

# Suicide and Self-Injury-Related Implicit Cognition: A Large-Scale Examination and Replication

Jeffrey J. Glenn and Alexandra J. Wertz  
University of Virginia

S. J. Katarina Slama  
University of California, Berkeley

Shari A. Steinman  
West Virginia University

Bethany A. Teachman  
University of Virginia

Matthew K. Nock  
Harvard University

Suicide and self-injury are difficult to predict because at-risk individuals are often unable or unwilling to report their intentions. Therefore, tools to reliably assess risk without reliance on self-report are critically needed. Prior research suggests that people who engage in suicidal and nonsuicidal self-injury (NSSI) often implicitly (i.e., outside conscious control) associate themselves with self-harm and death, indicating that self-harm-related implicit cognition may serve as a useful behavioral marker for suicide risk. However, earlier studies left several critical questions about the robustness, sensitivity, and specificity of self-harm-related implicit associations unaddressed. We recruited a large sample of participants ( $N = 7,015$ ) via a public web-based platform called Project Implicit Mental Health (PIMH) to test several hypotheses about self-harm-related implicit associations using the Implicit Association Test (IAT). Participants were randomly assigned to complete 1 of 3 self-harm IATs (Self + Cutting using picture stimuli, Self + Suicide using word stimuli, Self + Death using word stimuli). Results replicated prior studies demonstrating that self-harm-related implicit associations were stronger among individuals with (vs. without) a history of suicide attempt and NSSI. Results also suggested that self-harm-related implicit associations are robust (based on internal replication), are sensitive to recency and severity of self-harm history (e.g., stronger associations for more recent and more lethal prior suicide attempts), and correlate with specific types of self-harm behaviors. These findings clarify the nature of self-harm-related implicit cognition and highlight the IAT's potential to track current risk for specific types of self-harm in ways that more fixed risk factors cannot.

## **General Scientific Summary**

Tools to objectively and reliably assess risk of suicide and self-injury are critically needed, given the serious limitations of self-report in this domain. This study supports the clinical potential of using measures of implicit associations (associations that are difficult to consciously control) between the self and death or self-harm to detect more recent and severe histories of self-harm, and to differentiate between distinct types of self-harm. These associations may ultimately help identify who is at most imminent risk for future self-harm behaviors.

**Keywords:** suicide, nonsuicidal self-injury (NSSI), Implicit Association Test (IAT), risk assessment, replication

This article was published Online First December 19, 2016.

Jeffrey J. Glenn and Alexandra J. Wertz, Department of Psychology, University of Virginia; S. J. Katarina Slama, Department of Psychology, University of California, Berkeley; Shari A. Steinman, Department of Psychology, West Virginia University; Bethany A. Teachman, Department of Psychology, University of Virginia; Matthew K. Nock, Department of Psychology, Harvard University.

Bethany A. Teachman has a significant financial interest in Project Implicit, Inc., which provided services in support of this project under contract with the University of Virginia. This research was supported by an

NIMH Grant (R34MH106770) and a Templeton Science of Prospection Award to Bethany A. Teachman, and by the Norlien Foundation and John D. and Catherine T. MacArthur Foundation to Matthew K. Nock. We thank the Project Implicit Mental Health collaborators, including Brian Nosek, Mahzarin Banaji, and Tony Greenwald. We would also like to thank Frederick Smyth, Emily Umansky, and Nauder Namaky for their technical assistance.

Correspondence concerning this article should be addressed to Jeffrey J. Glenn, Department of Psychology, University of Virginia, P.O. Box 400400, Charlottesville, VA 22904-4400. E-mail: [jglenn@virginia.edu](mailto:jglenn@virginia.edu)

Suicide is a leading cause of death worldwide, accounting for more deaths than by all interpersonal violence combined (Nock, Borges, & Ono, 2012). In the United States, suicide is responsible for over 40,000 deaths each year (1.6% of all mortality), making it the 10th leading cause of death (Centers for Disease Control and Prevention, 2013). In addition, the lifetime rate of Nonsuicidal self-injury (NSSI; self-harming behavior performed with no intent to die) is estimated at 5–6% among adults (Swannell, Martin, Page, Hasking, & St John, 2014). Although rarely lethal in itself, NSSI can cause severe lasting damage to body tissue and also is a robust predictor of future suicide attempt, even greater than many other well-established risk factors, such as depression and anxiety (Asarnow et al., 2011; Wilkinson, Kelvin, Roberts, Dubicka, & Goodyer, 2011).

Despite the serious nature and relatively high prevalence of self-harm (a term we will use to encompass self-harming behaviors both with and without intent to die), clinical psychological science has not yet produced the tools to accurately measure and predict suicidal thoughts and behavior. Much of the difficulty in doing so can be attributed to the near universal reliance on self-report, which is a crucial source of information for assessing suicide risk but also is limited for several reasons. First, unlike for other less sensitive or stigmatized behaviors, those at greatest risk may be intentionally motivated to conceal their thoughts (e.g., to avoid or gain release from hospitalization). Indeed, one study found that 78% of inpatients denied suicidal thoughts during the last verbal communication they had prior to killing themselves (Busch, Fawcett, & Jacobs, 2003). Second, people may lack the ability to accurately assess the factors impacting their current risk, either due to the transient nature of suicidal thoughts or a lack of conscious awareness (Wilson, 2009). As such, there is an urgent need for risk assessment tools that do not exclusively rely on, and can serve to supplement, self-report.

To overcome limitations of self-report, researchers have attempted to identify objective markers of suicide risk using behavioral tools, such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The IAT is a computer-administered implicit measure that is intended to tap into attitudes operating more automatically, that is, outside of conscious control and at a different level of processing, than explicit measures (Fazio & Olson, 2003). Major advantages of the IAT are that it is brief, easy to administer, and less susceptible to social desirability bias than explicit self-reports (Greenwald, Nosek, & Banaji, 2003). Indeed, a meta-analysis of studies using the IAT across a wide variety of social domains found that the IAT was incrementally predictive beyond self-report, especially so for socially sensitive topics such as racially discriminatory behavior (Greenwald, Pohlman, Uhlmann, & Banaji, 2009). Although there is debate in the social psychology literature regarding the strength and significance of the IAT's predictive validity and its ability to reliably predict behaviors on an individual level (e.g., for a recent exchange on this topic, see Greenwald, Banaji, & Nosek, 2015 and Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2015), prior research has consistently supported the notion that implicit association measures provide unique information above, and at least partially distinct from, explicit self-report.

Several recent studies have suggested that the IAT can be used to measure implicit cognition about self-harm and, in doing so, may help to improve the prediction and prevention of this behav-

ior. These initial studies have suggested that self-harm IATs measuring the extent to which people associate NSSI and death/suicide with the self can both distinguish those engaging in self-harm from non-self-harming control participants and be used to improve the prediction of future self-harm above and beyond other well-known predictors (Cha et al., 2016; Glenn, Kleiman, Cha, Nock, & Prinstein, 2016; Nock & Banaji, 2007a, 2007b; Nock et al., 2010; Randall, Rowe, Dong, Nock, & Colman, 2013).

Although findings to date suggest that measures of implicit cognition can be used as behavioral markers that increase our ability to identify those at greatest risk for self-harm, several key questions remain regarding the measurement of self-harm-related implicit cognition. The greatest limitation of prior studies in this area is that they have used relatively small samples. It is well-documented that studies with low statistical power can lead not only to a failure to detect a true effect, but also to a reduction in the likelihood that any significant effect is representative of the true effect in the population (Button et al., 2013). This is especially concerning in the study of suicidal and self-injurious behaviors, where the outcomes of interest can involve serious human injury or death. Thus, it is imperative to have an accurate estimate of the magnitude and nature of the associations between implicit cognition and different forms of self-harm. The primary aim of this study was to collect data on the association between self-harm-related implicit cognition and a history of engaging in self-harm. Consistent with recent recommendations regarding the importance of sufficiently powered studies as well as replication of observed effects (Button et al., 2013; Nosek, Spies, & Motyl, 2012; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012), we sought to collect data from a large unselected sample of participants, and to replicate all tests conducted to assess the reliability of our findings.

