American Psychological Association

Journal of Abnormal Psychology

0021-843X/18/$12.00

http://dx.doi.org/10.1037/abn0000395

Suicidal Thoughts and Behaviors Are Associated With an Increased Decision-Making Bias for Active Responses to Escape Aversive States

Alexander J. Millner
Harvard University

Hanneke E. M. den Ouden
Radboud University

Samuel J. Gershman
Harvard University

Catherine R. Glenn and Jaclyn C. Kearns
University of Rochester

Aaron M. Bornstein
Princeton University

Brian P. Marx and Terence M. Keane
Veterans Affairs Boston Healthcare System, Boston, Massachusetts, and Boston University School of Medicine

Matthew K. Nock
Harvard University

Suicide is a leading cause of death worldwide. Despite decades of clinical and theoretical accounts that suggest that suicidal thoughts and behaviors are efforts to escape painful emotions, little prior research has examined decision making involved in escaping aversive states. We compared the performance of 85 suicidal participants to 44 nonsuicidal psychiatric patients on a novel reinforcement learning task with choices to make either active (i.e., "go") or passive responses (i.e., "no-go") to either escape or avoid an aversive stimulus. We used a computational cognitive model to isolate decision-making biases. We hypothesized that suicidal participants would exhibit a relatively elevated bias for making active responses to escape an aversive state and would show worse performance when escape required a passive response (i.e., "doing nothing" to escape). Our hypotheses were supported: The computational model revealed that suicidal participants exhibited a higher bias for an active response to escape compared with nonsuicidal psychiatric controls, suggesting that this finding was not just the result of the presence of psychopathology. The bias parameter also accounted for unique variance in predicting group status among several constructs previously related to suicidal thoughts and behaviors. This study provides a new method for testing escape decision making and does so using a computational cognitive model, allowing us to precisely index processes underlying suicidal and related behaviors. Future research examining escape decision making from a computational perspective could help link neural processes or environmental stressors to suicidal thoughts or behaviors.

Alexander J. Millner, Department of Psychology, Harvard University; Hanneke E. M. den Ouden, Donders Institute for Brain, Cognition and Behaviour, Radboud University; Samuel J. Gershman, Department of Psychology, Harvard University; Catherine R. Glenn and Jaclyn C. Kearns, Department of Psychology, University of Rochester; Aaron M. Bornstein, Neuroscience Institute, Princeton University; Brian P. Marx and Terence M. Keane, National Center for PTSD, Veterans Affairs Boston Healthcare System, Boston, Massachusetts, and Department of Psychiatry, Boston University School of Medicine; Matthew K. Nock, Department of Psychology, Harvard University.

This research was funded by the Military Suicide Research Consortium (MSRC), an effort supported by the Office of the Assistant Security of Defense for Health Affairs (Award W81XWH-10-2-0181; Alexander J. Millner, Catherine R. Glenn, Brian P. Marx, Terence M. Keane, Matthew K. Nock). Opinions, interpretations, conclusions, and recommendations are those of the authors and not necessarily endorsed by MSRC or the Department of Defense. This project also was supported by Grant PDF-1-094-15 awarded to Alexander J. Millner from the American Foundation for Suicide Prevention. The content is solely the responsibility of the authors and does not necessarily represent the official views of the American Foundation for Suicide Prevention.

All materials, data, code, results, and the results of a full Bayesian analysis at the group level are located on the Alexander J. Millner's Open Science Framework site (https://osf.io/hz3ag/).

Correspondence concerning this article should be addressed to Alexander J. Millner, Department of Psychology, Harvard University, William James Hall 1206, 33 Kirkland Street, Cambridge, MA 02138. E-mail: amillner@fas.harvard.edu

General Scientific Summary
Longstanding theories suggest that suicidal thoughts and behaviors are attempts to escape emotional pain, yet little work has examined decision making involved in escaping aversive states. We found that when in an aversive state, suicidal people show a decision-making bias to “do something” to escape the aversive state resulting in poorer performance when “doing nothing” was the best option to gain relief.

Keywords: suicide, suicidal thoughts and behaviors, escape, decision making, computational modeling

Supplemental materials: http://dx.doi.org/10.1037/abn0000395.supp

Suicide is a leading cause of death worldwide (Naghavi et al., 2017) and nonfatal suicidal thoughts and behaviors (STB) are associated with risk of future suicide (Franklin et al., 2017). Despite decades of research seeking to identify and treat risk factors, rates of suicide remain virtually unchanged over the last century (Carter et al., 2006; Centers for Disease Control and Prevention, 2018). Furthermore, over the last 15 years, the suicide rate has increased substantially (Gibbons, Hur, & Mann, 2017), particularly among veterans (Lyon, 2017). Novel approaches are needed to improve the understanding, prediction and prevention of STB. Clinical anecdotes and theoretical work consistently suggest that suicide is an attempt to escape intolerable emotional pain (Baumeister, 1990; Robins, 1981; Shneidman, 1998). A prominent suicide scholar argued that every case of suicidal thoughts and behaviors is driven by a desire to escape mental pain (Shneidman, 1998). Suicidal thoughts may provide relief because they allow people to imagine a state (i.e., death) where they have escaped their psychological pain. A recent real-time monitoring study supported this idea by showing that suicidal thoughts are associated with reductions in negative affect, potentially leading to the reinforcement of such thoughts (Kleiman et al., 2018). Despite this purportedly important role of escape in suicidal behaviors, there has been almost no research on psychological processes related to decision making to escape aversive states that might provide insight into the question of why people think about killing themselves and attempt suicide.

We recently developed a behavioral measure to assess decision processes involved in escaping an aversive state (i.e., “escape decision making”). This earlier study, using a normative sample, showed that when people are in an aversive state and are acting to escape, they exhibit a fairly automatic bias for active, as opposed to passive, responses (Millner, Gershman, Nock, & den Ouden, 2018). That is, people showed a bias for “doing something” (i.e., pressing a key) to escape an aversive state, and therefore performed more poorly when “doing nothing” was more likely to provide relief (i.e., an “active-escape bias”). The bias was reversed when people were in a neutral state, acting to avoid an impending aversive outcome; there was a bias for passive relative to active responses (i.e., an “inhibitory-avoid bias”). Prior work showed a similar interaction when acting to obtain reward or avoid punishment, where like for escape, reward was associated with active responses (Guitart-Masip et al., 2012).

These types of condition-by-action interactions are likely due to a relatively automatic form of behavioral control that selects responses by rigidly coupling certain actions to particular stimuli (i.e., Pavlovian control). An alternate form of behavioral control uses prior feedback to adaptively select actions that maximize advantageous outcomes over the long-run (i.e., trial-and-error learning or instrumental control). Perhaps due to evolution, the more automatic mechanism often selects the same advantageous choices as the adaptive form of control, making learning efficient. However, sometimes the presence of a stimulus (e.g., a reward or punishment) causes the more automatic mechanism to rigidly select an action that is disadvantageous in a given context. This is the purported cause of the Condition × Action interactions mentioned previously. For example, the presence of an ongoing aversive stimulus causes the more automatic active response to escape, even in contexts where passive responses are more likely to terminate the aversive state.

The disadvantageous behaviors purportedly caused by aberrant forms of this more automatic decision-making mechanism are tied to some psychiatric conditions, such as anxiety (Mkrtchian, Aylward, Dayan, Roiser, & Robinson, 2017; Mkrtchian, Roiser, & Robinson, 2017), trauma-related conditions (Ousdal et al., 2017), and major depression (Huys et al., 2016). In these disorders, aberrant biases vary in specific and mostly expected ways, with anxiety/trauma associated with increased avoidance, and depression with a decreased reward bias. Moreover, in depression, the degree of an intact bias to avoid punishment predicted recovery 4 months to 6 months after assessment (Huys et al., 2016). Crucially, all prior studies have used paradigms that examine biases to reward and punishment; none have examined escape, which may be important in understanding STB. To be clear, during “punishment,” choices are made in a neutral state to avoid a possible aversive state, whereas during “escape” choices are made while in an aversive state to gain relief from such a state.

These more automatic biases often are evident from average choice and reaction time (RT) performance on behavioral tasks (Millner et al., 2018). However, above and beyond average behavioral indices, computational models provide more fine-grained, latent constructs of task performance. Computational models are quantified hypotheses (i.e., a model) about the psychological processes underlying task performance. Models allow researchers to analyze each subject’s task performance dynamically, on a trial-by-trial basis, to parse task variance into different model parameters representing hidden/latent psychological processes. Thus, our prior study included a computational model to isolate the active-escape and inhibitory-avoid biases.

