in the original study (6) with R, version 4.1.1, and RStudio, version 1.4.1717. Hierarchical agglomerated clustering was conducted with data from the ROIs extracted from the two fMRI tasks (i.e., threat and reward reactivity fMRI tasks) with the *cluster* package (version 2.1.2) following Ward’s criterion (*agnes* function). This bottom-up method is designed to preserve the existing data structure, does not impose any assumptions of linearity, and is appropriate for exploratory analyses. The optimal number of clusters was determined using both the silhouette width metric (31) and Hartigan’s distance metric (32).

Cluster analytic algorithms are prone to find clusters even when the underlying data do not contain clusters but are multivariate and normally distributed (33). Therefore, we expanded the analytic plan of Stevens et al. (6) by testing whether the results of our cluster analysis meaningfully differed from the null hypothesis that our data did not contain clusters. For this purpose, we used a procedure reported by Dinga et al. (34) that was based on a procedure originally proposed by Liu et al. (33). In this approach, the null hypothesis is that the data come from a single multidimensional Gaussian distribution, that is, a distribution with no underlying clusters, with the number of dimensions equal to the number of features included in the clustering analysis (for full details, see methods in the online supplement).

### Analysis of Posttrauma Outcomes by Cluster

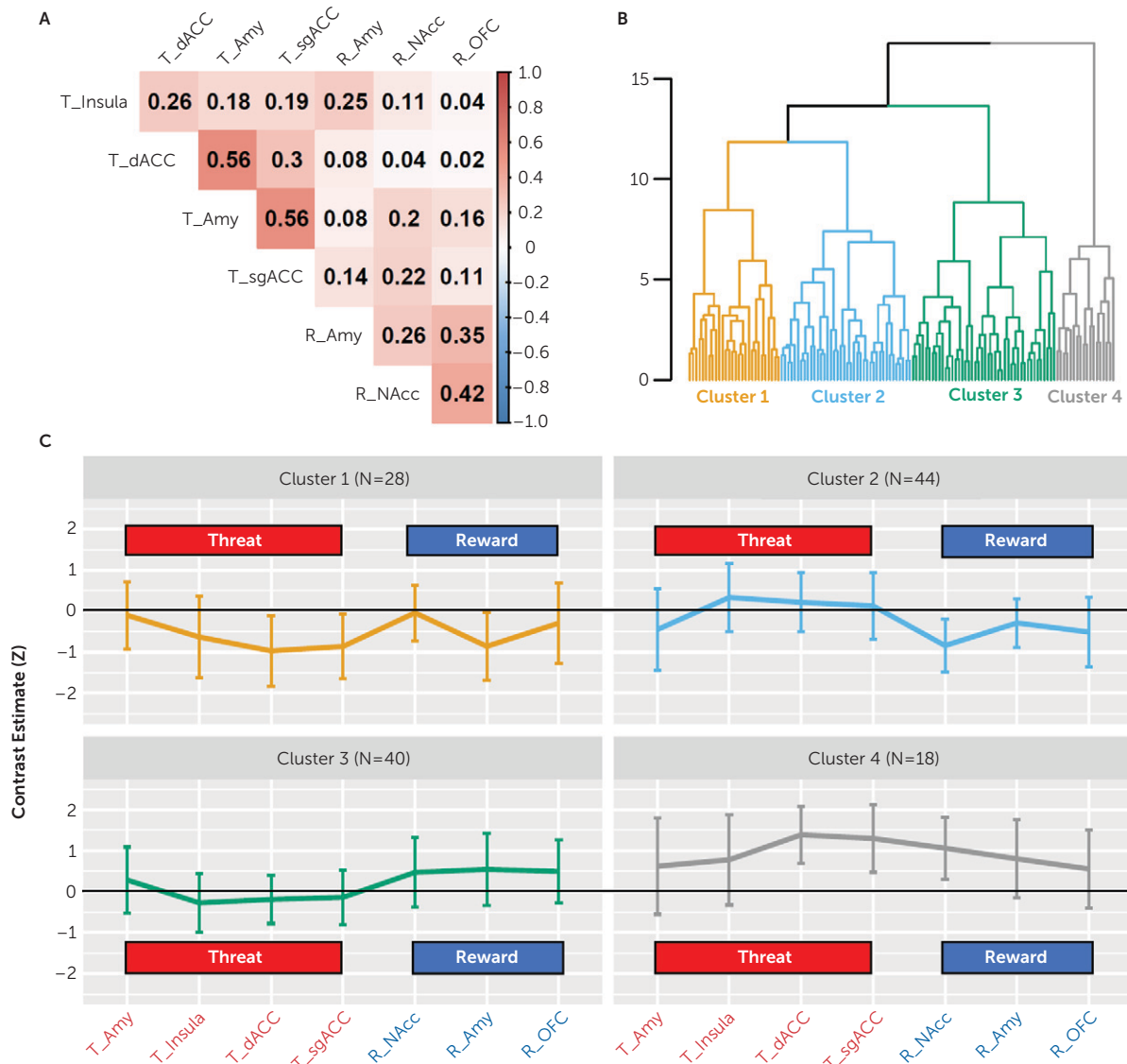
First, because different demographic characteristics might influence the cluster solution due to the unconstrained analytic approach, chi-square tests (categorical variables) and analyses of variance (ANOVAs) (continuous variables) were used to assess whether demographic characteristics differed between the clusters. Second, chi-square tests (categorical variables) and ANOVAs (continuous variables) were used to assess whether the clusters were significantly different in PTSD dichotomous diagnosis (i.e., PTSD or no PTSD), PTSD symptom severity (i.e., CAPS-4 and CAPS-5 total scores), self-reported PTSD symptoms (i.e., PCL scores), and self-reported anxiety symptoms (i.e., BAI scores). Benjamini-Hochberg (35) false discovery rate (FDR) correction ( $q < 0.05$ ) was calculated to control for multiple comparisons of these different clinical measures.