Prior studies suggest that implicit cognition about self-harm differs between those who engage in this behavior and those who do not; however, a vital and unanswered question is the extent to which implicit cognition about self-harm is *malleable* and changes over time. Although there is some debate about the nature of observed malleability in the IAT literature (e.g., attitude malleability vs. measurement malleability; Han, Czellar, Olson, & Fazio, 2010), numerous studies have shown that the IAT can track attitudinal change (although often weakly; Joy-Gaba & Nosek, 2010) and show change over treatment in clinical populations (e.g., Teachman, Marker, & Smith-Janik, 2008; Teachman & Woody, 2003). We do not yet know whether implicit self-harm attitudes change over time. The finding that implicit associations with self-harm differ as a function of recency of behavioral engagement (e.g., how long ago one attempted suicide) would indicate the IAT's potential capability to track history of self-harm and, by extension, imminence of risk (given more recent engagement may mean higher current risk). A related question about self-harm-related implicit cognition is whether it is associated with the frequency or severity of self-harm behavior.

Due to small sample sizes and a limited number of IAT versions administered per study, no studies have yet addressed the issue of the self-harm IAT's ability to distinguish between specific types of self-harm (e.g., with vs. without intent to die). One prior study found that a self-injury IAT using NSSI pictures differentiated between individuals with a prior suicide attempt, those with prior suicide ideation but no attempt, and those with no history of

ideation or attempt (Nock & Banaji, 2007b). However, given the high comorbidity between suicide ideation/attempts and NSSI, it is not clear whether self-harm in general, or suicidal thoughts and behaviors (or NSSI) specifically, explain these prior findings demonstrating group differences. Better understanding the specificity of the associations between different versions of the IAT and specific self-harm behaviors would allow us to tailor IATs to produce the strongest effects within a given population and thereby maximize their utility as risk assessment tools.

The purpose of this study was to test the magnitude and nature of the associations between self-harm-related implicit cognition and engagement in self-harm to answer four open questions: (a) do self-harm-related implicit cognitions differ between those with versus without a history of different forms of self-harm (replicating the results from earlier studies?); (b) do these effects differ by recency of self-harm?; (c) do these effects differ by frequency or severity of self-harm?; and (d) is there specificity between different versions of the IAT and specific forms of self-harm (e.g., suicidal vs. nonsuicidal)? We tested these questions by collecting data from a large, unselected sample of participants via a publicly accessible, web-based platform (PIMH), and examined the robustness of the findings by internally replicating all analyses.

## Method

### Participants

Participants were 7,015 adult volunteers who consented to participate and completed at least one IAT at the PIMH research website ([www.ImplicitMentalHealth.com](http://www.ImplicitMentalHealth.com)) between March 2012 and October 2014. The PIMH website allows users to explore mental health-related IATs. Participants included in this study are those who chose to complete a Self-harm IAT and were randomly assigned to one of three IATs (described below): Self-Injury ( $n = 2,332$ ), Death ( $n = 2,298$ ), and Suicide ( $n = 2,385$ ). Demographic characteristics by IAT subsample are provided in Table 1. Adult participants who consented but did not complete an IAT ( $n = 2,905$ ) were excluded from analyses. Part way through data collection (i.e., after 776 consenting participants), the consent form was updated to include a question regarding the participant's age; individuals under the age of 18 ( $n = 169$ ) also were excluded due to minors' inability to give informed consent.

### Measures

**Sociodemographics.** Sociodemographic information was collected from participants, including age, ethnicity, race, education, and country of residence.<sup>1</sup>

**Implicit cognition.** Implicit cognition about self-harm was measured using the IAT (Greenwald et al., 1998), a task that uses reaction times (RTs) when classifying different stimuli to measure the automatic mental associations they hold between concepts and attributes (e.g., suicide = me). Three different self-harm IATs used in prior studies were examined: the Self-Injury IAT, Death IAT, and Suicide IAT, each of which used the standard IAT structure to measure the extent to which participants associated self-harm-related constructs (i.e., death/suicide and self-injury) with the self (see the Appendix for complete stimulus sets).

During the IAT, participants sort words (or images) one at a time into superordinate categories presented in the top left and right corners of the screen by pressing corresponding keys on the keyboard. For example, in one critical category pairing condition, participants may see *death + me* paired in the upper left corner and *life + not me* paired in the upper right corner. As stimuli related to each of these four categories appear one at a time in the center of the screen, participants sort them as quickly as possible into their superordinate categories by pressing a key on the left of the keyboard ("E") for *death* or *me* words (e.g., "dead," "self," "suicide," "myself") and a key on the right ("I") for *life* or *not me* words (e.g., "alive," "them," "thrive," "their"). They then complete an analogous sorting task for the other critical category pairing condition in which *life + me* are paired together and *death + not me* are paired together. We followed the standard seven-block IAT structure and scoring procedure (Greenwald et al., 2003), and calculated a *D* score for each person on each IAT in which positive scores represent a stronger association between self-harm and self (e.g., faster responding when *death* and *me* are paired to the same key relative to when *life* and *me* are paired), and negative scores represent a stronger association between non-self-harm and self.

**Self-harming thoughts and behaviors.** Participants' history of self-injurious thoughts and behaviors was assessed using an abbreviated (20-item) self-report version of the Self-Injurious Thoughts and Behaviors Interview (SITBI; Nock, Holmberg, Photos, & Michel, 2007). Questions assessed history of NSSI ("Have you ever done anything to purposely hurt yourself without wanting to die [for example cutting or burning your skin]?"), suicidal thoughts ("Have you ever had thoughts of killing yourself?"), suicide plans ("Have you ever actually made a plan to kill yourself?"), and suicide attempts ("Have you ever made an actual suicide attempt, where you wanted to kill yourself, even just a little?" and, to identify the subset of attempts requiring medical attention, "If you have ever attempted suicide, did any attempt result in an injury, poisoning, or overdose that had to be treated by a doctor or nurse?"). Individuals were asked to rate the frequency of each behavior within their lifetime, the past year, and the past week. The interview version of the SITBI, which uses the same wording as this self-report version, has shown strong test-retest and interrater reliability, as well as concurrent validity in United States, German, and Spanish samples (Fischer et al., 2014; García-Nieto, Blasco-Fontecilla, Yepes, & Baca-García, 2013; Nock et al., 2007). Further, these items are very similar to those included in the most widely used self-report measures of the same constructs (e.g., Youth Risk Behavior Survey, Columbia Suicide Severity Rating Scale, World Mental Health Composite International Diagnostic Interview Suicide Module). Participants were also asked about current desire to harm oneself and desire to die as part of managing current risk (see the Procedure section).

Notably, while we recognize the limits of relying on self-report measures of these behaviors, prior research suggests participants are more comfortable and willing to report sensitive information online (e.g., Shapiro, Chandler, & Mueller, 2013). Therefore, we

<sup>1</sup> Additional sociodemographic information and state assessment of mood were also collected, but are not examined here given they are not central to current hypotheses or were reported elsewhere. Those data are available upon request from the corresponding author.

Table 1  
*Participant Characteristics by Self-Harm IAT Version Combined Across Samples*

Variable	Self-Injury IAT ( <i>n</i> = 2,063)	Death IAT ( <i>n</i> = 2,042)	Suicide IAT ( <i>n</i> = 2,124)
Mean ( <i>SD</i> ) age in years	27.04 (10.43)	27.34 (10.83)	27.58 (10.89)
Sex (% female)	67.9	66.9	67.0
Citizenship (%)			
U.S.	60.9	59.6	60.3
Non-U.S.	34.7	36.1	35.5
Did not report	4.4	4.3	4.2
Ethnicity (%)			
Non-Hispanic	80.4	80.5	80.4
Hispanic	8.2	8.4	8.9
Did not report	11.4	11.1	10.6
Race (%)			
Caucasian	75.4	73.6	72.9
Asian	6.2	7.3	7.3
African American	3.9	2.4	3.5
Multiracial	4.5	5.0	5.2
Other	6.5	8.1	7.9
Did not report	3.5	3.4	3.1
Education (%)			
Graduate degree	15.6	17.8	16.0
Bachelor's degree	21.9	22.3	22.3
Some college	34.0	32.2	33.9
High school	10.6	11.3	11.1
Less than high school	5.5	5.0	5.2
Did not report	12.4	11.5	11.6
Lifetime history of self-harm (%)			
Nonsuicidal self-injury (NSSI)	58.7	59.2	56.0
Suicide ideation	82.2	82.4	81.2
Suicide plan	38.8	37.5	36.9
Suicide attempt	29.0	27.8	29.3
Any self-harm thought/behavior	87.0	86.7	85.1

Note. IAT = Implicit Association Test.

believe the anonymous nature of the study actually makes it less likely that participants' responses were influenced by efforts at impression management compared with clinical or lab settings (where participants might fear their responses could lead to unwanted interventions or hospitalization).

