J. Behav. Ther. & Exp. Psychiat. 52 (2016) 17–28

Contents lists available at ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep

Characterizing implicit mental health associations across clinical domains

Alexandra J. Werntz [a,*], Shari A. Steinman [b], Jeffrey J. Glenn [a], Matthew K. Nock [c], Bethany A. Teachman [a]

[a] Department of Psychology, University of Virginia, PO Box 400400, Charlottesville, VA, 22904, United States
[b] New York State Psychiatric Institute, 1051 Riverside Drive, Unit 69, New York, NY, 10032, United States
[c] Psychology Department, Harvard University, William James Hall, 33 Kirkland St., Cambridge, MA, 02138, United States

ARTICLE INFO

Article history:
Received 13 April 2015
Received in revised form
19 January 2016
Accepted 24 February 2016
Available online 3 March 2016

Keywords:
Implicit associations
Alcohol use/abuse
Depression
Eating disorders
Anxiety
Internet research

ABSTRACT

Background and objectives: Implicit associations are relatively uncontrollable associations between concepts in memory. The current investigation focuses on implicit associations in four mental health domains (alcohol use, anxiety, depression, and eating disorders) and how these implicit associations: a) relate to explicit associations and b) self-reported clinical symptoms within the same domains, and c) vary based on demographic characteristics (age, gender, race, ethnicity, and education).
Methods: Participants (volunteers over age 18 to a research website) completed implicit association (Implicit Association Tests), explicit association (self + psychopathology or attitudes toward food, using semantic differential items), and symptom measures at the Project Implicit Mental Health website tied to: alcohol use ($N = 12{,}387$), anxiety ($N = 21{,}304$), depression ($N = 24{,}126$), or eating disorders ($N = 10{,}115$).
Results: Within each domain, implicit associations showed small to moderate associations with explicit associations and symptoms, and predicted self-reported symptoms beyond explicit associations. In general, implicit association strength varied little by race and ethnicity, but showed small ties to age, gender, and education.
Limitations: This research was conducted on a public research and education website, where participants could take more than one of the studies.
Conclusions: Among a large and diverse sample, implicit associations in the four domains are congruent with explicit associations and self-reported symptoms, and also add to our prediction of self-reported symptoms over and above explicit associations, pointing to the potential future clinical utility and validity of using implicit association measures with diverse populations.

1. Introduction

If asked to indicate the extent to which you think of yourself as a sad person, you may think about the last time you were sad or about what it means to be sad. Such assessments are *explicit* in that consciously controlled processing informs your reports. To measure less controlled, more spontaneous assessments that may reside outside conscious awareness, researchers have developed tools for measuring *implicit* cognition to capture ways that past experiences (that may not be explicitly remembered) influence performance or behaviors (for a comprehensive definition, see Greenwald & Banaji, 1995). *Implicit associations* are a key component of implicit cognition and reflect the relative strength of associations between concepts held in memory (see Lane, Banaji, Nosek, & Greenwald, 2007; Nosek, Greenwald, & Banaji, 2007). Importantly, implicit associations are relatively difficult to consciously control, in contrast to *explicit associations*, which require conscious, more deliberate reflection of one's attitudes towards the concepts. In the current study, implicit and explicit associations across four mental health-related domains (alcohol use, anxiety, depression, and eating disorders) are examined.

Implicit and explicit associations are not perfectly correlated

* Corresponding author.
E-mail addresses: ajw3x@virginia.edu (A.J. Werntz), haxelsh@nyspi.columbia.edu (S.A. Steinman), jglenn@virginia.edu (J.J. Glenn), nock@wjh.harvard.edu (M.K. Nock), bteachman@virginia.edu (B.A. Teachman).

http://dx.doi.org/10.1016/j.jbtep.2016.02.004
0005-7916/

and the magnitude of correlation strength between the two can depend on many moderators (e.g., social desirability of the content area; see Nosek, 2005). Although there is a rich literature evaluating implicit–explicit relationships in the social cognition field, less research has been devoted to implicit–explicit relationships specifically in the clinical domain (though see Roefs et al., 2011; for a useful review). Moreover, most studies in the clinical field have relied on small samples, making it difficult to evaluate the influence of individual differences on the expression of implicit associations (e.g., within small samples there may be minimal representation of a given ethnic group, limiting the opportunity to conduct high powered tests of ethnic group effects, and one cannot assume random assignment of other factors across participants so it may be difficult to disentangle the influence of co-occurring participant attributes, such as a discrepant gender ratio across different ethnic groups). Thus, using a large sample, the present study describes associations among implicit and explicit associations and self-reported symptoms in four highly prevalent mental health domains. We also investigate how implicit and explicit associations and symptom measures vary as a function of demographic variables (e.g., age, race).