All tests were two-tailed and used a significance threshold of  $p = 0.05$ .

## RESULTS

### Participants’ Demographic and Clinical Characteristics

A sample of 130 recent trauma survivors (mean age, 33.61 years, SD=11.21, range=18–64 years; 62 [48%] women) were included in all the analyses reported below. The most common trauma type among participants was a motor vehicle accident (N=115, 88%); 10 (8%) participants experienced an assault or brawl, and five (4%) participants experienced other traumatic events (for full demographic and clinical characteristics, see Table 1). Similarities and differences in these characteristics between our sample (N=130)

**FIGURE 2. Functional MRI (fMRI) profiles of the four identified clusters of recent trauma survivors (N=130)<sup>a</sup>**

<sup>a</sup> Panel A shows the region-of-interest (ROI) covariance matrix revealing linear associations between z-scored contrast estimates extracted from seven ROIs across the two fMRI tasks. For threat reactivity ("T"), fMRI activation was extracted from the right and left amygdala (Amy), insula, and dorsal and subgenual anterior cingulate cortex (dACC and sgACC, respectively) for the contrast of watching fearful > neutral faces. For reward reactivity ("R"), fMRI activation was extracted from the right and left nucleus accumbens (NAcc), amygdala, and orbitofrontal cortex (OFC) for the contrast of receiving rewards > punishments. The matrix is ordered in the same order as in Figure 2 panels A and B in the original paper (6). Correlation coefficients (R values) are indicated on a scale ranging from -1 (blue) to +1 (red). Panel B shows the dendrogram illustrating the final cluster solution with four clusters (marked by different colors). Panel C shows cluster differences (means and standard deviations) for standardized contrast estimates extracted from the ROIs across the threat (fearful > neutral faces) and reward (rewards > punishments) contrasts of the two different tasks.

and the original cohorts (6) (discovery: N=69; replication: N=77) are reported in the results in the online supplement.