## Procedure

After launching the PIMH website, we posted links to it on the main Project Implicit site ([www.ProjectImplicit.com](http://www.ProjectImplicit.com)) and our academic websites, and we referenced it in academic and public talks and in news stories about the topics assessed by the PIMH tests. Participants visiting the PIMH site were presented with informed consent agreements, which included information regarding PIMH's privacy policy, mental health resources, and contact information for the primary investigators.

Following informed consent, participants were randomly assigned to complete one of the three IATs (Self-Injury IAT, Death IAT, or Suicide IAT), sociodemographic questionnaire, and SITBI in a random order. After the final task, participants were shown the first debriefing page, with information regarding the purpose of the study. Participants were then given the option of viewing their IAT feedback on the following page (e.g., "Your responses show that you sorted words much faster when CUTTING and ME were paired on the same key (relative to CUTTING and NOT ME),

which suggests that you may have a strong implicit association between self-injury and yourself"). If an individual responded to any item regarding current desire to die/hurt oneself with "extremely," he or she received a special note on the first debriefing page:

Your responses on the survey you filled out suggest that you may want to hurt yourself or die. We are concerned that you are having these thoughts. We encourage you to contact someone who can help you cope with these thoughts and with any stressful events that may have caused them.

Additional mental health resource information was provided below this note. This study was approved by the Harvard University, University of Virginia, and University of Washington institutional review boards.

## Data Analytic Plan

All implicit association measures were scored following recommendations from Greenwald et al. (2003). Scoring algorithms create an IAT *D* score, which is the difference between response latencies for the two critical category pairing conditions, divided by the standard deviations across all blocks. The *D* score is conceptually similar to a Cohen's *d* effect size. Based on these scoring guidelines (e.g., cut points to exclude data with high error

rates or too many fast trials<sup>2</sup>), we excluded the following: 269 (11.5%) participants' Self-Injury IAT data, 256 (11.1%) participants' Death IAT data, and 261 (10.9%) participants' Suicide IAT data. Exclusion rates were comparable to previous web-based IAT studies (see Nosek et al., 2007, for the typical range of exclusion rates).

Prior to performing analyses, all study variables were checked for missing data, normality, and outliers. To adjust for moderate positive skewness among IAT  $D$  scores, a log transformation (with the addition of a constant of 2.58 so that the lowest value was equal to 1) was performed. To adjust for moderate positive skewness among the NSSI and suicide attempt frequency variables (restricted to participants reporting lifetime NSSI and suicide attempt, respectively, only), log transformations were performed. In each case, the resulting transformed data closely approximated a normal distribution (skewness between  $-2$  and  $+2$ ) and were used in all statistical analyses (Nolan, Heinzen, & Weathersby, 2008). Figures reflect participants' untransformed IAT  $D$  scores in order to retain the conceptual meaning of the  $D$  score (e.g., score greater than zero indicates stronger association between self-harm and the self, compared with non-self-harm and the self).

Median absolute deviation (MAD) was used to identify extreme outliers in the IAT  $D$  score and frequency of self-harming behaviors variables. MAD is a more robust dispersion statistic than standard deviation, which is particularly influenced by outliers, and offers an alternative to traditional outlier detection practices (Leys, Ley, Klein, Bernard, & Licata, 2013). Following Leys et al. (2013), the following cases with a MAD of 3 or more were identified from the variables for final analyses: 1 ( $<0.1\%$ ) of IAT  $D$  scores overall, 29 (0.9%) of NSSI lifetime frequency, and 11 (0.7%) of suicide attempt lifetime frequency. To reduce their influence, outliers were reassigned values of the closest most extreme nonoutlier (Tabachnick & Fidell, 2001). Subsequent results did not significantly change when analyses were performed with and without outliers included; therefore, we included the outliers in our analyses. Internal consistency of the IAT was very high; the Spearman-Brown corrected split-half reliability, determined from  $D$  scores calculated separately for odd and even trials, was  $r(6,628) = .91$ , similar to other IATs administered previously using a similar platform ( $r_s = .89-.92$ ; Greenwald & Nosek, 2001).

Following data reduction and cleaning, preliminary analyses were performed to test whether demographic differences existed among the entire sample based on lifetime presence of self-harm (i.e., NSSI, suicide ideation, suicide attempt) using chi-square tests for categorical measures and  $t$  tests for continuous measures. We then conducted analyses to test each of our four research questions.

First, we tested whether self-harm-related implicit cognition differs between those with versus without a history of self-harm (i.e., NSSI vs. no such history; suicide attempt vs. no such history) using independent samples  $t$  tests. To provide additional context regarding the degree to which the IAT is predictive of group status based on self-harm history, receiver operating characteristic (ROC) analyses were performed as follow-up analyses for IATs exhibiting the strongest effects. In addition to determining the area under the curve (AUC) for each ROC plot, four additional terms were calculated: *sensitivity* (proportion of participants with an actual history of self-harm correctly identified by the IAT [i.e.,  $D$  score  $> 0$ ]), *specificity* (proportion of participants without a his-

tory of self-harm correctly identified by the test [i.e.,  $D$  score  $< 0$ ]), *positive predictive value* (proportion of individuals with a positive test who were correctly classified as having a history of self-harm), and *negative predictive value* (proportion of individuals with a negative test who were correctly classified as having no history of self-harm). To perform these calculations, IAT scores were dichotomized using a cutpoint of 0 to indicate neither an association between self-harm and the self ( $D$  score  $> 0$ ) nor between non-self-harm and the self ( $D$  score  $< 0$ ), a similar approach to those used in prior studies (e.g., Nock et al., 2010; Randall et al., 2013).

Second, we tested whether implicit associations between self-harm and the self differ as a function of recency of participants' engagement in self-harm using factorial analyses of variance (ANOVAs) and post hoc comparisons. Specifically, for NSSI, we performed a 4 (NSSI recency group: past week NSSI, past year but not past week NSSI, lifetime but not past year NSSI, no lifetime NSSI)  $\times$  3 (IAT version: Self-Injury IAT, Death IAT, Suicide IAT) independent factorial ANOVA with both NSSI recency group and IAT version as between-subjects factors with the interaction term included in the model. For suicide attempt, a 3 (suicide attempt recency group<sup>3</sup>: Past year attempt, lifetime but not past year attempt, no lifetime attempt)  $\times$  3 (IAT version) independent factorial ANOVA with both suicide attempt recency group and IAT version as between-subjects factors. IAT version was entered as a factor to examine whether the pattern of associations within the self-harm recency groups differed as a function of specific IAT.

Third, we tested whether the IAT was associated with frequency and severity of self-harm. Zero-order Pearson correlations were calculated examining the relations between implicit associations and lifetime frequency of NSSI, lifetime frequency of suicide attempt, and lifetime frequency of severe attempts (i.e., those requiring medical attention). Given the majority of participants had neither a lifetime history of NSSI nor a lifetime history of suicide attempt, correlations were performed only among lifetime self-injurers or attempters. We also tested whether IAT effects differed by severity of suicidality by comparing results across three groups: lifetime suicide attempt, lifetime suicide ideation (but no attempt), and neither lifetime ideation nor attempt, using one-way ANOVAs.

Fourth and finally, we examined the extent to which performance on different versions of the IAT maps on to specific types of self-harm behaviors. To do so, we created subgroups of participants who engaged in: Only NSSI (i.e., no history of suicide ideation or attempt) or only suicide ideation or attempts (i.e., but never NSSI) and tested whether differences on the Self-Injury IAT were specific to the former and differences on the Death and Suicide IATs were specific to the latter.