Launched in September 2011, Project Implicit Mental Health (PIMH; a sister-site to Project Implicit, a popular implicit social cognition research website launched in 1998) is a demonstration and research website that allows individuals to learn more about their implicit biases associated with mental illness, such as anxiety and depression (To take a demonstration test, visit www.ImplicitMentalHealth.com.). The size and relative diversity of visitors to the public website provides a unique opportunity to characterize these implicit associations across multiple mental health domains and evaluate variability in implicit associations relating to differences in demographic characteristics.

The Implicit Association Test (IAT) was selected to operationalize implicit mental health associations because of its strong psychometric properties. The IAT has been found to correlate with mental health symptoms, such as specific phobias (e.g., Teachman, Gregg, & Woody, 2001), social anxiety (e.g., de Jong, 2002), alcohol use (e.g., Lindgren, Neighbors, et al., 2013), and even suicidality (e.g., Nock & Banaji, 2007). Further, the IAT uniquely predicts disorder-related behavior (e.g., an alcohol-arousing IAT predicts alcohol use; Houben & Wiers, 2008) and treatment outcomes (e.g., change in panic disorder symptoms; Teachman, Marker, & Smith-Janik, 2008). In a review of psychopathology research involving implicit measures, Roefs et al. (2011) conclude that although results across disorders are mixed, implicit association measures like the IAT provide insight into psychopathology not captured by explicit measures. However, the majority of these studies examined single problem areas only and often used relatively homogenous, small samples in the lab, limiting our ability to make more reliable estimates of implicit–explicit relationships and determine how demographic variables moderate the implicit evaluations.

Theoretical support for the use of implicit measures in psychopathology research follows from numerous models highlighting how various features of automatic processing biases are key components of mental health disorders (for review on automaticity in anxiety and depressive disorders, see Teachman, Joormann, Steinman, & Gotlib, 2012). For instance, in the anxiety domain, Beck and Clark's (1997) model of anxiety pathology posits that information processing in anxious individuals is a combination of automatic and strategic processes. In the mostly automatic stage of threat stimulus recognition, individuals process the threat outside of conscious awareness before using more elaborated or strategic processing to act on the threat. Similarly, Beevers' (2005) dual-process model for depression suggests that individuals who are unable to strategically correct for negative automatic, associative biases may be more vulnerable to depressive disorders. In the alcohol literature, Wiers, Gladwin, Hofmann, Salemink, and Ridderinkhof (2013) suggest that focusing on iterative reprocessing and interactions between relatively automatic and controlled processing will lead to a more nuanced understanding of addiction and psychopathology. Finally, Vitousek and Hollon (1990) hypothesize that automatic processing biases arise among individuals with eating disorders given the association between the self and unhealthy weight-related beliefs. The role of automatic processing is central to a broad range of psychopathology theories, suggesting implicit associations may be transdiagnostic maintaining or vulnerability factors. Therefore, understanding how implicit–explicit relationships vary across disorder domains is critical.

1.1. Overview

This is the first study that examines the relationship among implicit and explicit associations and self-reported symptoms across multiple disorder domains in large and relatively diverse samples (though see important work by Glashouwer & de Jong, 2010, examining implicit anxious and depressed associations in a large Dutch sample recruited from different health settings). Our primary aim is to describe these relationships and examine how implicit associations vary based on demographic variables, so the approach is mainly descriptive, rather than conducting formal tests of theory. Four clinical domains were selected because of their high prevalence and the prominence of uncontrollable processing biases in the models for each of these problem areas: alcohol use (hazardous drinking), anxiety, depression, and eating disorder symptoms. In each case, self-reported symptoms were measured, as opposed to diagnoses, in part for logistical reasons given the online data collection method, but also to better understand implicit evaluations across persons with a broad range of symptom severity. We also included measures of explicit associations to allow for a more direct comparison with the implicit measures because both capture simple associations based on analogous relative construct comparisons. We recognize that the selected mental health domains encompass some heterogeneity (e.g., panic disorder and social anxiety disorder differ in meaningful ways, despite both being part of the anxiety domain). However, we focus on the broader domains as opposed to diagnostic categories for this initial evaluation, consistent with a Research Domain Criteria (RDoC; Insel et al., 2010) perspective, and because of our primary interest in implicit associations that we expect to be fundamental across diagnostic categories.