The goal of the current study was to examine aberrant decision-making biases among suicidal people, using the empirical and computational framework established in our prior study (Millner et al., 2018). We focused exclusively on decision making within a negative context, using a negative reinforcement learning paradigm, in which participants had to choose between active (i.e., "go") and passive (i.e., "no-go") responses to either escape an ongoing aversive state or avoid an impending aversive state and fit a computational model established in our prior study (Millner et al., 2018). Given that suicidal thoughts and behaviors are theorized to be efforts to escape aversive states, we anticipated that, compared with people with psychiatric conditions but without suicidal thoughts, suicidal individuals would show a stronger active-escape bias (i.e., a bias to "do something" to escape) relative to a bias for active responses to avoid. Finally, we conducted this study among a sample of military veterans given that this group is at elevated risk of suicide and in need of greater suicide prediction and prevention efforts (Bossarte, Claassen, & Knox, 2010).

Method

Participants

We recruited 158 veteran participants from Boston-area veteran treatment centers. Most of the sample (n = 127) were recruited from a Department of Veterans Affairs (VA) hospital psychiatric inpatient unit (93 with prior STB) and VA outpatient treatment provider (n = 31; 15 with prior STB). Having participants from outside the inpatient unit helps mitigate Berkson's bias where increased severity required to be admitted to an inpatient unit biases the sample, potentially causing erroneous conclusions (Westreich, 2012). Four participants either withdrew from the study (n = 3) or declined to complete the task (n = 1). The first 25 participants were excluded because low accuracy caused us to provide more detailed instructions and add more practice and task trials. After all exclusions, data from 129 participants were analyzed; 85 with a lifetime history of STB and 44 with a psychiatric history, but no lifetime history of STB.

The study was administered as part of a larger battery of measures and interviews described more fully in the online supplemental material. Originally, study recruitment goals were to collect equal numbers of those with and without last year suicidal ideation. After recruitment was completed but prior to any data analyses, we reassigned participants to groups based on the lifetime presence or absence of STB because prior studies have shown that those with lifetime history of STB show differences from those without any STB history on tasks relevant to STBs (Glenn et al., 2017). Within the lifetime suicidal group, more participants had last year STB (n = 56) than lifetime but not last year STB (n = 29). Additionally, 41 suicidal participants had previously attempted suicide, with most having attempts within the prior year (n = 26) but also nearly a quarter (n = 10) having their most recent attempt greater than 10 years prior. The lifetime STB and psychiatric control groups were well matched on intelligence and current diagnosed mental disorders, but the suicidal group was significantly younger, had fewer years of education, and more clinically severe scores on every continuous clinical measure (see Table 1). Participants were compensated $60. The VA Boston Healthcare System (Institutional Review Board [IRB] 2773) and Harvard University (IRB 13-1390) IRBs approved the study. Participants were contacted 1 and 3 months after testing to assess STB occurring since baseline. Retention was low

Table 1
Group Demographic and Clinical Information

Demographic or clinical variable	Suicidal (lifetime STB) (n = 85)	Nonsuicidal (psychiatric history and no history of STB) (n = 44)	Statistic	Effect size (95% CI)
Age[a]	41.79 (12.7)	43.33 (14.8)	−7.7*	−.11 (−.51−.28)
Sex (male)[b]	68.2 (58)	83.7 (36)	3.5	.027 (<.001−.10)
Race (White)[b]	81.2 (69)	81.4 (35)	>.001	>.001 (<.001−.04)
Education[a]	3.40 (1.7)	3.70 (1.7)	−11.7*	−.17 (−.54−.21)
Intelligence[a]	105.7 (10.4)	105.7 (10.5)	−.2	−.003 (−.37−.37)
PTSD EMR[b]	70.6 (60)	76.2 (32)	.4	.006 (<.001–.05)
MDD EMR[b]	54.1 (46)	47.6 (20)	.5	.004 (<.001–.06)
Bipolar EMR[b]	15.3 (13)	19 (8)	.3	.002 (<.001–.05)
Substance dep/abuse EMR[b]	74.1 (63)	76.2 (32)	.1	>.001 (<.001–.04)
Mood disorder NOS EMR[b]	11.8 (10)	2.4 (1)	3.1	.02 (<.001−.06)
Suicidal ideation (SSI)[a]	9.17 (9.5)	.26 (1.1)	77.7*	1.14 (.92−1.40)
Depression (PHQ-9)[a]	17.99 (7.2)	13.56 (5.7)	44.7*	.66 (.29−1.03)
Hopelessness (BHS)[a]	10.62 (2.4)	9.84 (2.4)	21.7*	.32 (−.11−.69)
Borderline personality disorder traits (MSI-BPD)[a]	6.34 (2.4)	3.72 (2.2)	74.7*	1.10 (.70−1.51)
PTSD symptoms (PCL-5)[a]	49.63 (16.2)	40.79 (16.3)	36.7*	.54 (.16−.93)

Note. Sex is presented as the percentage that are male; Race is presented as the percentage that are White. STB = suicidal thoughts and behaviors; PTSD = Post-Traumatic Stress Disorder; MDD = Major Depressive Disorder; Substance dep/abuse = Substance Dependence Disorder or Substance Abuse Disorder; EMR = electronic medical record, current diagnosis, mood disorder; NOS = mood disorder, not otherwise specified; SSI = Beck Suicide Scale for Ideation; PHQ-9 = Patient Health Questionnaire; BHS = Beck Hopelessness Scale; MSI-BPD = McLean Screening Instrument for Borderline Personality Disorder; PCL-5 = The Post-Traumatic Stress Disorder Checklist for *DSM–5*.
[a] Reported as mean (standard deviation), groups compared with a t test and Cohen's d effect size reported. [b] Reported as percentage (number), groups compared with a χ^2 test and Cramer's V effect size reported.
* $p < .05$.

(44% to 55%) and all results from the follow-up were null (see the online supplemental material).

Measures

Behavioral task. The task (see Figure 1) is nearly identical to our previous study (Millner et al., 2018). Four cues (fractal images) were presented either with the simultaneous onset of an aversive sound (escape condition) or not (avoid condition). The aversive sound stimulus was the sound of a fork scraping on slate altered with a high-frequency sound presented over headphones at 80 dB to 85 dB. After making a choice, participants received feedback, which consisted of either the aversive sound or silence. Participants had to learn which response (press a button, i.e., go, or withhold a button press, i.e., no-go) was more likely to result in silence. Thus, in the escape condition, participants chose the response that turned off the sound, whereas, in the avoid condition, they chose the response that prevented the sound from coming on. There were four total cues, go-to-avoid, go-to-escape, no-go-to-avoid, no-go-to-escape, that resulted in a 2 (response: go, no-go) × 2 (condition: escape, avoid) design. Each cue was associated with one optimal response (go or no-go) that resulted in silence 80% of the time and the aversive sound 20%. The other response resulted in the sound 80% and silence 20% of the time. Like in Millner et al. (2018), for no-go responses cue duration was 2 s, and feedback duration was 2 s when the selected choice resulted in the aversive sound. The main differences in the current paradigm from the one in Millner et al. (2018) are (1) if feedback was silence (i.e., no sound), then it was presented for 750 ms to shorten total task duration, (2) if a go response immediately led to feedback, however (3) if a go response resulted in the aversive sound, then feedback duration was adjusted so that that the entire target and feedback duration was 4 s total (so that go responses that led to aversive sound feedback were not shorter than no-go responses with sound feedback). The fact that go responses moved trials to the feedback phase was a possible confound as, during escape trials, go responses could terminate the sound faster than no-go responses; however, effects were similar when the target duration was held constant (Millner et al., 2018).

Demographics. Participants completed a brief questionnaire with demographic and socioeconomic indicators (e.g., age, gender, ethnicity, and education).

Intelligence. Intelligence was measured with two subtests (Vocabulary and Matrix Reasoning) of the Wechsler Abbreviated Scale of Intelligence (2nd ed.; WASI) that are highly correlated with full-scale scores (Wechsler, 1999). Education was measured on an eight-point scale where one was no college courses and eight was completed graduate school.