Due to the benefit of a very large sample size and the unique opportunity for internal replication, we split our dataset (6,229

<sup>2</sup> Specifically, IAT scores meeting any of the following criteria were excluded: (a) missing all trials from a block, (b) greater than 10% of trial RTs faster than 300 ms overall, (c) greater than 25% of trial RTs faster than 300 ms in any critical block, (d) greater than 30% trial errors overall, and (e) greater than 40% trial errors in any critical block.

<sup>3</sup> Past week attempters were not treated as a separate group, given that only 33 participants attempted suicide during the past week.

participants) in half using random selection and conducted our planned analyses on the first sample (i.e., Sample 1;  $N = 3,115$ ) and then conducted identical confirmatory analyses on the second sample (i.e., Sample 2;  $N = 3,114$ ) to cross-validate our results (Wagenmakers et al., 2012). Alpha was set at .05 for all two-tailed significance tests and post hoc pairwise comparisons were adjusted using Bonferroni correction for all analyses.

## Results

### Sociodemographic Differences Based on Self-Harm History

Across participants in Samples 1 and 2, lifetime self-injurers ( $M = 25.41$ ,  $SD = 8.75$ ) were significantly younger than noninjurers<sup>4</sup> ( $M = 29.85$ ,  $SD = 12.39$ ),  $t(5,958) = 16.21$ ,  $p < .001$ ,  $d = 0.42$ , but lifetime suicide attempters ( $M = 26.89$ ,  $SD = 10.29$ ) were not significantly younger than suicide nonattempters ( $M = 27.47$ ,  $SD = 10.84$ ). Women had higher rates of NSSI (64.3%) than men (44.3%),  $\chi^2(1, N = 5,945) = 213.65$ ,  $p < .001$ ,  $\phi = .19$ , and suicide attempt (32.9%) than men (19.7%),  $\chi^2(1, N = 5,913) = 111.21$ ,  $p < .001$ ,  $\phi = .14$ . Hispanic/Latino participants had a higher rate of suicide attempt (33.5%) than non-Hispanic/Latino participants (28.2%),  $\chi^2(1, N = 5,442) = 6.35$ ,  $p = .012$ ,  $\phi = .03$ , but there was no significant difference in rate of NSSI between Hispanic/Latino (56.4%) and non-Hispanic/Latino participants (58.1%),  $\chi^2(1, N = 5,470) = 0.76$ ,  $p = .384$ ,  $\phi = .01$ . White participants had a higher rate of NSSI (59.5%) compared with non-Whites (53.4%),  $\chi^2(1, N = 5,896) = 15.66$ ,  $p < .001$ ,  $\phi = .05$ , and non-White participants had a slightly higher rate of suicide attempt (30.9%) compared with Whites (28.1%),  $\chi^2(1, N = 5,864) = 4.04$ ,  $p = .044$ ,  $\phi = .03$ .

To ensure that any observed effects were not due to the sociodemographic differences noted above, we created case-control subsamples that were matched on age, gender, race, and ethnicity (with cases being those with a history of each type of self-harm and controls those with no such history) and conducted all study analyses on these sociodemographically equivalent groups. Results were the same for these matched samples (results available upon request), suggesting that group differences on the IAT are not the result of observed sociodemographic differences between self-harm and non-self-harm participants. Therefore, all analyses reported subsequently are from original nonmatched samples.

### Between-Group Differences in Implicit Cognition

Results revealed the hypothesized between-groups differences on the Self-Injury IAT, Death IAT, and Suicide IAT (i.e., higher scores for those with vs. without a history of self-harm; Table 2). The largest effect size in the NSSI versus control (no history of NSSI) comparison was observed on the Self-Injury IAT ( $d = 0.81$ ), whereas the largest effect size in the suicide attempter versus control (no attempt history) comparison was observed on the Suicide IAT ( $d = 0.54$ ), suggesting that effects were strongest when IAT stimuli matched the specific target behaviors. The Self-Injury IAT was the only IAT version in which lifetime self-injurers and suicide attempters had a positive score (indicating a stronger identification with self-injury than noninjury), whereas these self-harm groups had less negative scores than their respec-

tive controls on the Death IAT and Suicide IAT (indicating a weaker identification with life, but not a stronger identification with death than life; Figure 1A and 1B). As shown in Table 2, results from the replication sample confirmed these findings.

When predicting lifetime presence of NSSI, the Self-Injury IAT produced an AUC of 0.72 (95% CI [0.69, 0.75],  $p < .001$ ), indicating moderate accuracy. The dichotomized Self-Injury IAT, using a cutoff score of 0, produced fair sensitivity (53%) and negative predictive value (59%) and strong specificity (82%) and positive predictive value (82%). When predicting a lifetime history of suicide attempt, the Suicide IAT produced an AUC of 0.64 (95% CI [0.60–0.67],  $p < .001$ ), indicating only fair accuracy. The dichotomized Suicide IAT score produced poor sensitivity (27%), fair positive predictive value (47%), and strong specificity (87%) and negative predictive value (73%). AUCs were slightly larger in the replication sample (Self-Injury IAT AUC = 0.76, Suicide IAT AUC = 0.66) but analyses were otherwise similar.

### Recency of Behavior and Self-Harm-Related Implicit Cognition

Examination of these effects as a function of time revealed that IAT scores were consistently higher for those who had engaged in self-harm more recently (see Figures 1C and 1D). In the case of NSSI (Figure 1C), a 4 (NSSI recency group)  $\times$  3 (IAT version) independent factorial ANOVA revealed significant main effects of NSSI recency group,  $F(3, 3005) = 161.10$ ,  $p < .001$ ,  $\eta_p^2 = .14$ , and IAT version,  $F(2, 3005) = 124.64$ ,  $p < .001$ ,  $\eta_p^2 = .08$ . Post hoc pairwise comparisons revealed significant differences between each group, such that participants who engaged in NSSI during the past week ( $n = 283$ ) had the strongest implicit associations (vs. past year,  $p < .001$ ), followed in order by past year NSSI ( $n = 706$ ; vs. lifetime NSSI,  $p < .001$ ), lifetime NSSI ( $n = 761$ ; vs. no lifetime NSSI,  $p < .001$ ), and no history of NSSI ( $n = 1,267$ ). Post hoc comparisons between IAT versions revealed that effects for the Self-Injury IAT were stronger (i.e., larger association with self-harm) than those for the Death IAT and Suicide IAT (both  $p < .001$ ), which also significantly differed from one another but to a weaker degree ( $p = .005$ ). On balance, the significant but negligible effect size of the NSSI Group  $\times$  IAT Version interaction,  $F(6, 3005) = 4.62$ ,  $p < .001$ ,  $\eta_p^2 = .01$ , suggests the patterns of implicit associations among NSSI groups were largely not different across IATs, indicating comparable sensitivity to recency of self-harm behavior across IATs.

In the case of suicide attempts (Figure 1D), a 3 (suicide attempt recency group)  $\times$  3 (IAT version) independent factorial ANOVA revealed significant main effects of suicide attempt recency group,  $F(2, 2988) = 72.73$ ,  $p < .001$ ,  $\eta_p^2 = .05$ , and IAT version,  $F(2, 2988) = 60.32$ ,  $p < .001$ ,  $\eta_p^2 = .04$ . Post hoc pairwise comparisons indicated significant differences between each group, such that participants who attempted during the past year ( $n = 223$ ) had the strongest implicit associations (vs. lifetime attempt,  $p < .001$ ), followed in order by lifetime attempt ( $n = 645$ ; vs. no lifetime

<sup>4</sup>This is consistent with prior studies, which report higher rates of lifetime NSSI in younger respondents than older respondents (e.g., Swanell et al., 2014), likely due to a combination of problems with memory/reporting bias in adults and increasing rates of NSSI over the past few decades.