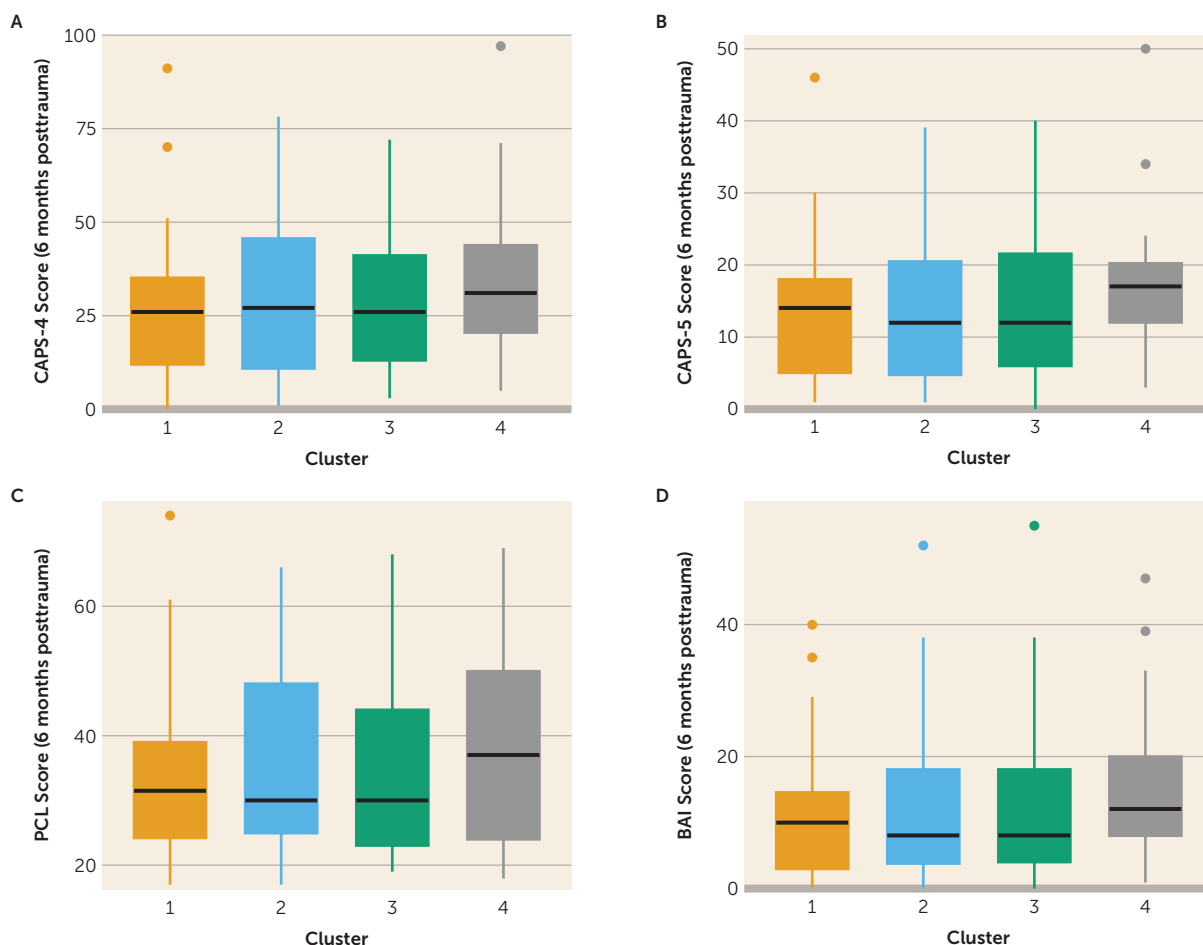
### Covariance Among fMRI Tasks and ROIs