Clinical Information

We obtained current psychiatric diagnoses from the medical record and data on several clinical and personality constructs to assess whether the groups were clinically similar and to test whether task effects were present above and beyond any clinical/personality differences. The online supplemental material includes detailed information on the scales described below as well as several additional measures that were included as part of the larger

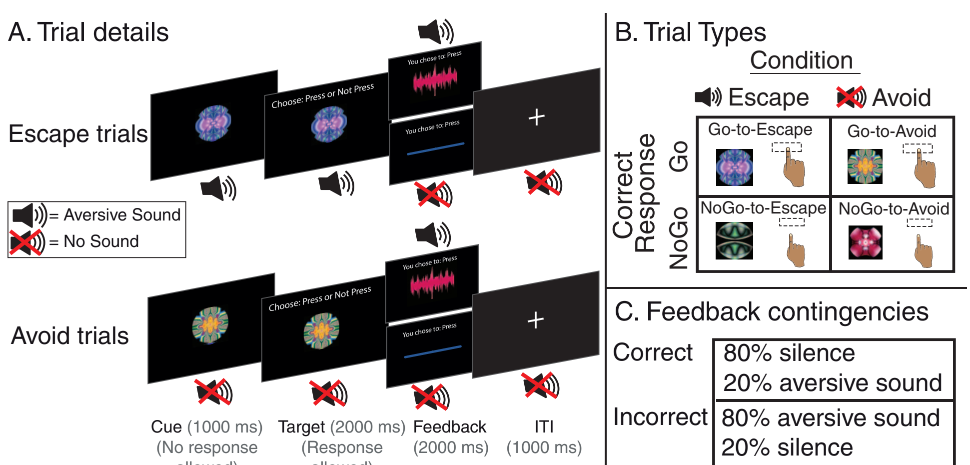

Figure 1. Experimental paradigm. (A) On each trial, participants were presented one of four fractal images and had to learn whether pressing a button (i.e., go) or withholding a button press (i.e., no-go) resulted in silence, rather than an aversive sound, during feedback. For 1 s after the image onset participants were unable to make a response (i.e., the cue). This was followed by the presentation of the text "Choose: Press or Not Press," which represented the onset of a 2-s target where participants could choose to go or no-go, followed by 2 s of feedback that consisted of an aversive sound (and a pink sound wave image) or silence (and a blue line image) as well as text reading either "You chose: Press" or "You chose: Not Press." Feedback was followed by a 1-s intertrial stimulus. On escape, but not avoid, trials the aversive sound played during the cue and target. (B) The four stimuli mapped on to a condition (i.e., escape or avoid) and a correct response (i.e., go or no-go) that more frequently led to silence during feedback. (C) Feedback was probabilistic such that a correct response resulted in silence 80% of the time and the aversive sound 20% of the time and vice versa for an incorrect response. See the online article for the color version of this figure.

study and were used in regularized regressions (described in the following text).

Self-Injurious Thoughts and Behaviors Interview (SITBI). The SITBI is a semistructured interview that assess a range of STB and as well as the recency, frequency, and other characteristics of each STB endorsed. We used the SITBI to assign participants to the suicidal or nonsuicidal groups.

Posttraumatic stress disorder (PTSD) symptoms. The PTSD Checklist for *DSM–5* (PCL-5) is a 20-item self-report measure that assesses the 20 *DSM–5* symptoms of PTSD (Weathers et al., 2013).

Depressive symptoms. The nine-item Patient Health Questionnaire (PHQ-9) is a self-report instrument that measures the frequency of the nine criteria for major depressive disorder in the last 2 weeks, each on a 0 (*not at all*) to 3 (*nearly every day*) scale (Kroenke, Spitzer, & Williams, 2001).

Hopelessness. The Beck Hopelessness Scale (BHS) is a measure of self-reported hopelessness that includes 20 true–false items with total scores ranging from 0 to 20 (Beck, Weissman, Lester, & Trexler, 1974).

Borderline personality disorder (BPD) traits. The McLean Screening Instrument for BPD (MSI-BPD) contains 10 true–false items assessing criteria of BPD, which were summed (Zanarini et al., 2003).

Beck Suicide Scale for Ideation (SSI; Beck, Steer, & Ranieri, 1988). The SSI is a 21-item self-report instrument that assesses several aspects of suicidal thoughts and desires in the last week. The SSI was used as a measure of suicide ideation severity to correlate computational parameters.

Data Analyses

Behavioral data. Behavioral data were analyzed using generalized linear mixed-effects regression (GLMER) models with the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). In our prior study, both choice and RT data supported a bias for active responses to escape an aversive stimulus (“active-escape bias”) and a similar bias for inhibitory responses to avoid an aversive stimulus (“inhibitory-avoid bias”). Thus, accuracy was higher for go-to-escape than no-go-to-escape but this pattern was reversed in Avoid, such that accuracy was higher for no-go-to-avoid than go-to-avoid, resulting in a significant Condition × Required Response interaction. For RT, relative to avoid, escape elicited faster RTs, leading to a main effect of condition (i.e., escape produced more vigorous responding than avoid). In the current study, we tested whether the Condition (escape, avoid) × Response (go, no-go) interaction for accuracy and the main effect of condition for RT differed between the groups (lifetime suicidal, psychiatric control). For accuracy, we conducted a logistic GLMER with the dependent variable as trial accuracy (i.e., 0 = incorrect choice; 1 = correct choice; correct choice defined as choosing the response (go/no-go) associated with a higher probability of silence during feedback). For RT, we conducted a gamma GLMER. For all GLMERs, within-subject factors were added as a random factor for the intercept and slopes for all fixed factors and interactions (i.e., maximal models (Barr, 2013). Fixed factors were condition, response, group and their interaction. We used deviation coding so that each fixed effect coefficient is at the unweighted mean of the other fixed effects rather than arbitrary reference levels, allowing coefficients to be interpreted as main effects (Menard, 2010; Politzer-Ahles, 2015). Coefficient confidence intervals and statistical significance, set at .05 with a two-tailed test, was determined using a Wald test to compute p values. To understand differences between groups, we followed significant Group × Condition × Response three-way interactions with two-way interaction contrasts in the R package phia, which also uses a Wald test to determine significance (Martınez, 2015). We used the R package simr to determine observed power in the mixed-effects models.

Computational model.

Model description. We used a computational model to obtain two parameters that captured the active-escape and inhibitory-avoid biases, respectively. We assessed the degree to which these biases were present in both choice and RT for each participant, allowing us then to compare groups on these biases. Our a priori hypothesis was that lifetime suicidal participants would show a higher active-escape bias relative to an active-avoid bias (i.e., active responses to avoid a punishment). The computational model used both a reinforcement learning (RL) model and drift-diffusion model (DDM).

Briefly, the RL model operationalized how on each trial the value of each response (go, no-go) to each cue was updated based on feedback from previous choices. These values were then used in a DDM to estimate choice and RT probabilities. Two-alternative DDMs have a decision variable that begins a trajectory from a starting point and progresses over time based on a drift rate parameter and Gaussian noise component until it reaches one of two decision boundaries, which represents a choice and therefore a response. The difference in value between the two response options is parameterized as the drift rate in the DDM. On the basis of prior work using the current task (see Figure 2; Millner et al., 2018), we parameterized the active-escape bias and the inhibitory-avoid bias (together referred to as *bias parameters*) by allowing the starting point of the DDM to vary by condition. Allowing the starting point to vary means the decision variable starts closer to a decision boundary, requiring less “evidence” of a higher (RL) value to reach that boundary (i.e., select that response). The no-go decision boundary was modeled similarly to the go decision boundary, except that the boundary was implicit. Prior research suggests that this is a valid assumption (Ratcliff, Huang-Pollock, & McKoon, 2018). We used the identical model described in Millner et al. (2018) and model details are supplied in the online supplemental material.

Model fitting. We used a hierarchical model fitting procedure, similar to prior studies (Guitart-Masip et al., 2012; Mkrtchian, Aylward, et al., 2017). This is an iterative process whereby (1) the model is fit to each participant’s data, then (2) each participant’s parameter estimates contribute to mean estimates for each parameter, and (3) these mean estimates are then used as priors for another round of model fitting to each participant’s data. This process continues until the log model evidence converges or after 50 iterations. We provide a formal explanation of the process in the online supplemental material.