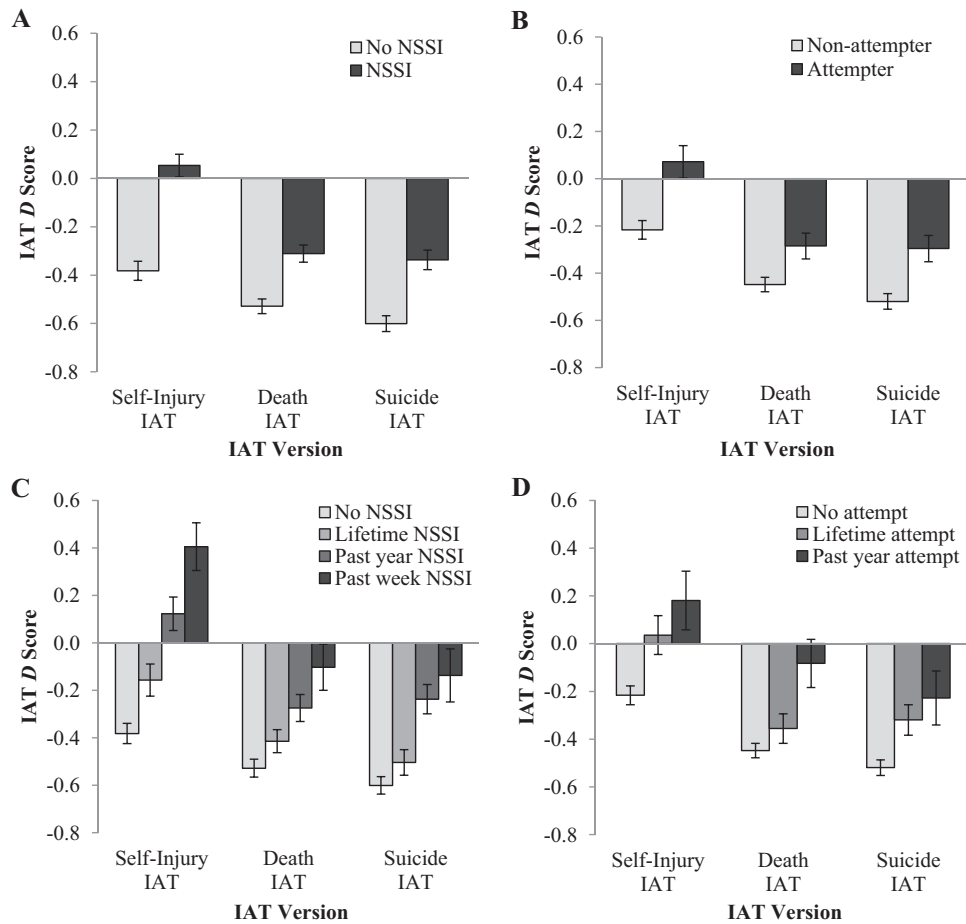
**Table 2**  
*Group Differences in Implicit Association (IAT D Scores) by Self-Harm IAT Version for Initial and Replication Samples*

Group comparison	Sample 1 (N = 3,115)			Sample 2 (N = 3,114)		
	Test	<i>d</i>	<i>p</i>	Test	<i>d</i>	<i>p</i>
<b>NSSI vs. control</b>						
Self-Injury IAT	<i>t</i> (985) = 12.34	.81	<.001	<i>t</i> (970) = 14.52	.97	<.001
Death IAT	<i>t</i> (989) = 8.09	.52	<.001	<i>t</i> (988) = 9.39	.61	<.001
Suicide IAT	<i>t</i> (1001) = 9.00	.58	<.001	<i>t</i> (1010) = 11.28	.72	<.001
<b>Suicide attempt vs. control</b>						
Self-Injury IAT	<i>t</i> (997) = 7.55	.52	<.001	<i>t</i> (986) = 7.91	.54	<.001
Death IAT	<i>t</i> (984) = 5.32	.37	<.001	<i>t</i> (982) = 4.49	.32	<.001
Suicide IAT	<i>t</i> (1017) = 8.16	.54	<.001	<i>t</i> (1029) = 10.01	.67	<.001

*Note.* IAT = Implicit Association Test; NSSI = nonsuicidal self-injury.

attempt,  $p < .001$ ), and no lifetime attempt ( $n = 2,129$ ). Post hoc comparisons between IAT versions revealed that self-harm-related implicit associations on the Self-Injury IAT were stronger (i.e., larger association with self-harm) than for the Death IAT and

Suicide IAT (both  $p < .001$ ) and the Death IAT was slightly stronger than the Suicide IAT ( $p = .011$ ). On balance, the non-significant Suicide Attempt Group  $\times$  IAT Version interaction,  $F(6, 2988) = 2.11, p = .077, \eta_p^2 = .003$ , again suggests that the



**Figure 1.** Differences in performance on Self-Harm IAT versions by lifetime presence of (A) NSSI and (B) suicide attempt, and recency of engagement in (C) NSSI and (D) suicide attempt for the initial sample. Positive D scores indicate a stronger association between self-harm and self (vs. non-self-harm and self); negative D scores indicate a stronger association between non-self-harm and self (vs. self-harm and self). IAT = Implicit Association Test; NSSI = nonsuicidal self-injury. Error bars represent 95% confidence intervals.

patterns of implicit associations among suicide attempt groups largely were not different across IATs.

Results from the replication sample confirmed these findings. In the case of NSSI, analyses again revealed significant main effects of NSSI recency,  $F(3, 3002) = 214.41, p < .001, \eta_p^2 = .18$ , and IAT version,  $F(2, 3002) = 118.30, p < .001, \eta_p^2 = .07$ . Participants who engaged in NSSI during the past week ( $n = 307$ ) had the strongest implicit associations (vs. past year,  $p = .037$ ), followed in order by past year NSSI ( $n = 707$ ; vs. lifetime NSSI,  $p < .001$ ), lifetime NSSI ( $n = 731$ ; vs. no lifetime NSSI,  $p < .001$ ), and no history of NSSI ( $n = 1,269$ ). Self-harm-related implicit associations on the Self-Injury IAT were stronger compared to both the Death IAT and Suicide IAT (both  $p < .001$ ), which, unlike in the initial sample, did not significantly differ from one another ( $p = 1.000$ ). The significant but negligible effect size of the NSSI Group  $\times$  IAT Version interaction,  $F(6, 3002) = 3.95, p = .001, \eta_p^2 = .01$ , again suggests the patterns of implicit associations among NSSI groups were largely not different across IATs.

In the case of suicide attempts, analyses again revealed significant main effects of suicide attempt recency,  $F(2, 2983) = 77.67, p < .001, \eta_p^2 = .05$ , and IAT version,  $F(2, 2983) = 57.27, p < .001, \eta_p^2 = .04$ . Participants who attempted during the past year ( $n = 211$ ) had the strongest implicit associations (vs. lifetime attempt,  $p < .001$ ), followed, in order, by lifetime attempt ( $n = 641$ ; vs. no lifetime attempt,  $p < .001$ ), and no lifetime attempt ( $n = 2,140$ ). Self-harm-related implicit associations on the Self-Injury IAT were stronger than for the Death IAT and Suicide IAT (both  $p < .001$ ), which, unlike in the initial sample, did not significantly differ from one another ( $p = 1.000$ ). The significant but negligible Suicide Attempt Group  $\times$  IAT Version interaction,  $F(6, 2983) = 2.45, p = .044, \eta_p^2 = .003$ , again suggests the patterns of implicit associations among suicide attempt groups largely did not differ across IATs.

### Frequency and Severity of Behavior and Self-Harm-Related Implicit Cognition

Among self-injurers, lifetime frequency of NSSI was significantly correlated with all three IATs and most strongly with the Self-Injury IAT ( $r = .39$ ; Table 3). Among suicide attempters, lifetime frequency of suicide attempt was significantly correlated with the Death IAT ( $r = .12$ ), but not the Suicide IAT or Self-Injury IAT, and the number of lifetime attempts requiring serious medical attention was not significantly correlated with any of the

IATs. As shown in Table 3, results from the replication sample confirmed the strong correlation between lifetime frequency of NSSI and the Self-Injury IAT. However, correlations between lifetime frequency of suicide attempt and IAT scores were stronger in the replication sample. Among attempters, lifetime frequency of suicide attempt was significantly correlated with the Death IAT ( $r = .27$ ), and, unlike in the initial sample, also with both the Suicide IAT ( $r = .12$ ) and Self-Injury IAT ( $r = .14$ ). Unlike in the initial sample, the number of lifetime attempts requiring serious medical attention was significantly correlated with the Death IAT ( $r = .25$ ).