In the present study, it is expected that positive correlations will emerge among implicit associations, explicit associations, and self-reported symptom measures across domains; however, given prior meta-analyses showing a small mean effect size of the implicit–explicit relationships (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Roefs et al., 2011), these correlations are expected to be relatively modest in magnitude. In addition to examining correlations across the measures, we will also present how implicit, explicit, and symptom measures vary by demographic variables, and whether implicit associations are uniquely predictive of symptoms (above and beyond explicit associations). Establishing incremental predictive validity is important for determining the ultimate clinical utility of these measures.

2. Methods

2.1. Participants

A total of 67,932 participants consented and completed at least one task among the Alcohol (12,387), Anxiety (21,304), Depression

Table 1
Participant characteristics.

Mental health study:	Alcohol (n = 12,387)	Anxiety (n = 21,304)	Depression (n = 24,126)	Eating (n = 10,115)
Mean (*SD*) age in years	28.3 (11.5)	28.5 (11.6)	28.6 (11.8)	27.6 (11.4)
Sex (% female)	58.3	71.1	68.2	76.6
Ethnicity (%)				
Non-Hispanic	79.8	78.8	78.3	78.0
Hispanic	9.9	10.6	10.4	10.5
Unknown/did not report	10.3	10.7	11.3	11.5
Race (%)				
Caucasian	72.6	69.4	67.1	68.4
Asian	4.9	7.3	8.6	7.1
African American	6.2	6.1	6.5	6.5
Multiracial	4.8	5.2	5.5	5.1
Other	7.2	8.0	8.4	8.3
Did not report	4.2	4.0	3.9	4.6
Education (%)				
Graduate degree	16.0	16.2	17.4	15.9
Bachelor's degree	26.2	24.7	24.7	24.7
Some college	42.4	42.8	39.9	41.0
High school	10.2	10.8	11.8	12.0
Less than high school	2.7	3.4	3.7	3.9
Did not report	2.4	2.3	2.5	2.5

(24,126), and Eating (10,115) studies on the PIMH website between September 2011 and May 2014 (www.ImplicitMentalHealth.com). Visitors could select which studies they wished to complete at the site; thus, disorder topics were not assigned randomly. In addition to the four disorder domains examined presently, individuals could choose to complete studies on mental health treatment, self-esteem, stigma toward persons with mental illness, and suicide and self-harm (latter described in Glenn et al., 2016). Research on this site was approved by affiliated Institutional Review Boards. PIMH samples are more diverse than typical collegiate samples; however, they are not representative of any one, defined population. Participants were citizens of 173 different countries; 71.2% were United States citizens. See Table 1 for participant characteristics for each mental health study. Comparing our participant characteristics to US Census data, we see that our sample is modestly more diverse than the general US population in terms of race (i.e., 2013 US Census data indicated that 77.7% of the US was white, 13.2% Black or African American, 2.4% more than one race), but not in terms of ethnicity (17.1% of the US reported being Latino or Hispanic in 2013; United States Census Bureau, 2015b). Our sample is also more educated than the general US population (United States Census Bureau, 2015a, data indicates approximately 10% of adults 18 and older have a Master's, professional, or Doctoral degree; 19% have a Bachelor's degree; and 30% have a high school degree).

We do not have formal data on how participants are directed to the site; however, we have a good idea of how most volunteers hear about the site. Project Implicit (www.projectimplicit.net), a very popular research and education website, has a link to PIMH on its site, which likely contributes a large portion of our site traffic. We also mention the PIMH site in talks and news coverage, and it is likely that some visitors come to the site through 'surfing' the Internet, or a class assignment.

2.2. Measures

2.2.1. Demographics

[1] Participants completed a brief demographics form, in which they were asked to report their age, gender, ethnicity, race, and educational background.