Group differences on bias model parameter estimates. Hierarchical model fitting provides more stable estimation of group parameters, but it is unclear whether the two clinical groups should be considered separate populations, as hypothesized, and therefore be fit with separate group-level priors, or

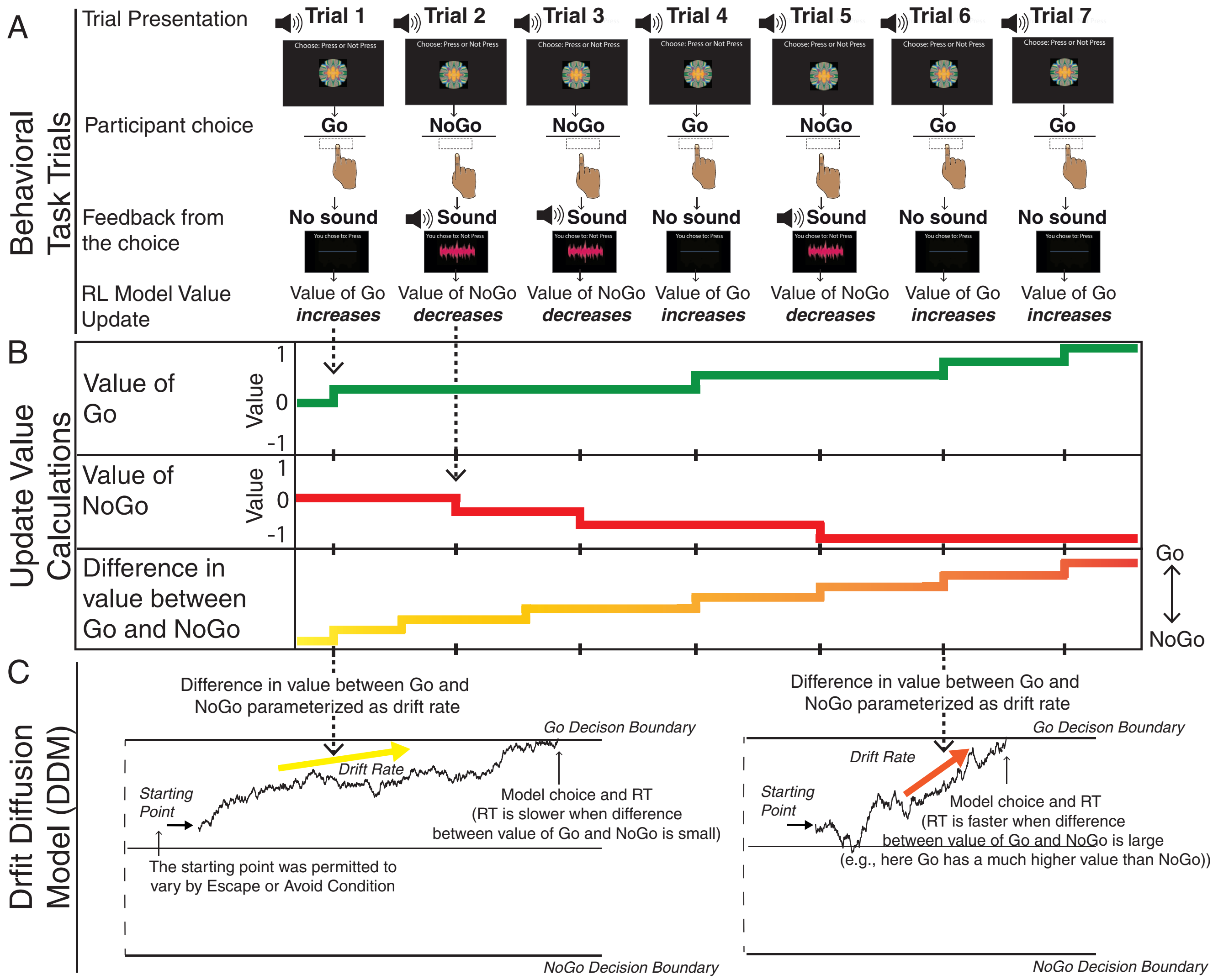

Figure 2. Schematic of RL-DDM model. (A) Example of escape-to-go trials, responses (go or no-go) and feedback. (B) When a response results in "no sound" feedback, the value associated with that response is increased, whereas value decreases when a response results in the aversive sound. (C) After the value of a response is updated on each trial, the value difference between the go and no-go is parameterized as the drift rate in the DDM. Early, when it is unclear which response is better, the value difference is smaller, resulting in a lower drift rate and longer response times (on the left). When the difference in value between the two responses is greater, the drift rate is higher and decision response times are faster (on the right). The plots show two DDM parameters, the starting point and drift rate. The starting point was the main bias parameter as it was the only parameter that was permitted to differ between avoid and escape conditions. See the online article for the color version of this figure.

one population, in which case all participants will contribute to and use the same group-level priors. Fitting the groups separately could pull parameter estimates toward different group-level priors and therefore promote group differences. On the other hand, a single group-level prior has the opposite effect, pushing all parameter estimates toward single group-level priors and therefore potentially underestimating group differences. Thus, each approach could bias subject-level parameter estimates. Similar to prior articles (Mkrtchian, Aylward, et al., 2017), we quantitatively compared which population model (one or two populations) provided higher model evidence using Bayesian model selection (BMS; Rigoux, Stephan, Friston, & Daunizeau, 2014). However, given that BMS results were not conclusive and both model-fitting approaches could be biased, we report the results from both approaches, providing effectively an upper and lower bound on the group differences. To compare groups on the bias parameters, our main parameters of interest, we conducted independent *t* tests of the active-escape bias and inhibitory-avoid bias parameters. We used the Python package statsmodels to calculate statistics and determine power.

As exploratory analyses, we conducted *t* tests on the other parameters as well.

Exploratory analyses to test for unique variance and correlations with bias parameters. We conducted exploratory analyses to test whether the active-escape and inhibitory-avoid bias parameters explained unique variance in suicidal group status after accounting for age only (the suicidal group was significantly younger than the nonsuicidal controls), education only (education was significantly higher among nonsuicidal controls), and after accounting for age and several clinical measures (on which the suicidal group scored higher). We first ran logistic regressions where group was entered as the dependent variable and bias parameter values and age were entered as the independent variables (separately for parameters fit with one and two populations). A similar but separate logistic regression we enter education instead of age as an independent variable. We then conducted similar regressions controlling for age, depression, hopelessness, PTSD symptoms, and borderline personality traits (entering age and education together resulted in larger coefficients for active-escape bias parameters that entering only age, suggesting there were suppression effects). Finally, we used a lasso regularized logistic regressions. By penalizing regression coefficients and setting many coefficients to zero, lasso regressions are often used to identify variables that account for relevant variance among a large number of predictor variables. We sought to test whether the bias parameters, particularly the active-escape bias, would remain a relevant predictor of group status among 27 clinical, personality and demographic constructs collected in the study. To do this, we entered the bias parameters as well as the majority of self-report scales collected during the study, including scales on free will and determinism, optimism, impulsiveness, emotion regulation, and a measure of risk and resilience from military deployment experiences, such as combat exposures (see the online supplemental material) into a lasso regularized logistic regression. The sparsity parameter, which influences the penalty on the regression coefficients, was selected through *k*-fold stratified cross-validation across the entire sample. We used the Python package scikit-learn (Pedregosa et al., 2011).

Finally, to assess what clinically relevant constructs correlated with the latent active-escape bias computational constructs and how the clinical constructs might help explain variation across the groups, we conducted exploratory bivariate correlations between the bias parameters and all clinical scales. We correlated 27 scales with the computational parameters and none survived corrections for multiple comparisons. Therefore, we report significant correlations uncorrected for multiple comparisons.

Results

Behavioral Data

Overall effects. Participants showed an overall accuracy of 73.4% (SD = 16.6; condition M range = 83.0%–63.8%) on the task, suggesting they learned the required responses. Continuing to look across groups, we replicated prior accuracy results from Millner et al. (2018) showing a significant Condition × Response interaction (b = 1.7 [2.1–1.2], $p < .001$), where during escape, accuracy was higher when go, rather than no-go, was the required response ($p < .001$). During avoid, trials where no-go, compared with go, was the required response showed nonsignificantly higher accuracy ($p = .09$). These decision-making biases also purportedly affect the vigor with which choices are made and should be revealed by faster RT when biases promote go choices. Consistent with this, the current RT results across all participants also replicated prior results showing a main effect of condition ($b = -0.08$ [-0.08–0.08], $p < .001$), such that escape had faster RTs than avoid ($p < .001$; for trials where no-go was the required response, RT was collected when participants erroneously chose go).

Group effects of accuracy. Both groups showed the expected Group × Condition interaction where escape was associated with higher accuracy for go, compared with no-go, but avoid was associated with higher accuracy for no-go, compared with go (see Figure 3). Additionally, these interactions differed among the groups (i.e., Condition × Response × Group interaction (b = 1.1 [0.2–2.1], $p = .01$, observed power = .80) with a larger magnitude interaction for suicidal participants ($b = 2.1$, $p < .001$) than psychiatric controls ($b = 0.9$, $p = .02$). To understand whether the larger interaction among suicidal participants was driven by escape or avoid conditions, we looked at Response × Group interactions within escape and avoid conditions separately. These interactions were not significant (escape $b = -0.6$, $p = .19$; avoid $b = 0.5$, $p = .19$), suggesting that, although both groups showed biases as evinced by significant Condition × Response interactions, the three-way interaction was due to both stronger active-escape and inhibitory-avoid biases on accuracy for lifetime suicidal participants than psychiatric controls.