In terms of severity of suicidality, one-way ANOVAs comparing participants with a lifetime suicide attempt ( $n = 868$ ), lifetime suicide ideation without attempt ( $n = 1,611$ ), and neither lifetime ideation nor attempt ( $n = 512$ ) demonstrated an overall difference among groups on the Self-Injury IAT,  $F(2, 990) = 34.66, p < .001, \eta^2 = .07$ , Death IAT,  $F(2, 980) = 33.90, p < .001, \eta^2 = .07$ , and Suicide IAT,  $F(2, 1012) = 47.24, p < .001, \eta^2 = .09$ . Post hoc pairwise comparisons revealed that for all three IATs, lifetime attempters had stronger self-harm-related implicit associations than lifetime ideators (all  $ps < .001-.002$ ), and lifetime ideators had stronger self-harm-related implicit associations than those without any lifetime suicide ideation or suicide attempt (all  $ps < .001$ ). These results indicate the self-harm IAT's ability to specifically isolate participants with more severe suicide histories (suicide attempt vs. ideation).

Results from the replication sample confirmed these findings. One-way ANOVAs comparing participants with a lifetime suicide attempt ( $n = 852$ ), lifetime suicide ideation without attempt ( $n = 1,581$ ), and neither lifetime ideation nor attempt ( $n = 552$ ) were significant for the Self-Injury IAT,  $F(2, 978) = 35.37, p < .001, \eta^2 = .07$ , Death IAT,  $F(2, 978) = 33.08, p < .001, \eta^2 = .06$ , and Suicide IAT,  $F(2, 1020) = 59.40, p < .001, \eta^2 = .10$ . Post hoc pairwise comparisons revealed significant differences such that lifetime attempters had stronger implicit associations than lifetime ideators for the Self-Injury IAT and Suicide IAT (both  $ps < .001$ ; for the Death IAT,  $p = .050$ ), and lifetime ideators had stronger implicit associations than those without any lifetime suicide ideation or attempt for all three IATs ( $ps < .001$ ).

### Specificity of Self-Harm-Related Implicit Cognition

Lastly, we examined the specificity of associations between performance on the IAT and distinct types of self-harm. After

Table 3

Correlations Between Implicit Association (IAT D Scores) and Lifetime Frequencies of NSSI, Suicide Attempt, and Suicide Attempts Requiring Serious Medical Attention for Initial and Replication Samples

Self-harming behavior	Sample 1 ( $N = 3,115$ )			Sample 2 ( $N = 3,114$ )		
	Self-Injury IAT	Death IAT	Suicide IAT	Self-Injury IAT	Death IAT	Suicide IAT
NSSI <sup>a</sup>	.39***	.20***	.26***	.38***	.32***	.21***
Suicide attempts (total) <sup>b</sup>	.03	.12*	.02	.14*	.27***	.12*
Suicide attempts (requiring serious medical attention) <sup>b</sup>	.03	.02	.04	.11	.25*	.11

Note. IAT = Implicit Association Test; NSSI = non suicidal self-injury.

<sup>a</sup> Performed among lifetime self-injurers only. <sup>b</sup> Performed among lifetime suicide attempters only.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

removing those with a lifetime history of suicide ideation and/or attempt from the sample, we found significant differences between those with a lifetime history of NSSI ( $n = 112$ ) and those without a history of NSSI ( $n = 398$ ) on the Self-Injury IAT,  $t(168) = 3.75$ ,  $p < .001$ ,  $d = 0.58$ , but not on the Death IAT,  $t(158) = 1.93$ ,  $p = .055$ ,  $d = 0.31$ , or Suicide IAT,  $t(178) = 0.69$ ,  $p = .489$ ,  $d = 0.10$ , indicating that those engaging in NSSI had an implicit identification with NSSI but not with death or suicide. In separate analyses, after removing participants with a history of NSSI from our sample, we found significant differences between those with lifetime suicide ideation and/or attempt ( $n = 861$ ) and those with no such history ( $n = 398$ ) on both the Death IAT,  $t(400) = 4.38$ ,  $p < .001$ ,  $d = 0.44$ , and Suicide IAT,  $t(433) = 4.33$ ,  $p < .001$ ,  $d = 0.42$ , but not on the Self-Injury IAT,  $t(420) = 1.84$ ,  $p = .066$ ,  $d = 0.18$ , indicating that those with a history of suicidal thoughts and behaviors had a weaker implicit identification with life, but not an implicit identification with NSSI.

Results from the replication sample largely confirmed these findings. After removing suicide ideators and attempters from the sample, those with a history of NSSI ( $n = 128$ ) differed significantly from those without a history of NSSI ( $n = 424$ ) on the Self-Injury IAT,  $t(173) = 5.26$ ,  $p < .001$ ,  $d = 0.80$ , and, unlike in the initial sample, also on the Suicide IAT,  $t(199) = 2.87$ ,  $p = .005$ ,  $d = 0.41$ , but not the Death IAT,  $t(174) = 1.76$ ,  $p = .081$ ,  $d = 0.27$ . After removing those with a history of NSSI, suicide ideators and attempters ( $n = 832$ ) differed significantly from those with no suicide history ( $n = 424$ ) on the Suicide IAT,  $t(459) = 4.37$ ,  $p < .001$ ,  $d = 0.41$ , and Death IAT,  $t(398) = 3.52$ ,  $p < .001$ ,  $d = 0.35$ , but not the Self-Injury IAT,  $t(393) = 1.77$ ,  $p = .077$ ,  $d = 0.18$ .

## Discussion

This study provided important evidence of replication and novel findings about implicit self-harm associations. First, people with a history of suicide attempt and NSSI had a stronger implicit identification with self-harm (i.e., a stronger association between self-harm and the self) than those with no such history—replicating the results from earlier, smaller studies on this topic. Second, implicit associations about self-harm tracked the recency of self-harm behavior, such that the more recent the self-harm, the higher the scores on the IAT. Third, implicit associations about self-harm were stronger, in general, among those with more frequent and more severe self-harm, although not all of these effects replicated across both samples. Fourth, the results of this study yielded evidence for specificity, such that when excluding from analyses participants with a history of suicidal thoughts or behaviors, between-groups differences for those who did versus did not engage in NSSI were largest for the Self-Injury IAT and were nonsignificant for the Death IAT and, in the initial sample, the Suicide IAT. Likewise, when excluding from analyses participants with a history of NSSI, between-groups differences for those who did versus did not have a history of suicidal thoughts and behaviors were largest for the Suicide IAT and Death IAT, and were nonsignificant for the Self-Injury IAT. Several aspects of these findings warrant additional comment.

There has been increasing scientific and clinical interest in recent years in the use of measures of implicit cognition to learn about cognitive processes associated with suicidal and NSSI. Prior

studies have shown that scores on self-harm IATs not only distinguish between those engaging versus not engaging in self-harm but also predict future self-harm above and beyond other established predictors. However, small sample sizes among these studies have limited our ability to more precisely clarify the association between implicit associations and self-harming behaviors. The current study utilized a large-scale web-based platform as a means to replicate prior laboratory and hospital results and to recruit the numbers of participants necessary to address several novel questions.

First, we found that individuals with a lifetime history of NSSI and suicide attempt associate self-harm more strongly with the self than do those without a history of self-injury and attempt, respectively. These results not only replicate prior laboratory and hospital studies demonstrating similar group differences for NSSI (Nock & Banaji, 2007a) and suicide attempt (Nock & Banaji, 2007b; Nock et al., 2010), but also indicate the feasibility of disseminating the self-harm IAT across a variety of settings, given that it can be successfully self-administered and remains valid even for less controlled environments (e.g., online). IAT versions utilizing stimuli more specific to the target behaviors assessed produced the strongest effects, such that the largest effect size in the NSSI versus control comparison was observed on the Self-Injury IAT and the largest effect size in the suicide attempt versus control comparison was observed on the Suicide IAT. The fact that specific words related to suicide methods (Suicide IAT) produced stronger effects for lifetime suicide attempt than general words related to death (Death IAT) suggests that self-harm IATs are strongest when assessing behavioral constructs precisely. When using the IAT to predict the presence of past self-harm, its false-negative rate was low but its sensitivity to detecting a history of self-harm performed only moderately well, especially for suicide attempts.