To assess feature redundancy, we examined the covariance structure among the tasks and ROIs. In line with Stevens et al. (6), different ROIs showed high positive within-task covariance but low between-task covariance (Figure 2A). Within the same task, there were moderate to high correlations between different ROIs, with the highest correlations being between reactivity to threat observed in the amygdala and reactivity

to threat observed in both the sgACC and dACC (for both,  $r=0.56$ ,  $p<0.01$ ) (Figure 2A). Similar to the original findings (6), reactivity in a particular region was uncorrelated across the two tasks. Specifically, threat reactivity in the amygdala was not correlated with reward reactivity in the amygdala ( $r=0.08$ ,  $p=0.77$ ) (Figure 2A).

### Clustering of Individuals Using Task-Based fMRI at 1 Month Posttrauma

Hierarchical clustering performed with data from all 130 participants suggested an optimal solution of four clusters ( $k=4$ )

**FIGURE 3. PTSD and anxiety symptom scores at 6 months posttrauma among the four clusters of recent trauma survivors (N=130)<sup>a</sup>**

<sup>a</sup> Box plots presenting clinical symptoms at 6 months posttrauma among the four clusters created based on neuroimaging data at 1 month posttrauma: total scores on the Clinician-Administered PTSD Scale (CAPS-4 [for DSM-IV] in panel A, CAPS-5 [for DSM-5] in panel B), total scores on the PTSD Checklist (PCL) in panel C, and total scores on the Beck Anxiety Inventory (BAI) in panel D.

according to Hartigan's distance metric (see Figure S1a in the online supplement) and two clusters ( $k=2$ ) according to the silhouette width metric (see Figure S1b in the online supplement); we therefore tested both solutions. To avoid redundancy, and in line with Stevens et al. (6), we report the four-cluster solution here (Figure 2B) and the two-cluster solution in the supplemental results (see Figure S2 in the online supplement).

Assessment of different fMRI activation patterns revealed a subgroup of 18 individuals showing high reactivity of all brain regions to threat, predominantly in the dACC and sgACC, and to reward, predominantly in the NAcc (cluster 4 in Figure 2C). The other three clusters were more similar to each other and were indeed part of the same cluster according to the two-cluster solution (clusters 1–3 in Figure 2C). While individuals in cluster 1 ( $N=28$ ) showed low reactivity to both threat (in the dACC and sgACC) and reward (in the amygdala), those in cluster 2 ( $N=44$ ) showed low reward reactivity (predominantly in the NAcc) and those in cluster 3 ( $N=40$ ) showed high reward reactivity (similar

levels across the three regions), with relatively no threat reactivity (clusters 1–3 in Figure 2C).

Notably, as in the original study, the clusters were unrelated to any of the demographic characteristics. There was no significant association between cluster assignment and participants' age ( $F=0.492$ ,  $p=0.688$ ), years of education ( $F=0.412$ ,  $p=0.745$ ), gender ( $\chi^2=0.865$ ,  $p=0.834$ ), or marital status ( $\chi^2=2.032$ ,  $p=0.236$ ).

Finally, following Dinga et al. (34) and Liu et al. (33), we tested the statistical significance of the observed Hartigan distance index. In our data set, the four-cluster solution resulted in the optimal Hartigan's distance index. Using a simulation approach (described in the methods section in the online supplement), we found that this index was not statistically significant (Hartigan's distance index=18.15,  $p=0.371$ ) (see Figure S1c in the online supplement). In other words, it is not unusual to observe such an index even when the hierarchical clustering is performed on a multivariate, normally distributed data set with no clusters.