Group effects of RT. As mentioned, effects of the biases on RT are reflected by a main effect of Condition because the bias should increase vigor (as indexed by faster RT) when promoting go responses during escape and reduce vigor when promoting inhibitory responses during avoid. Consistent with the idea that the bias differed between the groups, we found a Condition × Group (b = 0.06 [0.05–0.06], $p < .001$, observed power = .99) interaction, such that suicidal participants had significantly faster RT during the escape condition ($b = 0.03$, $p < .001$) and significantly slower RT during avoid ($b = -0.03$, $p < .001$; Figure 3).[1] Like accuracy, this result again suggests that there were stronger active-escape and inhibitory-avoid effects on RT for lifetime suicidal participants than psychiatric controls.

There also was a main effect of response for RT because RTs during no-go trials (i.e., when no-go is correct) are inherently errors and therefore have much slower RT than go trials Avoid ($b = 0.13$, $p < .001$). Interestingly, there also was a significant Condition × Response × Group interaction ($b = -0.01$ [-0.1–0.0], $p < .001$). Here, both groups showed significant Condition × Response interactions where the difference between escape and avoid was larger for go than for no-go trials, but this interaction effect was larger among suicidal participants ($b = -0.02$, $p < .001$) compared with nonsuicidal participants ($b = -0.01$, $p < .001$).

Atypical behavioral performance. Several participants had extremely low accuracy within a given trial type (see Figure 3), whereas others had very high accuracy across all four trial types. We ran multiple analyses removing data from people with atypical

[1] Power based on traditional analysis of variance with wide format data because simr returned errors.

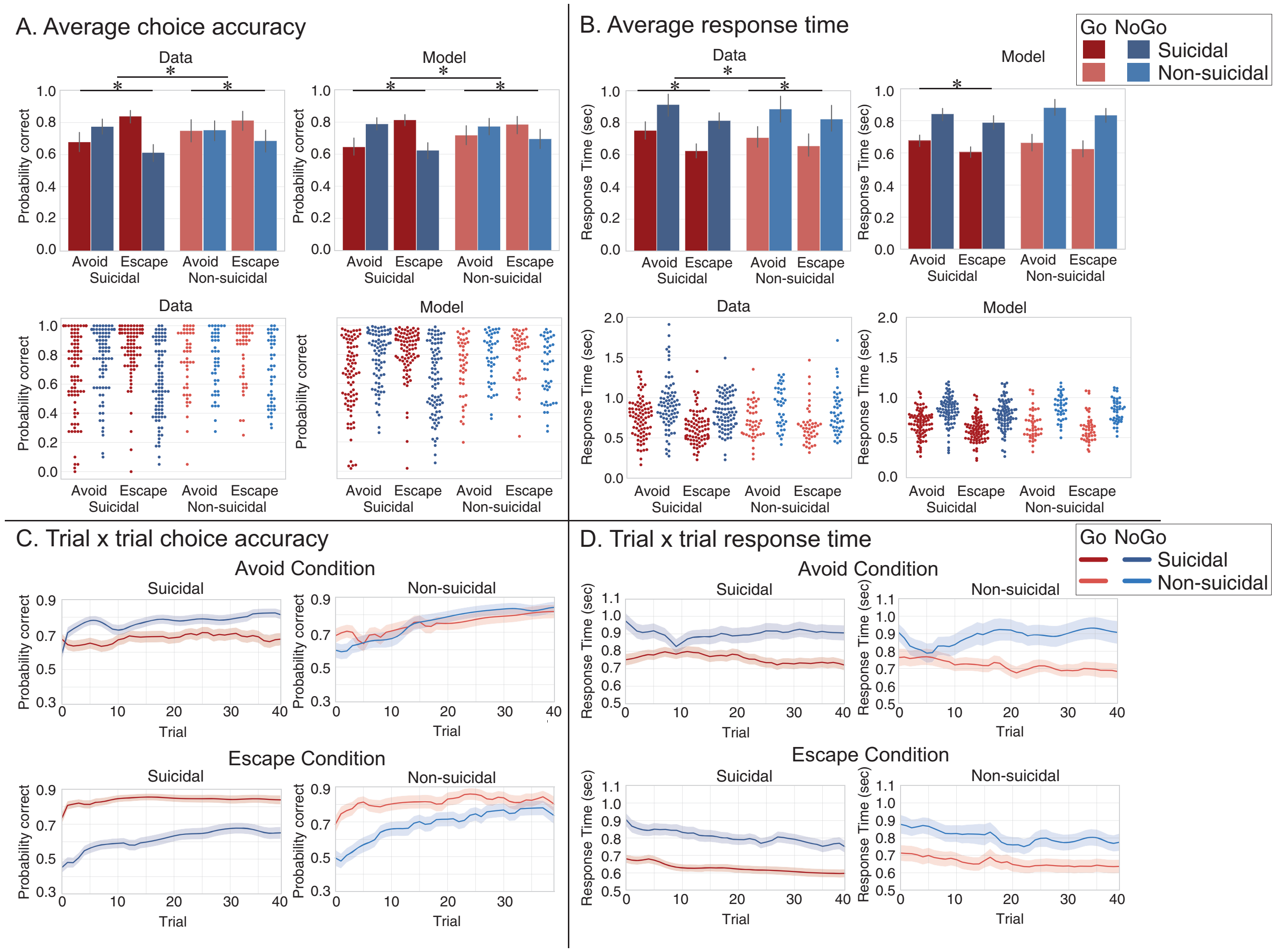

Figure 3. Accuracy and response time (RT) results for each group and each participant from the empirical data and the model. (A) For accuracy, both groups showed the expected Condition × Response interaction where accuracy was higher for go than no-go during escape but not during avoid. The magnitude of the interaction was larger for suicidal participants, leading to a three-way interaction. The model recapitulates these results well. (B) No-go trials with RT are inherently errors and slower than go trials. Both groups showed slower RT for avoid versus escape trials but this effect was larger among lifetime suicidal participants. Qualitatively, the model captures both the within- and between-group patterns among the RT results with one exception: in the empirical data, compared with nonsuicidal group, the suicidal group showed slower RT for avoid no-go trials whereas the model produced slower RT among the nonsuicidal group for this condition. (C) Proportion correct. (D) Average RT for each trial with smoothing based on robust spline smoothing (Garcia, 2010). Error bars represent standard error of the mean. * $p < .05$. See the online article for the color version of this figure.

performance to ensure they did not cause confounds. Results were largely consistent with those from the full sample (see the online supplemental material for more information).

Computational Model

Model description. Choice and RT data provided some evidence in support of stronger biases for lifetime suicidal participants compared with psychiatric controls, but these results could not be attributed specifically to either avoid or escape conditions, as neither condition alone was significantly different between the groups. This could be due, in part, to a power issue. Given that we hypothesized that a single mechanism affects both choice and RT, we used a computational model to obtain latent bias parameters that captures group differences in condition effects across RT and choice, thus enhancing the overall power of the analysis.

Model fitting. BMS revealed a small advantage for fitting a single population distribution versus two separate population distributions for each group (protected exceedance probabilities: single population = 0.52, two populations = 0.48). Thus, fitting the model as if all participants were drawn from a single population was slightly favored over fitting the model as if the participants in the two groups were drawn from two different populations.

Group differences on bias model parameter estimates. Group differences for all model parameters are shown in Table 2.

Table 2
Group Differences on Model Parameters

Group difference statistics	T	b0	b1	w1	w2	α	ω
Model fit with one population							
Suicidal *M* (*SD*)	.02 (.15)	.86 (.30)	2.22 (1.31)	.32 (.1)	.21 (.09)	.16 (.11)	1.84 (.32)
Nonsuicidal *M* (*SD*)	.01 (.15)	.86 (.36)	2.6 (1.37)	.28 (.1)	.23 (.08)	.14 (.1)	1.9 (.28)
t	.13	.10	−1.52	2.22	−1.12	1.11	−1.07
p	.90	.92	.13	.03	.27	.27	.29
Cohen's *d*	.02	.02	−.28	.41	−.21	.21	−.20
95% CI	[−.34−.39]	[−.38−.4]	[−.66−.1]	[.03−.78]	[−.56−.14]	[−.15−.56]	[−.54−.14]
Observed power	.05	.05	.32	.59	.2	.2	.19
Model fit with two populations							
Suicidal *M* (*SD*)	.02 (.15)	.84 (.27)	1.99 (1.31)	.33 (.1)	.21 (.09)	.23 (.15)	1.83 (.31)
Nonsuicidal *M* (*SD*)	.02 (.16)	.85 (.31)	2.73 (1.39)	.28 (.08)	.23 (.06)	.12 (.08)	1.89 (.28)
t	.11	−.14	−2.99	3.14	−.76	4.64	−1.04
p	.91	.89	.003	.002	.45	>.001	.30
Cohen's *d*	.02	−.02	−.55	.58	−.15	.86	−.20
95% CI	[−.36−.39]	[−.41−.37]	[−.94−.16]	[.2−.94]	[−.47−.19]	[.55−1.16]	[−.54−.16]
Observed power	.05	.05	.84	.89	.13	.99	.18

Note. The parameters w1 and w2 represent condition biases (i.e., escape and avoid) and were the main parameters of interest. Suicidal = lifetime suicidal thoughts and behaviors; nonsuicidal = psychiatric history and no history of STB; T = nondecision time; b0 = constant go bias; b1 = shared go bias; w1 = starting point for escape trials; w2 = starting point for avoid trials; α = learning rate; ω = boundary separation; CI = Cohen's *d* confidence interval.