2.2.2. Implicit associations

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) compares reaction times when classifying stimuli into superordinate categories that have been paired to match vs. contradict disorder-linked automatic associations to measure the relative association strengths between constructs. Both traditional IATs, which require an explicitly labeled comparison category to contrast with the target construct, and Brief IATs (BIAT; Sriram & Greenwald, 2009), which are conceptually similar to the IAT but do not require an explicitly labeled comparison category, were used, depending on the study. See Table 2a for all category labels and stimuli. Three of the four domains (alcohol use, anxiety, and depression) measure self-concept implicit associations (i.e., self + a disorder characteristic), whereas the eating task measures individuals' food-related implicit associations.

The alcohol BIAT[2] was designed to assess drinking associations with the self, based on prior research showing that implicit association strength between "drinker" + "me" predicts alcohol problems and consumption better than other alcohol IAT variants (Lindgren, Neighbors, et al., 2013). The task consisted of two classification conditions, one in which the category labels "drinker" and "me" appeared together at the top of the screen, and one in which "abstaining" and "me" appeared together. Words related to "drinker," "abstaining," "me," or "others" were then displayed one at a time in the center of the screen. Note, the "others" stimuli are background stimuli to contrast with the "me" stimuli, but there is no labeled "others" category in the BIAT design. Participants were asked to press the "i" key if the word belonged to one of the displayed word categories and the "e" key if the word did not fit in the

[1] Other demographic items not central to this study were included in the form, but not reported here. Income or socioeconomic status was not collected. The full list of demographic items is available from the first author.

[2] Two changes were made midway through the first year of data collection. In the Alcohol study, the classification categories in the alcohol BIAT were revised from "alcohol-irresistible" to "drinker-me" given evidence for the latter's stronger psychometric properties (Lindgren, Neighbors, et al., 2013). For the sake of consistency, only data from participants who took the drinker-me alcohol BIAT (n = 9424) were included in these analyses. In the Anxiety study, the anxiety symptom questionnaire (DASS-21 anxiety subscale) was temporarily exchanged for another symptom questionnaire (DASS-21 stress subscale). The DASS-21 anxiety subscale is currently being administered and is the scale reported here (n = 15,516 in the current analysis).

Table 2a
Labels and stimuli for implicit association measures.

Label	Stimuli
Me	me, self, I, my
Others	not me, other, them, they
(Me and Others categories were used for Alcohol, Anxiety, & Depression studies)	
Alcohol	
Drinker	alcohol, drunk, intoxicated, drinking
Abstaining	abstaining, abstain, sober, refrain
Anxiety	
Anxious	anxious, panicked, scared, frightened
Calm	calm, relaxed, serene, tranquil
Depression	
Sad	sad, miserable, depressed, gloomy
Happy	happy, joyful, content, cheerful
Eating	
High-fat food	fries, cake, candy, chocolate
Low-fat food	salad, carrots, fruit, cucumber
Shameful	disgraceful, bad, embarrassing, shameful
Acceptable	suitable, good, appropriate, acceptable

Table 2b
Explicit association questions and response options.

Domain	Questions and response options
Alcohol	To what extent do you think of yourself as drinking or abstaining?
	To what extent do you think of others as drinking or abstaining?
	Completely as drinking, mostly as drinking, moderately as drinking, slightly as drinking, slightly as drinking, neither as drinking or abstaining, slightly as abstaining, moderately as abstaining, mostly as abstaining, completely as abstaining
Anxiety	To what extent do you think of yourself as anxious or calm?
	To what extent do you think of others as anxious or calm?
	Extremely anxious, very anxious, moderately anxious, neither anxious nor calm, slightly calm, moderately calm, very calm, extremely calm
Depression	To what extent do you think of yourself as happy or sad?
	To what extent do you think of others as happy or sad?
	Extremely sad, very sad, moderately sad, slightly sad, neither happy nor sad, slightly happy, moderately happy, very happy, extremely happy
Eating	To what extent do you think of high-fat foods as acceptable or shameful?
	To what extent do you think of low-fat foods as acceptable or shameful?
	Extremely acceptable, very acceptable, moderately acceptable, slightly acceptable, neither acceptable nor shameful, slightly shameful, moderately shameful, very shameful, extremely shameful