Second, these results raise the possibility that implicit associations may not be stable traits but rather are more malleable or state-dependent, decreasing with time since the most recent self-harm engagement. Performance on the Self-Injury IAT achieved considerable temporal resolution for NSSI, distinguishing between participants who self-injured during the past week and those who self-injured during the past year (but not the past week). Whereas longitudinal data are necessary to draw firm conclusions, these findings provide some preliminary indication that the IAT may serve as a way to assess imminence of current risk, given its capability of tracking one's recent history of self-harm. Nonetheless, we reiterate that caution is needed, given that a similar pattern of results could be observed even if implicit cognitions were more stable in nature (e.g., those with stronger implicit cognitions tend to act on such cognitions more often and thus have typically experienced self-harm more recently). Third, we found that self-harm-related implicit associations typically were stronger among those with more frequent and severe histories of self-harm. The Self-Injury IAT was most strongly correlated with lifetime frequency of NSSI, and although correlations for suicide attempt frequency were weaker than NSSI, the Death IAT demonstrated the strongest implicit associations for frequency of suicide attempt history (i.e., lifetime frequency for both total and medically serious attempts, though results for medically serious attempts did not hold across both the initial and replication samples). We also found that implicit associations with self-harm were associated specifically

with suicide attempt (i.e., not explained by suicide ideation) and were strongest on the Suicide IAT, indicating the IAT's ability to differentiate those with particularly severe suicide histories.

Lastly, we found evidence of specificity between different IAT versions and types of self-harm. The Self-Injury IAT most strongly differentiated those with versus without a history of NSSI (excluding those with a history of suicide ideation or attempt), and the Suicide and Death IATs most strongly differentiated those with versus without a history of suicide ideation or attempt (excluding those with NSSI history). Those with a history of NSSI had consistently elevated scores on the Self-Injury IAT but not the Death or Suicide IATs (though results were significant for the Suicide IAT in the replication sample), suggesting that implicit associations among self-injurers are associated with NSSI, specifically, and not self-harm, generally. Similarly, those with a history of suicidal thoughts and behaviors were consistently elevated on the Suicide IAT and Death IAT but not the Self-Injury IAT, indicating the Suicide/Death IAT's ability to capture suicide history independent from NSSI history. Future research is needed to gain even greater clarification of the IAT's precision.

### Clinical Implications

These findings have important implications for both researchers and clinicians. Prior laboratory and hospital studies have established the IAT's ability to uniquely predict future self-harm behaviors above and beyond known risk factors, as well as clinician and patient self-report. Many robust risk factors for suicidal behavior, such as depression, have high sensitivity but poor specificity for predicting future behaviors (i.e., most suicide attempters have a history of depression but very few depressed individuals go on to attempt suicide). This study and others have shown that the self-harm IAT not only has good sensitivity but even better specificity (Nock et al., 2010), suggesting the IAT has unique predictive value not captured by other known risk factors. The fact that this study replicated prior main findings and found that self-harm-related implicit associations may be temporally malleable indicates the feasibility of validly administering the IAT across diverse, less controlled settings, and the IAT's ability to track current self-harm risk in a way that more fixed-risk factors cannot. As such, the IAT may be a tool that clinicians and patients can use remotely on an ongoing basis as a means to gather more information about one's risk of NSSI or suicide attempt. Examples include follow-up assessments completed by patients at home between outpatient visits, or tracking risk for the increasing number of patients using eHealth interventions such as computer-based cognitive behavior therapies. Given its sensitivity to temporal patterns, the IAT also may serve a critical function to assess ongoing risk for patients following hospital discharge (e.g., following a suicide attempt), given research showing that individuals are at especially heightened risk during the first year following a suicide attempt (Hawton & Fagg, 1988).

To maximize the IAT's clinical utility, these results suggest it may be most useful for researchers and clinicians to tailor IAT stimuli to the specific target behaviors of interest. Specifically, there is a clinical advantage to having two separate IATs: (a) a Self-Injury IAT to assess potential presence and risk of NSSI, and (b) a Death/Suicide IAT to assess potential presence and risk of suicide attempts. The Self-Injury IAT produced strong and specific effects when assessing NSSI history, and the Death/Suicide IATs may be a preferred means to assess

suicide attempt, given its stronger ability to differentiate suicide attempt from either suicide ideation or comorbid NSSI. For example, the Suicide IAT could be used in the emergency department prior to making decisions regarding discharge with high-risk suicidal patients, especially considering the IAT's resistance to faking (Greenwald et al., 2009). An interesting future direction would be to compare picture versus word stimuli for suicide-related stimuli to see whether a similar pattern emerges and/or effects are even stronger (e.g., suicide images vs. words). It may even be possible to personalize self-harm IATs to the level of individual stimuli, such as using pictures of pills for individuals with a history of overdose. Also, results showing that past suicide attempters had stronger associations with life than death/suicide suggests it will be clinically necessary to determine cutpoints based on relative, rather than absolute, associations with self-harm.

Taken together, these results indicate the self-harm IAT's utility in measuring implicit associations outside of conscious control and the promise that this tool holds for assessing suicide risk with greater precision. It is important, however, to view implicit cognition about self-harm not as a panacea for assessing suicide risk but rather as one measure that, in conjunction with other important factors or markers, can increase the accuracy of risk assessments. Future work is needed to develop methods to combine available data (e.g., IAT and clinician- and patient-reports) in a clinically useful way. When such progress is achieved, the IAT represents a well-validated measure to include in the battery.

### Limitations and Conclusion

Several limitations of the current study should be noted. First, these data were collected cross-sectionally; therefore, self-harm associations could only be shown to predict group status (i.e., self-reported past self-harm) rather than future behaviors. Although prior studies with smaller samples have demonstrated the IAT's ability to predict future self-harm above and beyond other well-known predictors (e.g., past suicide attempt; Nock et al., 2010), large-scale longitudinal studies that assess participants at multiple time points are needed to make temporal predictions between self-harm-related implicit cognition and behavior. Specifically, such studies could clarify whether frequency of self-harm leads to increases in the strength of implicit cognition, whether stronger implicit cognitions about self-harm lead to greater engagement in or more severe self-harm behaviors, and the extent to which implicit associations significantly predict future self-harm beyond already established predictors.

Second, although these results are consistent with prior laboratory studies using community samples, it is important to note that our study sample was not representative of the general population and contained higher rates of self-harm than is commonly found in non-clinical populations (e.g., 58% participants had lifetime NSSI vs. 13% reported prevalence among young adults; Swannell et al., 2014). Participants with a history of self-harm, or psychopathology in general, may have self-selected into the study (e.g., more likely to visit the site, complete the study tasks after consent, or choose to complete the self-harm vs. other non-self-harm-related IATs on the site) due to personal experience or interest in self-harm. Therefore, interpretation of these results should be made with caution and may be limited to participants with a history of self-harm behavior or those with more prior exposure to self-harm-related concepts or stimuli. Even so, these findings support the IAT's potential utility, given that those with a

history of self-harm are at heightened risk for future engagement and thus are the very individuals for which better assessment tools are especially needed.

Third, participants were placed into past self-harm groups based on self-report to, in some cases, single items (e.g., lifetime suicide attempt). Although measures utilizing single-item questions to assess history of self-harm have been shown to be reliable (Nock et al., 2007) and the anonymous nature of the self-report makes it less likely to be subject to impression management, it is possible that participants would be categorized differently based on a more comprehensive interview. Mitigating this concern, a recent study directly testing this issue found that only 11% of participants who endorsed a single-item measure asking about a past suicide attempt engaged in behaviors that, as determined by coders, would not be considered an actual suicide attempt (Millner, Lee, & Nock, 2015).