The suicidal group showed a significantly increased active-escape bias (w1) parameter compared with the nonsuicidal group (w1; see Table 2 and Figure 4), but the groups were similar on the inhibitory-avoid bias parameter (w2; see Table 2 and Figure 4). To test whether suicidal subgroups drove the effect, we compared bias parameters for (1) those with last-year versus lifetime (but not last year) STB and (2) attempters versus ideators. Neither showed comparison significant differences (see the online supplemental material). When fitting the model with a single population distribution, no other parameter differed between suicidal and nonsuicidal groups whereas when fitting with two population distributions (see Table 2), the general go bias was higher among nonsuicidal participants (see Table 2) and the learning rate α of the reinforcement learning model was significantly higher among suicidal participants (see Table 2). This higher general go bias suggests that nonsuicidal participants had a higher tendency to make a go response across both escape and avoid trials. Within a RL model, a higher learning rate suggests that when making a decision, suicidal participants give more weight to recent outcomes compared with nonsuicidal participants.

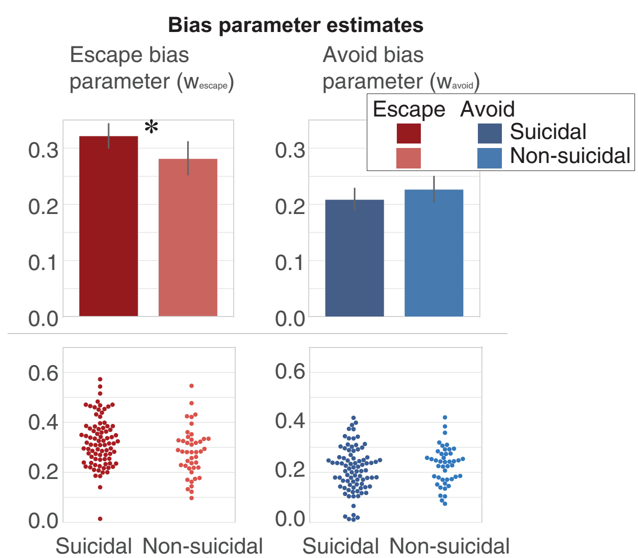

Figure 4. Bias parameter estimates. Lifetime suicidal participants showed significantly higher escape bias parameter estimate than nonsuicidal controls. The groups were comparable on the avoid bias parameter. * $p < .05$. See the online article for the color version of this figure.

Exploratory analyses to test for unique variance and correlations with bias parameters. As expected, bivariate logistic regressions with the inhibitory-avoid bias parameter predicting group status were not significant (single population: b = –2.4 [−6.7–1.8], p = .25; two populations: b = –1.7 [−6.1–2.7], p = .44). For active-escape, logistic regressions showed that the association between group status and the active-escape bias parameter remained significant when controlling for age (single population: b = 4.5 [0.3–8.7], p = .04; two populations: b = 6.9 [2.3–11.5], p = .004) or education (single population: b = 4.6 [0.6–8.7], p = .03; two populations: b = 7.0 [2.5–11.5], p = .002) and marginally significant when controlling for age, depression, hopelessness, PTSD symptoms, and symptoms of borderline personality disorder (single population: b = 4.4 [−0.4–9.1], p = .07; two populations: b = 7.2 [2.0–12.4], p = .007), suggesting that unique group difference variance was accounted for by the active-escape bias parameter, independent of the clinical measures. Regularized logistic regressions, which included 30 measures, retained the active-escape bias parameter (single population: ninth highest coefficient out of 15 variables retained; two populations: second highest coefficient out of 13 variables retained). However, the inhibitory-avoid bias parameter was not retained among the final coefficients. Thus, in a high dimensional variable space predicting group differences, the active-escape bias parameter was selected as a relevant variable. The highest coefficient in the lasso logistic regressions was borderline personality symptoms.

Finally, we assessed uncorrected correlations to understand which constructs were related with the bias parameters. The only variable significantly related to the active-escape bias parameter across both fitting with single and multiple populations was borderline personality symptoms (r = .20–.22, p = .02–.01), whereas the inhibitory-avoid bias parameter was significantly correlated with two aspects of deployment risk: combat experiences (r = .21, p = .02 across both fitting procedures) and consequences of combat (r = .23, p = .01 across both fitting procedures). Given that borderline personality disorder has been associated with self-injurious behaviors intended to escape aversive emotions (Zanarini et al., 2003) and avoidance behaviors are inherent in PTSD (i.e., necessary to meet criteria for; Mitchell et al., 2017), the bias parameters were associated with plausible clinical constructs. Finally, the active-escape parameter was not significantly related to suicidal ideation severity among those with current ideation (r = .13, p = .39).

Discussion

Clinical and theoretical accounts suggest that suicide represents an effort to escape aversive psychological states (Baumeister, 1990; Robins, 1981; Shneidman, 1998). However, there has been surprisingly little research examining decision making related to escaping aversive states. In this study, we used a reinforcement learning and decision-making framework to examine the influence of biases when selecting a response to either escape or avoid aversive states, among lifetime suicidal participants and nonsuicidal psychiatric controls. We found that both choice and RT suggested the presence of an increased bias for lifetime suicidal participants, but that this increased bias could not be attributed uniquely to either escape or avoid contexts. However, a computational modeling approach allowed us to capture processes that influenced both choice and RT data in a single model, and thus estimate these biases more precisely. Here we found that a bias for active responding when escaping an aversive state was increased among suicidal participants, compared with a non-STB psychiatric comparison group. In contrast, the parameter representing inhibitory responses when avoiding an aversive state was comparable among the groups. Further tests revealed that the active-escape bias continued to be associated with STB even after controlling for several psychiatric (e.g., depression, hopelessness, PTSD symptoms, and borderline personality traits) and demographic variables and was retained in a regularized logistic regression predicting group status. This latter analysis identifies the relevant variables in high-dimensional data, and suggests that, among an array of 27 clinical self-report measures and demographic variables, the active-escape bias provided nonredundant information in predicting which participants were suicidal. Finally, in exploratory analyses, this active-escape bias correlated with a continuous measure of borderline personality disorder symptoms, a clinical measure highly relevant to STB (Yen et al., 2004). Overall, this study suggests the computationally defined aberrant escape decision making measured here is associated with STB above and beyond just psychopathology or clinical severity.

There are multiple ways in which increased active-escape bias might be related to STB. Assuming that the active-escape bias applies to internal psychological processes as well as observable behaviors, one possibility is that an increased active-escape bias may play a causal role, pushing suicidal people to fairly automatically “do something” to escape aversive emotions, including fantasizing about suicide or, potentially, acting on such thoughts. That is, an increased active-escape bias may cause some people to have difficulty staying in an aversive state and instead, they (perhaps unconsciously) move to suicidal thoughts as the best way to gain relief. Perhaps suicidal thoughts are more likely if other options for relief are not perceived as viable options (e.g., the perception that depression symptoms will not subside). Alternatively, the increased escape bias among suicidal participants may reflect a type of conditioning caused by repeatedly engaging in STB to escape aversive emotional states. That is, the bias to escape aversive states is stronger among suicidal participants due to their repeated practice engaging in suicidal thoughts and behaviors to escape negative affect (Kleiman et al., 2018). A third possibility is that an active-escape bias measured here may be a stable vulnerability that requires other state-level stressors or factors to produce STB. This is supported by the fact that those with lifetime, but not last year, STB showed an active-escape bias on par with those with last year STB. Last, increased active-escape bias may be associated with STB but not directly related to their presence. For example, increased active-escape bias could be associated with just a higher general affective response that is also related to STB.