reduced preference for high-fat foods (Stoner, Fedoroff, Andersen, & Rolls, 1996), perhaps because high-fat foods are seen as shameful. Prior studies have shown that anorexic patients show greater negative associations with high-caloric food than control participants (de Jong & Veenstra, 2007). In the current study, in one condition, "high-fat food" and "shameful" category labels were paired together on one side of the screen while "low-fat food" and "acceptable" category labels were paired together on the other side of the screen. This differs from the BIAT described above where only 3 of the 4 categories are explicitly labeled. Words related to "high-fat food," "low-fat food," "shameful," or "acceptable" were then displayed one at a time in the center of the screen. Participants were asked to press the "i" key if the word belonged to the word categories displayed on the right of the screen and the "e" key if the word belonged to the word categories displayed on the left of the screen. The categories were then switched so that "high-fat food" was paired with "acceptable" and "low-fat food" with "shameful," and the classification task was repeated.

displayed categories. Six blocks of 16 trials each were presented, alternating between the "abstaining" + "me" and "drinker" + "me" conditions. The first two blocks were practice trials used to familiarize the participant with the test; the last four blocks were critical trials from which the scores were derived. Order of the "abstaining" + "me" condition and the "drinker" + "me" condition were randomized, such that individuals received two blocks of one condition and then two blocks of the other. To assess the relative strength of drinking identity associations, time to classify stimuli in the two conditions was compared. Faster responses when "drinker" was paired with "me," relative to when "drinker" was paired with "not me," indicated a higher implicit drinking identity.

The anxiety BIAT was designed to assess anxious associations with the self, based on prior research examining the self-concept of anxiety (Egloff & Schmukle, 2002). The depression BIAT was designed to assess negative mood associations with the self, based on prior research showing that depressed individuals have more dysfunctional and negatively biased views of themselves than non-depressed individuals (e.g., Hollon, Kendall, & Lumry, 1986). The implicit anxiety and depression tasks were similar to the BIAT for alcohol except that "drinking" and "abstaining" category labels were replaced with "anxious" and "calm" for the anxiety study and "sad" and "happy" for the depression task. In all three tasks, the same words for "me" and "others" were used.

Unlike the other three BIATs, the eating IAT followed a typical IAT format, consisting of two critical classification conditions, and was designed to assess the acceptability of high-fat foods, based on prior research in which anorexic individuals reported having a

For the eating IAT, seven total blocks were presented, consisting of three practice and four critical blocks. To familiarize participants with the task, the first two blocks consisted of 20 practice trials each in which only two categories from a single dimension (e.g., "high-fat food" and "low-fat food") appeared on the screen. Participants then completed blocks 3 (20 trials) and 4 (40 trials) that were identical critical blocks for one combined category pairing classification condition (e.g., high-fat food paired with acceptable and low-fat food paired with shameful). Block 5 was a practice block of 40 trials in which "high-fat food"/"low-fat food" remained in the same position but the other contrasting categories ("shameful"/"acceptable") switched sides of the screen. Finally, blocks 6 and 7 followed the same format as blocks 3 and 4, but the category pairings were reversed (e.g., high-fat food was now paired with shameful and low-fat food paired with acceptable). Again, order of the critical category pairing conditions was randomized.

Across tasks, higher implicit scores reflect relatively more unhealthy associations with the self (e.g., self as a drinker, or self as sad), and higher eating IAT scores reflect stronger high-fat food as shameful associations. Internal consistencies were calculated by correlating reaction times of alternating couplets within the implicit measures (following Nosek & Smyth, 2007). The alcohol BIAT ($r = .58$, $p < .001$), anxiety BIAT ($r = .60$, $p < .001$), depression BIAT ($r = .62$, $p < .001$), and eating IAT ($r = .76$, $p < .001$) exhibited modest to strong internal reliability.

2.2.3. *Explicit associations*

To provide explicit associations that mirror the IAT and BIAT's relative structure, participants were asked for ratings along dimensions that matched the implicit tasks' categories using semantic differential items, anchored by two opposing adjectives (i.e., semantic differential items; Greenwald et al., 1998; Ranganath & Nosek, 2008). See Table 2b for questions and response options.