**TABLE 2. Statistics on prospective clinical outcomes among the four different clusters<sup>a</sup>**

Variable	6 Months Posttrauma			14 Months Posttrauma		
	$\chi^2$	p	$P_{FDR}$	$\chi^2$	p	$P_{FDR}$
PTSD diagnosis (PTSD/no PTSD)	2.285	0.515	1	0.011	0.999	0.999
	F	p	$P_{FDR}$	F	p	$P_{FDR}$
Clinician-Administered PTSD Scale for DSM-IV (total score)	0.411	0.746	1	0.505	0.679	1
Clinician-Administered PTSD Scale for DSM-5 (total score)	0.558	0.644	1	0.335	0.800	1
PTSD Checklist (total score)	0.502	0.682	1	0.443	0.723	1
Beck Anxiety Inventory (total score)	0.555	0.646	1	0.312	0.816	0.907

<sup>a</sup> Results of chi-square tests to assess differences in clinical diagnosis (PTSD or no PTSD) and analyses of variance (continuous variables) to assess differences in symptom severity (total scores) among the four suggested clusters. FDR=false discovery rate.

### Prospective Trajectories of Mental Health Among the Different Clusters

The four clusters did not differ in prospective 6-month PTSD dichotomous diagnosis (i.e., PTSD or no PTSD), PTSD symptom severity (i.e., CAPS-4 or CAPS-5 scores), self-reported PTSD (i.e., PCL scores), or anxiety symptom severity (i.e., BAI scores) (Figure 3 and Table 2). Similarly, these four clusters did not differ in any clinical measure at 14 months posttrauma (Table 2; see also Figure S3 in the online supplement). Statistical significance further decreased after applying the FDR correction for multiple comparisons (35) (for all comparisons,  $0.907 \leq P_{FDR} \leq 1.000$ ) (Table 2). In summary, there was no association between individuals' cluster membership and PTSD or anxiety at 6 months posttrauma (the original study's endpoint) or at 14 months posttrauma (this study's endpoint).

### DISCUSSION

In this conceptual nonexact replication and extension of the Stevens et al. study (6), we failed to replicate the previously identified neuroimaging-based biotypes (6) or their association with prospective posttraumatic stress symptoms. Despite overall similarities in study design and aims, participant characteristics, and fMRI probes, ours is not an exact replication of the original study. Nevertheless, nonexact replications can provide strong evidence for robustness and external validity of previous findings and demonstrate generalization beyond specific study design choices and populations (36). On the other hand, when original findings are not replicated (37), it is hard to determine whether it is because of the methodological differences or because the original findings were false positive ones.

There are several potential explanations for our inability to replicate the results of Stevens et al. (6), mainly due to methodological differences between the studies. First, replication data were obtained from a single site in Israel, compared to original data collected at several different sites across the United States. Although some demographic

characteristics were similar in the two studies (e.g., participant ages), other characteristics differed (e.g., gender and trauma type) or were not collected in the present study (e.g., race/ethnicity and childhood trauma) (Table 1; see also results in the online supplement). Second, while this study specifically screened participants for early posttraumatic stress symptoms and excluded individuals with prior PTSD diagnoses, the original study did not have these constraints. Third, while our neuroimaging data were collected 1 month posttrauma (mean = 30.43 days [SD = 9.54] posttrauma), MRI data in the original study were obtained at a slightly earlier time point (mean = 21 days [SD = 6] posttrauma). Fourth, while Stevens et al. (6) assessed symptoms based on abbreviated self-report tools (a limitation noted by the authors), we used gold-standard structured interviews (CAPS) administered by trained and certified clinicians. These assessments were performed at 1, 6, and 14 months following trauma exposure, whereas the assessments in the original study ranged from 30 days pretrauma (queried retrospectively in the emergency department) to 6 months posttrauma (for a total of five assessments). Importantly, a 6-month follow-up duration is a dynamic time point in the course toward the tangible chronic PTSD subtype (38), whereas a 14-month follow-up is clinically stable and indicative of the chronic disease, as further recovery is marginal (12). Finally, while clustering in the original study was performed with nine neuroimaging measures from three different tasks, in the present study, clustering was based on seven measures from two different tasks (we did not include an fMRI inhibition task). Furthermore, although both studies used the same fMRI task to probe threat reactivity, different tasks were used to assess reward reactivity. Nevertheless, both reward tasks showed similar whole-brain activations in the contrast of reward-gain versus punishment-loss (see results in the online supplement).