By analyzing behavior with a computational model, we were able to detect a bias that was not otherwise evident in analyses using typical behavioral indices and which was uniquely related to being suicidal, above having a psychiatric condition and after accounting for several clinical and demographic covariates. Furthermore, using an RL–DDM model allowed us to analyze both choice and RT simultaneously, increasing our statistical power and accounting for a single bias that affected both behavioral indices. Computational approaches provide the possibility of identifying computationally defined behaviors that are basic enough to be tied to specific brain circuits but can also be associated with clinical outcomes, thus bridging an explanatory gap that currently exists between brain function and clinical phenomena (Montague, Dolan, Friston, & Dayan, 2012). For example, studies have found a range of anatomical and functional abnormalities in serotonergic circuits (e.g., dorsal raphe nucleus [DRN] and prefrontal cortex [PFC] associated with STB and suicide death (Oquendo et al., 2014; van Heeringen & Mann, 2014). Interestingly, animal studies have implicated similar circuits in active escape behaviors. For instance, studies using optogenetics to activate neurons connecting the PFC and DRN have experimentally demonstrated that these pathways are critically involved in actions to escape challenging circumstances (Ren et al., 2018; Warden et al., 2012). Thus, this line of research could connect anatomical or functional abnormalities with computationally defined active-escape behaviors that, in turn, are related to increased suicide risk. This improved understanding could then lead to important clinical benefits if the implicated circuits could be targeted and modulated.

The behavioral data, but not the modeling results, suggested that suicidal participants might have a stronger bias for passive responses to avoid an impending aversive outcome. The different conclusions are due to either a shortcoming in the model or ambiguity in the data. Collecting more data or a different paradigm that produces larger avoidance effects might help resolve this inconsistency. If suicidal people do have a stronger bias to avoid impending aversive stimuli by responding passively, the bias would result in more failures to avoid the aversive stimulus, leading to more interactions with the aversive stimulus and therefore more instances that require escape. This result potentially has clinical relevance as encounters with more negative states may be important for the development and/or maintenance of psychopathology or STB.

Beyond STB, aberrant escape decision making may be relevant for understanding other psychiatric conditions and clinically relevant constructs (and may explain their association with STB). For example, substance seeking in postaddiction substance dependence is motivated by escaping negative affect associated with withdrawal (Baker, Piper, McCarthy, Majeskie, & Fiore, 2004) and obsessive–compulsive rituals provide escape from increased anxiety because of obsessions (Brandt et al., 2018). We also found that the active-escape bias was correlated with borderline personality disorder symptoms, which is associated with suicidal and nonsuicidal forms of self-injurious behaviors (Zanarini et al., 2003) to escape aversive emotions (Chapman, Specht, & Cellucci, 2005). In terms of a clinically related construct, distress intolerance (i.e., the degree to which people report not being able to handle or bear distressing feelings) might be related to escape and has been previously associated with STB (Anestis, Pennings, Lavender, Tull, & Gratz, 2013). For example, higher distress intolerance may be associated with a stronger bias for active behaviors to escape an aversive state, even when those behaviors are maladaptive. Although distress intolerance was not collected here, future research in this area could lead to a better understanding of the association between distress intolerance and STB.

This study was limited by several factors. First, like all clinical, observational studies, we lacked experimental control over group assignment which has implications for understanding the association between STB and active-escape biases. For example, group differences on clinical, demographic or some unmeasured variables may have led to group differences on active-escape bias rather than the presence of lifetime STB, although the bias was still related to STB after statistically controlling for covariates. Second, although computational models improve the measurement of behaviors by providing precise, quantifiable constructs, understanding the role of such computationally defined constructs in psychiatric conditions still relies on group assignment and/or clinically relevant outcomes, which are assessed with self-report. Thus, although we agree with several commentaries (e.g., Adams, Huys, & Roiser, 2015; Montague et al., 2012) that computational models can aid in understanding of psychiatric conditions, the increased precision afforded by these techniques may be undermined by the lack of precision characterizing clinically relevant outcomes. This limitation, along with the lack of experimental control, hinders our understanding of the precise role of the active-escape bias in STB. Third, this sample was among veterans and it is unclear whether these results will extend to the general population. Fourth, the unpleasant sound, which was the main negative stimulus of the current paradigm, was intended to produce an aversive state that approximates aversive psychological experiences; however, we did not collect ratings of the aversiveness for the sound and do not know the extent to which it was analogous of aversive psychological experiences. Fifth, an important question is whether attempters show greater active-escape bias compared with ideators. The current sample had a small number of attempters with last year attempts ($n = 26$) and substantial proportion of attempters with their last attempt 10+ years prior ($n = 10$), which was concerning given that studies have found larger effects the more recent the STB (Glenn et al., 2017). We compared (all) attempters and ideators but future studies should test a larger sample of attempters with more recent attempts. Sixth, we measured a limited set of variables and therefore were unable to test for relationships with many constructs previously associated with STBs (e.g., childhood adversity).

In conclusion, we examined decision-making processes associated with a longstanding view that STB are efforts to escape aversive emotional states. We used a reinforcement learning and computational modeling framework that captured choice and RT effects to assess a latent/hidden bias for active responses to escape that would otherwise be undetectable. We found that this bias was larger for people with a history of STB compared with those with a psychiatric condition but no history of STB. Furthermore, we applied several controls for demographic and clinical covariates and the elevated active-escape bias among suicidal participants persisted. This line of research may help provide insight into how neural processes and environmental stressors influence people to consider and select suicide as an option to escape difficult emotions and life circumstances.

References

Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2015). Computational Psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry*. Advance online publication. http://dx.doi.org/10.1136/jnnp-2015-310737

Anestis, M. D., Pennings, S. M., Lavender, J. M., Tull, M. T., & Gratz, K. L. (2013). Low distress tolerance as an indirect risk factor for suicidal behavior: Considering the explanatory role of non-suicidal self-injury. *Comprehensive Psychiatry, 54,* 996–1002. http://dx.doi.org/10.1016/j.comppsych.2013.04.005

Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R., & Fiore, M. C. (2004). Addiction motivation reformulated: An affective processing model of negative reinforcement. *Psychological Review, 111,* 33–51. http://dx.doi.org/10.1037/0033-295X.111.1.33

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*. Advance online publication. http://dx.doi.org/10.3389/fpsyg.2013.00328

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. Advance online publication. http://dx.doi.org/10.18637/jss.v067.i01

Baumeister, R. F. (1990). Suicide as escape from self. *Psychological Review, 97,* 90–113. http://dx.doi.org/10.1037/0033-295X.97.1.90

Beck, A. T., Steer, R. A., & Ranieri, W. F. (1988). Scale for Suicide Ideation: Psychometric properties of a self-report version. *Journal of Clinical Psychology, 44,* 499–505. http://dx.doi.org/10.1002/1097-4679(198807)44:4<499::AID-JCLP2270440404>3.0.CO;2-6

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The hopelessness scale. *Journal of Consulting and Clinical Psychology, 42,* 861–865. http://dx.doi.org/10.1037/h0037562

Bossarte, R., Claassen, C. A., & Knox, K. (2010). Veteran suicide prevention: Emerging priorities and opportunities for intervention. *Military Medicine, 175,* 461–462. http://dx.doi.org/10.7205/MILMED-D-10-00050

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,* 1379–1391. http://dx.doi.org/10.1037/pas0000254

Brandt, V. C., Hermanns, J., Beck, C., Bäumer, T., Zurowski, B., & Münchau, A. (2018). The temporal relationship between premonitory urges and covert compulsions in patients with obsessive-compulsive disorder. *Psychiatry Research, 262,* 6–12. http://dx.doi.org/10.1016/j.psychres.2018.01.041

Callahan, C. M., Unverzagt, F. W., Hui, S. L., Perkins, A. J., & Hendrie, H. C. (2002). Six-item screener to identify cognitive impairment among potential subjects for clinical research. *Medical Care, 40,* 771–781. http://dx.doi.org/10.1097/00005650-200209000-00007

Carter, S. B., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R., & Wright, G. (2006). *Historical statistics of the United States: Millennial edition*. (Vol. 3). New York, NY: Cambridge University Press.

Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. (2018). *Web-based Injury Statistics Query and Reporting System (WISQARS)*. Retrieved from www.cdc.gov/injury/wisqars

Chapman, A. L., Specht, M. W., & Cellucci, T. (2005). Borderline personality disorder and deliberate self-harm: Does experiential avoidance play a role? *Suicide & Life-Threatening Behavior, 35,* 388–399. http://dx.doi.org/10.1521/suli.2005.35.4.388

Dennis, J. P., & Vander Wal, J. S. (2010). The Cognitive Flexibility Inventory: Instrument development and estimates of reliability and validity. *Cognitive Therapy and Research, 34,* 241–253. http://dx.doi.org/10.1007/s10608-009-9276-4

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., . . . Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin, 143,* 187–232. http://dx.doi.org/10.1037/bul0000084

Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis, 54,* 1167–1178. http://dx.doi.org/10.1016/j.csda.2009.09.020

Gibbons, R. D., Hur, K., & Mann, J. J. (2017). Suicide rates and the declining psychiatric hospital bed capacity in the United States. *Journal of the American Medical Association Psychiatry, 74,* 849–850. http://dx.doi.org/10.1001/jamapsychiatry.2017.1227

Glaesmer, H., Rief, W., Martin, A., Mewes, R., Brähler, E., Zenger, M., & Hinz, A. (2012). Psychometric properties and population-based norms of the Life Orientation Test Revised (LOT-R). *British Journal of Health Psychology, 17,* 432–445. http://dx.doi.org/10.1111/j.2044-8287.2011.02046.x

Glenn, J. J., Werntz, A. J., Slama, S. J. K., Steinman, S. A., Teachman, B. A., & Nock, M. K. (2017). Suicide and self-injury-related implicit cognition: A large-scale examination and replication. *Journal of Abnormal Psychology, 126,* 199–211. http://dx.doi.org/10.1037/abn0000230

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85,* 348–362. http://dx.doi.org/10.1037/0022-3514.85.2.348

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage, 62,* 154–166. http://dx.doi.org/10.1016/j.neuroimage.2012.04.024

Hoge, C. W., Riviere, L. A., Wilk, J. E., Herrell, R. K., & Weathers, F. W. (2014). The prevalence of post-traumatic stress disorder (PTSD) in U.S. combat soldiers: A head-to-head comparison of *DSM–5* versus *DSM–IV–TR* symptom criteria with the PTSD checklist. *The Lancet Psychiatry, 1,* 269–277. http://dx.doi.org/10.1016/S2215-0366(14)70235-4

Huys, Q. J. M., Gölzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., & Dolan, R. J. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological Medicine, 46,* 1027–1035. http://dx.doi.org/10.1017/S0033291715002597

Kleiman, E. M., Coppersmith, D. D. L., Millner, A. J., Franz, P. J., Fox, K. R., & Nock, M. K. (2018). Are suicidal thoughts reinforcing? A preliminary real-time monitoring study on the potential affect regulation function of suicidal thinking. *Journal of Affective Disorders, 232,* 122–126. http://dx.doi.org/10.1016/j.jad.2018.02.033

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16,* 606–613. http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x

Lyon, J. (2017). New data on suicide risk among military veterans. *Journal of the American Medical Association, 318,* 1531–1531.

MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. New York, NY: Cambridge University Press.

Martınez, H. D. R. (2015). *Analyzing interactions of fitted models*. Retrieved from https://cran.r-project.org/web/packages/phia/vignettes/phia.pdf

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. Atlanta, GA: SAGE.

Millner, A. J., Gershman, S. J., Nock, M. K., & den Ouden, H. E. M. (2018). Pavlovian control of escape and avoidance. *Journal of Cognitive Neuroscience, 30,* 1379–1390. http://dx.doi.org/10.1162/jocn_a_01224

Mitchell, K. S., Wolf, E. J., Bovin, M. J., Lee, L. O., Green, J. D., Rosen, R. C., . . . Marx, B. P. (2017). Network models of *DSM–5* posttraumatic stress disorder: Implications for ICD-11. *Journal of Abnormal Psychology, 126,* 355–366. http://dx.doi.org/10.1037/abn0000252

Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling avoidance in mood and anxiety disorders using reinforcement learning. *Biological Psychiatry, 82,* 532–539. http://dx.doi.org/10.1016/j.biopsych.2017.01.017

Mkrtchian, A., Roiser, J. P., & Robinson, O. J. (2017). Threat of shock and aversive inhibition: Induced anxiety modulates Pavlovian-instrumental interactions. *Journal of Experimental Psychology: General, 146,* 1694–1704. http://dx.doi.org/10.1037/xge0000363

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences, 16,* 72–80. http://dx.doi.org/10.1016/j.tics.2011.11.018

Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., . . . the GBD 2016 Causes of Death Collaborators. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet, 390,* 1151–1210. http://dx.doi.org/10.1016/S0140-6736(17)32152-9

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in Wiener diffusion models. *Journal of Mathematical Psychology, 53,* 222–230. http://dx.doi.org/10.1016/j.jmp.2009.02.003

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-Injurious Thoughts and Behaviors Interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment, 19,* 309–317. http://dx.doi.org/10.1037/1040-3590.19.3.309

Oquendo, M. A., Sullivan, G. M., Sudol, K., Baca-Garcia, E., Stanley, B. H., Sublette, M. E., & Mann, J. J. (2014). Toward a biosignature for suicide. *The American Journal of Psychiatry, 171,* 1259–1277. http://dx.doi.org/10.1176/appi.ajp.2014.14020194

Ousdal, O. T., Huys, Q. J., Milde, A. M., Craven, A. R., Ersland, L., Endestad, T., . . . Dolan, R. J. (2017). The impact of traumatic stress on Pavlovian biases. *Psychological Medicine*. Advance online publication. http://dx.doi.org/10.1017/S003329171700174X

Paulhus, D. L., & Carey, J. M. (2011). The FAD-Plus: Measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment, 93,* 96–104. http://dx.doi.org/10.1080/00223891.2010.528483

Pearlin, L. I., & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior, 19,* 2–21. http://dx.doi.org/10.2307/2136319

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12,* 2825–2830.

Politzer-Ahles, S. (2015). *Coding schemes for categorical variables in regression*. Retrieved from http://www.mypolyuweb.hk/~sjpolit/coding_schemes.html

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2018). Modeling individual differences in the go/no-go task with a diffusion model. *Decision, 5,* 42–62. http://dx.doi.org/10.1037/dec0000065

Ren, J., Friedmann, D., Xiong, J., Liu, C. D., DeLoach, K. E., & Ran, C. (2018). *Anatomical, physiological, and functional heterogeneity of the dorsal raphe serotonin system*. http://dx.doi.org/10.1101/257378

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory, 2,* 64–99.

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies—Revisited. *NeuroImage, 84,* 971–985. http://dx.doi.org/10.1016/j.neuroimage.2013.08.065

Ringer, F. B., Soberay, K. A., Rogers, M. L., Hagan, C. R., Chu, C., Schneider, M., . . . Joiner, T. E. (2018). Initial validation of brief measures of suicide risk factors: Common data elements used by the Military Suicide Research Consortium. *Psychological Assessment, 30,* 767–778. http://dx.doi.org/10.1037/pas0000519

Robins, E. (1981). *The final months: A study of the lives of 134 persons who committed suicide*. New York, NY: Oxford University Press.

Shneidman, E. S. (1998). Perspectives on suicidology. Further reflections on suicide and psychache. *Suicide & Life-Threatening Behavior, 28,* 245–250. http://dx.doi.org/10.1521/suli.2007.37.3.245

van Heeringen, K., & Mann, J. J. (2014). The neurobiology of suicide. *The Lancet Psychiatry, 1,* 63–72. http://dx.doi.org/10.1016/S2215-0366(14)70220-2

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,* 710–717. http://dx.doi.org/10.1002/jts.21868

Warden, M. R., Selimbeyoglu, A., Mirzabekov, J. J., Lo, M., Thompson, K. R., Kim, S.-Y., . . . Deisseroth, K. (2012). A prefrontal cortex-brainstem neuronal projection that controls response to behavioural challenge. *Nature, 492,* 428–432. http://dx.doi.org/10.1038/nature11617

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD Checklist for DSM–5 (PCL-5)*. Retrieved from www.ptsd.va.gov

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Westreich, D. (2012). Berkson's bias, selection bias, and missing data. *Epidemiology, 23,* 159–164. http://dx.doi.org/10.1097/EDE.0b013e31823b6296

Whiteside, S. P., & Lynam, D. R. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30,* 669–689. http://dx.doi.org/10.1016/S0191-8869(00)00064-7

Yen, S., Shea, M. T., Sanislow, C. A., Grilo, C. M., Skodol, A. E., Gunderson, J. G., . . . Morey, L. C. (2004). Borderline personality disorder criteria associated with prospectively observed suicidal behavior. *The American Journal of Psychiatry, 161,* 1296–1298. http://dx.doi.org/10.1176/appi.ajp.161.7.1296

Zanarini, M. C., Frankenburg, F. R., Hennen, J., & Silk, K. R. (2003). The longitudinal course of borderline psychopathology: 6-year prospective follow-up of the phenomenology of borderline personality disorder. *The American Journal of Psychiatry, 160,* 274–283. http://dx.doi.org/10.1176/appi.ajp.160.2.274

Zanarini, M. C., Vujanovic, A. A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., & Hennen, J. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): A continuous measure of *DSM–IV* borderline psychopathology. *Journal of Personality Disorders, 17,* 233–242. http://dx.doi.org/10.1521/pedi.17.3.233.22147

Received June 19, 2018
Revision received September 24, 2018
Accepted October 1, 2018 ■