Fourth, given that participation in the study was anonymous, it is possible that some participants took the IAT multiple times; however, the very large sample size minimizes the chances that such repeat participants significantly affect these results.

Fifth, the current study focused primarily on *behavioral* outcomes of self-harm (i.e., suicide attempt and engagement in NSSI). Although the ability to examine self-harming behaviors, as opposed to using proxies for such behaviors only (e.g., suicidal thoughts), is a strength of this work, future studies could examine other types of self-harming thoughts and behaviors—such as suicide plans and intent—to elucidate their respective implicit associations with different IAT variants.

Although the current study provides strong evidence for the validity and potential clinical utility of the IAT, further research is needed to develop clinical assessment tools that are able to integrate risk factors and behavioral markers as a means to reliably identify individuals at risk. Given that clinical judgment of suicide risk is fraught with practical challenges of integrating knowledge of numerous risk factors and individual circumstances, actuarial approaches to suicide risk assessment that include implicit cognition measures provide tremendous promise in the critical goal of helping those at risk of hurting themselves.

## References

- Asarnow, J. R., Porta, G., Spirito, A., Emslie, G., Clarke, G., Wagner, K. D., . . . Brent, D. A. (2011). Suicide attempts and nonsuicidal self-injury in the treatment of resistant depression in adolescents: Findings from the TORDIA study. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*, 772–781. <http://dx.doi.org/10.1016/j.jaac.2011.04.003>
- Busch, K. A., Fawcett, J., & Jacobs, D. G. (2003). Clinical correlates of inpatient suicide. *The Journal of Clinical Psychiatry, 64*, 14–19. <http://dx.doi.org/10.4088/JCP.v64n0105>
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*, 365–376. <http://dx.doi.org/10.1038/nrn3475>
- Centers for Disease Control and Prevention. (2013). *Web-based Injury Statistics Query and Reporting System (WISQARS)*. Retrieved from <http://www.cdc.gov/ncipc/wisqars>
- Cha, C. B., Augenstein, T. M., Frost, K. H., Gallagher, K., D'Angelo, E. J., & Nock, M. K. (2016). Using implicit and explicit measures to predict nonsuicidal self-injury among adolescent inpatients. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*, 62–68. <http://dx.doi.org/10.1016/j.jaac.2015.10.008>
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition: research: Their meaning and use. *Annual Review of Psychology, 54*, 297–327. <http://dx.doi.org/10.1146/annurev.psych.54.101601.145225>
- Fischer, G., Ameis, N., Parzer, P., Plener, P. L., Groschwitz, R., Vonderlin, E., . . . Kaess, M. (2014). The German version of the self-injurious thoughts and behaviors interview (SITBI-G): A tool to assess non-suicidal self-injury and suicidal behavior disorder. *BMC Psychiatry, 14*, 265. <http://dx.doi.org/10.1186/s12888-014-0265-0>
- García-Nieto, R., Blasco-Fontecilla, H., Yepes, M. P., & Baca-García, E. (2013). Translation and validation of the Self-Injurious Thoughts and Behaviors Interview in a Spanish population with suicidal behaviour. *Revista de Psiquiatría y Salud Mental, 6*, 101–108. <http://dx.doi.org/10.1016/j.rpsm.2012.07.001>
- Glenn, C. R., Kleiman, E. M., Cha, C. B., Nock, M. K., & Prinstein, M. J. (2016). Implicit cognition about self-injury predicts actual self-injurious behavior: Results from a longitudinal study of adolescents. *Journal of Child Psychology and Psychiatry, 57*, 805–813.
- Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology, 108*, 553–561. <http://dx.doi.org/10.1037/pspa0000016>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464–1480. <http://dx.doi.org/10.1037/0022-3514.74.6.1464>
- Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48*, 85–93.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197–216. <http://dx.doi.org/10.1037/0022-3514.85.2.197>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41. <http://dx.doi.org/10.1037/a0015575>
- Han, H. A., Czellar, S., Olson, M. A., & Fazio, R. H. (2010). Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology, 46*, 286–298. <http://dx.doi.org/10.1016/j.jesp.2009.11.011>
- Hawton, K., & Fagg, J. (1988). Suicide, and other causes of death, following attempted suicide. *The British Journal of Psychiatry, 152*, 359–366.
- Joy-Gaba, J. A., & Nosek, B. A. (2010). The surprisingly limited malleability of implicit racial evaluations. *Social Psychology, 41*, 137–146. <http://dx.doi.org/10.1027/1864-9335/a000020>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*, 764–766. <http://dx.doi.org/10.1016/j.jesp.2013.03.013>
- Millner, A. J., Lee, M. D., & Nock, M. K. (2015). Single-item measurement of suicidal behaviors: Validity and consequences of misclassification. *PLoS ONE, 10*, e0141606. <http://dx.doi.org/10.1371/journal.pone.0141606>
- Nock, M. K., & Banaji, M. R. (2007a). Assessment of self-injurious thoughts using a behavioral test. *The American Journal of Psychiatry, 164*, 820–823. <http://dx.doi.org/10.1176/ajp.2007.164.5.820>
- Nock, M. K., & Banaji, M. R. (2007b). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of Consulting and Clinical Psychology, 75*, 707–715. <http://dx.doi.org/10.1037/0022-006X.75.5.707>
- Nock, M. K., Borges, G., & Ono, Y. (2012). *Suicide: Global perspectives from the WHO World Mental Health Surveys*. New York, NY: Cambridge University Press.

- 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>
- Nock, M. K., Park, J. M., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science, 21*, 511–517. <http://dx.doi.org/10.1177/0956797610364762>
- Nolan, S. A., Heinzen, T. E., & Weathersby, R. (2008). *Statistics for the Behavioral Sciences Study Guide & SPSS Manual*. New York, NY: Worth.
- Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., . . . Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*, 36–88. <http://dx.doi.org/10.1080/10463280701489053>
- Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*, 615–631. <http://dx.doi.org/10.1177/1745691612459058>
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology, 108*, 562–571.
- Randall, J. R., Rowe, B. H., Dong, K. A., Nock, M. K., & Colman, I. (2013). Assessment of self-harm risk using implicit thoughts. *Psychological Assessment, 25*, 714–721. <http://dx.doi.org/10.1037/a0032391>
- Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Science, 1*, 213–220.
- Swannell, S. V., Martin, G. E., Page, A., Hasking, P., & St John, N. J. (2014). Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression. *Suicide and Life-Threatening Behavior, 44*, 273–303. <http://dx.doi.org/10.1111/sltb.12070>
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.
- Teachman, B. A., Marker, C. D., & Smith-Janik, S. B. (2008). Automatic associations and panic disorder: Trajectories of change over the course of treatment. *Journal of Consulting and Clinical Psychology, 76*, 988–1002. <http://dx.doi.org/10.1037/a0013113>
- Teachman, B. A., & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112*, 100–109. <http://dx.doi.org/10.1037/0021-843X.112.1.100>
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*, 632–638. <http://dx.doi.org/10.1177/1745691612463078>
- Wilkinson, P., Kelvin, R., Roberts, C., Dubicka, B., & Goodyer, I. (2011). Clinical and psychosocial predictors of suicide attempts and nonsuicidal self-injury in the Adolescent Depression Antidepressants and Psychotherapy Trial (ADAPT). *The American Journal of Psychiatry, 168*, 495–501. <http://dx.doi.org/10.1176/appi.ajp.2010.10050718>
- Wilson, T. D. (2009). Know thyself. *Perspectives on Psychological Science, 4*, 384–389. <http://dx.doi.org/10.1111/j.1745-6924.2009.01143.x>

**Appendix**  
**Categories and Stimuli for Implicit Association Measures**

Category	Stimuli
Me	I, mine, myself, self
Not me	they, them, their, other
Death	suicide, die, deceased, dead
Suicide	gunshot, hanging, overdose, cutting
Life	alive, thrive, breathing, living
Cutting	
No Cutting	

*Note.* See the online article for the color version of this figure.

Received June 17, 2016  
 Revision received August 13, 2016  
 Accepted September 20, 2016 ■