The large clinical heterogeneity of posttraumatic psychopathology, together with recent advances in statistical and computational methods, motivated the search for homogeneous PTSD subtypes through data-driven approaches (39). Nevertheless, the presumption of distinct and homogeneous

subgroups might not be clinically useful or represent the biology underlying PTSD (34). For example, most clustering approaches will always yield clusters regardless of the data structure, even if there are no clusters at all (33). Hence, it is crucial to distinguish biologically and clinically meaningful subtypes from random data fluctuations or noise (33). Here, following a procedure by Dinga et al. (34), we showed that the four-cluster solution could be found even if the data came from a single Gaussian distribution with no underlying clusters; this finding supports the fact that four clusters (as shown in Figure 2C) were not more likely than no clusters at all. This statistical test was not performed by Stevens et al. (6). Another way to deal with the disadvantages of data-driven approaches is the use of hybrid analytic methods (16), which combine prior knowledge and assumptions (theory-driven and supervised) with data-driven (unsupervised) approaches (40).

In conclusion, our results highlight that slight changes in sample characteristics or experimental tasks can have a critical impact on the replicability of neuroimaging-based biotypes and their association with posttraumatic stress symptoms. This is in line with recent findings suggesting smaller than expected brain-phenotype associations and large variability across population subsamples (41), as would be expected from a disorder with over 600,000 potential phenotypes (42). Therefore, studies should carefully specify their design and methodology to define populations to which the results could be generalized until more stable and unified measures are established in psychiatry. Importantly, caution is warranted when attempting to define PTSD subtypes using neuroimaging data before treatment implications can be fully realized. Future replication studies may assist in closing the translational gap between basic psychiatric research and practice, advancing the development of meaningful biological tools to assist diagnosis or predict clinical outcomes (3).

#### AUTHOR AND ARTICLE INFORMATION

Yale School of Medicine (Ben-Zion, Spiller, Levy, Harpaz-Rotem) and Wu Tsai Institute and Department of Psychology (Levy, Harpaz-Rotem), Yale University, New Haven; U.S. Department of Veterans Affairs National Center for PTSD Clinical Neuroscience Division, VA Connecticut Healthcare System, West Haven (Ben-Zion, Spiller, Harpaz-Rotem); Sagol Brain Institute, Wohl Institute for Advanced Imaging, Tel Aviv Sourasky Medical Center (Ben-Zion, Shalev, Hendler), Sagol School of Neuroscience (Ben-Zion, Hendler), and Faculty of Social Sciences and Sackler Faculty of Medicine (Hendler), Tel Aviv University, Tel Aviv, Israel; Department of Consultation-Liaison Psychiatry and Psychosomatic Medicine, University Hospital Zurich, University of Zurich, Zurich (Spiller); Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford (Keynan); School of Psychological Sciences and Integrated Brain and Behavior Research Center, University of Haifa, Haifa, Israel (Admon); Department of Psychiatry, College of Medicine, Texas A&M University, College Station (Liberzon); Department of Psychiatry, NYU Grossman School of Medicine, New York (Shalev).

Send correspondence to Dr. Ben-Zion (ziv.ben-zion@yale.edu).

Supported by NIMH grant R01-MH-103287 to Drs. Shalev, Liberzon, and Hendler.

Dr. Ben-Zion was supported by fellowships from the Sagol School of Neuroscience at Tel Aviv University, Sagol Brain Institute at Tel Aviv Sourasky Medical Center, Yale School of Medicine, and Fulbright U.S.-Israel Program. Dr. Spiller was supported by a fellowship from the Yale School of Medicine.

The authors thank Jennifer Stevens, Ph.D., for her collaboration and willingness to share her materials and methods (namely, the R code used for the clustering analysis and the anatomical regions of interest). The authors also thank the research team at Tel Aviv Sourasky Medical Center (including Naomi Fine, Nili Green, Mor Halevi, Sheli Luvton, Yael Shavit, Olga Nevenchannaya, Iris Rashap, Efrat Routledge, and Ophir Leshets) for their significant contributions to participant screening, enrollment, and assessments. The authors also extend their gratitude to all the participants of this study, who completed assessments at three different time points after experiencing a traumatic event, thus contributing to scientific research of posttraumatic psychopathology.

Dr. Hendler is the chief medical officer of GrayMatters Health, Haifa, Israel. The other authors report no financial relationships with commercial interests.

ClinicalTrials.gov identifier: NCT03756545.

Received March 26, 2022; revisions received June 6 and July 6, 2022; accepted July 11, 2022.

#### REFERENCES

- Insel TR, Cuthbert BN: Medicine. Brain disorders? Precisely. *Science* 2015; 348:499–500
- Hyman SE: Can neuroscience be integrated into the DSM-V? *Nat Rev Neurosci* 2007; 8:725–732
- Kapur S, Phillips AG, Insel TR: Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry* 2012; 17:1174–1179
- Marquand AF, Wolfers T, Mennes M, et al: Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2016; 1:433–447
- Maron-Katz A, Zhang Y, Narayan M, et al: Individual patterns of abnormality in resting-state functional connectivity reveal two data-driven PTSD subgroups. *Am J Psychiatry* 2020; 177:244–253
- Stevens JS, Harnett NG, Lebois LAM, et al: Brain-based biotypes of psychiatric vulnerability in the acute aftermath of trauma. *Am J Psychiatry* 2021; 178:1037–1049
- McLean SA, Ressler K, Koenen KC, et al: The AURORA Study: a longitudinal, multimodal library of brain biology and function after traumatic stress exposure. *Mol Psychiatry* 2020; 25:283–296
- Stevens JS, Kim YJ, Galatzer-Levy IR, et al: Amygdala reactivity and anterior cingulate habituation predict posttraumatic stress disorder symptom maintenance after acute civilian trauma. *Biol Psychiatry* 2017; 81:1023–1029
- Speer ME, Bhanji JP, Delgado MR: Savoring the past: positive memories evoke value representations in the striatum. *Neuron* 2014; 84:847–856
- Jovanovic T, Ely T, Fani N, et al: Reduced neural activation during an inhibition task is associated with impaired fear inhibition in a traumatized civilian sample. *Cortex* 2013; 49:1884–1891
- Ben-Zion Z, Fine NB, Keynan NJ, et al: Neurobehavioral moderators of post-traumatic stress disorder (PTSD) trajectories: study protocol of a prospective MRI study of recent trauma survivors. *Eur J Psychotraumatol* 2019; 10:1683941
- Shalev AY, Freedman S: PTSD following terrorist attacks: a prospective evaluation. *Am J Psychiatry* 2005; 162:1188–1191
- McNutt M: Reproducibility. *Science* 2014; 343:229
- Baker M: 1,500 scientists lift the lid on reproducibility. *Nature* 2016; 533:452–454



15. Botvinik-Nezer R, Holzmeister F, Camerer CF, et al: Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* 2020; 582:84–88
16. Feczko E, Miranda-Dominguez O, Marr M, et al: The heterogeneity problem: approaches to identify psychiatric subtypes. *Trends Cogn Sci* 2019; 23:584–601
17. McEwan B, Carpenter CJ, Westerman D: On replication in communication science. *Commun Stud* 2018; 69:235–241
18. Esteban O, Markiewicz CJ, Blair RW, et al: fMRIPrep: a robust pre-processing pipeline for functional MRI. *Nat Methods* 2019; 16:111–116
19. Friston KJ, Ashburner JT, Kiebel SJ, et al: *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. London, Academic Press, 2007
20. Blake DD, Weathers FW, Nagy LM, et al: The development of a Clinician-Administered PTSD Scale. *J Trauma Stress* 1995; 8:75–90
21. Weathers FW, Bovin MJ, Lee DJ, et al: The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans. *Psychol Assess* 2018; 30:383–395
22. Hoge CW, Yehuda R, Castro CA, et al: Unintended consequences of changing the definition of posttraumatic stress disorder in DSM-5: critique and call for action. *JAMA Psychiatry* 2016; 73:750–752
23. Hoge CW, Riviere LA, Wilk JE, et al: The prevalence of post-traumatic stress disorder (PTSD) in US combat soldiers: a head-to-head comparison of DSM-5 versus DSM-IV-TR symptom criteria with the PTSD checklist. *Lancet Psychiatry* 2014; 1:269–277
24. Stein DJ, McLaughlin KA, Koenen KC, et al: DSM-5 and ICD-11 definitions of posttraumatic stress disorder: investigating “narrow” and “broad” approaches. *Depress Anxiety* 2014; 31:494–505
25. Weathers FW, Ruscio AM, Keane TM: Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychol Assess* 1999; 11:124–133
26. Weathers FW, Litz BT, Herman DS, et al: The PTSD checklist: reliability, validity, and diagnostic utility. Annual meeting of the International Society for Traumatic Stress Studies, San Antonio, Tex 1993
27. Beck AT, Epstein N, Brown G, et al. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*. 1988;56:893–897
28. Hariri AR, Bookheimer SY, Mazziotta JC: Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport* 2000; 11:43–48
29. Ben-Zion Z, Shany O, Admon R, et al: Neural responsivity to reward versus punishment shortly after trauma predicts long-term development of posttraumatic stress symptoms. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2022; 7:150–161
30. Shalev AY, Ankri Y, Israeli-Shalev Y, et al: Prevention of post-traumatic stress disorder by early treatment: results from the Jerusalem Trauma Outreach and Prevention study. *Arch Gen Psychiatry* 2012; 69:166–176
31. Rousseeuw PJ: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987; 20:53–65
32. Hartigan JA: *Clustering Algorithms*. New York, John Wiley & Sons, 1975
33. Liu Y, Hayes DN, Nobel A, et al: Statistical significance of clustering for high-dimension, low-sample size data. *J Am Stat Assoc* 2008; 103:1281–1293
34. Dinga R, Schmaal L, Penninx BWJH, et al: Evaluating the evidence for biotypes of depression: methodological replication and extension of Drysdale et al (2017). *Neuroimage Clin* 2019; 22:101796
35. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological)* 1995; 57:289–300
36. Rosenthal R: Replication in behavioral research. *J Soc Behav Personal* 1990; 5:1–30
37. Esterman M, Stumps A, Jagger-Rickels A, et al: Evaluating the evidence for a neuroimaging subtype of posttraumatic stress disorder. *Sci Transl Med* 2020; 12:eaaz9343
38. Kessler RC, Sonnega A, Bromet E, et al: Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1995; 52:1048–1060
39. Sheynin S, Wolf L, Ben-Zion Z, et al: Deep learning model of fMRI connectivity predicts PTSD symptom trajectories in recent trauma survivors. *Neuroimage* 2021; 238:118242
40. Ben-Zion Z, Zeevi Y, Keynan NJ, et al: Multi-domain potential biomarkers for post-traumatic stress disorder (PTSD) severity in recent trauma survivors. *Transl Psychiatry* 2020; 10:208
41. Marek S, Tervo-Clemmens B, Calabro FJ, et al: Reproducible brain-wide association studies require thousands of individuals. *Nature* 2022; 603:654–660
42. Galatzer-Levy IR, Bryant RA: 636,120 ways to have posttraumatic stress disorder. *Perspect Psychol Sci* 2013; 8:651–662