For the alcohol, anxiety, and depression tasks, explicit associations were calculated by subtracting the response to the question regarding ‘others’ from the response to the question regarding the ‘self.’ Higher scores reflect greater self-as-unhealthy explicit associations, matching the directional valence of the implicit association score. Similarly, for the eating task, an explicit association was calculated by subtracting the question regarding ‘high-fat foods’ from the question regarding ‘low-fat foods,’ so that higher scores reflected a greater high-fat food-as-shameful explicit association. These relative measures of explicit associations have been used in previous implicit/explicit association research across different clinical domains (e.g., people with mental illness vs. being on welfare, Peris, Teachman, & Nosek, 2008; fat vs. thin people, Teachman, Gapinski, Brownell, Rawlins, & Jeyaram, 2003 Study 2a). Additionally, relative measures of explicit self-anxious and self-depressed associations correlate with corresponding symptom measures (Glashouwer et al., 2010; Glashouwer, de Jong, & Penninx, 2012).

Note that these explicit association measures were included as simple explicit parallels to the implicit measure, given both assess the same relative associations between constructs. If the explicit association measures were excluded, the study would only examine a single-target measure (symptoms) and its correlation with an implicit, dual-target relative measure (implicit association tests). As a result, implicit and explicit differences would be confounded with single- and dual-target differences. Further, there would not be a clear way to assess incremental validity of the implicit measures in predicting symptoms.

2.2.4. Symptom measures

2.2.4.1. Alcohol. The Alcohol Use Disorders Identification Test (AUDIT; Babor, Higgins-Biddle, Saunders, & Monteiro, 1992) is a 10-item questionnaire that measures alcohol consumption and is widely used to identify individuals who may exhibit hazardous drinking patterns (Allen, Litten, Fertig, & Babor, 1997; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). Higher scores reflect greater risk of alcohol-related disorders (e.g., dependence). This measure has been shown to have acceptable test-retest reliability and internal consistency (Reinert & Allen, 2002). Cronbach's $\alpha = .85$ in the current study.

2.2.4.2. Anxiety. The Depression, Anxiety, and Stress Scale-21 item version (DASS-21; Lovibond & Lovibond, 1995) 7-item anxiety subscale focuses on worry, fear, and the physiological responses associated with anxiety (over the past week), such as “I felt I was close to panic.” Its psychometric properties are strong in both clinical and community samples (Antony, Bieling, Cox, Enns, & Swinson, 1998). It also correlates with other established measures of anxiety (Antony et al., 1998; Brown, Chorpita, Korotitsch, & Barlow, 1997; Crawford & Henry, 2003). Cronbach's $\alpha = .83$ in the current study.

2.2.4.3. Depression. The Depression, Anxiety, and Stress Scale-21 item version (DASS-21; Lovibond & Lovibond, 1995) 7-item depression subscale focuses on feelings of dysphoria and hopelessness over the past week, such as “I felt down-hearted and blue.” The subscale demonstrates adequate internal consistency (Osman et al., 2012) and concurrent validity with other measures of depression (Brown et al., 1997). Cronbach's $\alpha = .90$ in the current study.

2.2.4.4. Eating. The Eating Attitudes Test-26 item version (EAT-26; Garner, Olmsted, Bohr, & Garfinkel, 1982) is a self-report measure of symptoms and concerns related to eating disorders. Higher scores reflect more eating disorder symptoms. The EAT-26 has been shown to have strong internal consistency in non-clinical samples (Garner et al.) and diagnostic support for its construct validity (e.g., Mintz & O'Halloran, 2000). Cronbach's $\alpha = .91$ in the current study.

2.3. Procedure

Following self-selection of a mental health domain (i.e., alcohol, anxiety, depression, or eating study), visitors to the PIMH website completed the informed consent form, which included information about the upcoming tasks (e.g., “this study examines attitudes, preferences, and beliefs related to mental health issues”). Following consent, the domain's measures (IAT/BIAT, explicit association semantic differential items, and symptom questionnaire, all specific to that domain) and demographic questions were administered in random order, with the exception of the semantic differential items that always immediately followed the symptom questionnaire. Participants were then fully debriefed, and given the opportunity to view their IAT score and feedback if they wished (e.g., “Your implicit data suggest that you strongly identify more with HAPPY than SAD”). Eighty-six percent of participants that consented saw all tasks and were debriefed. Mental health resource information (e.g., links to treatment referrals) was provided during both consent and debriefing.

3. Results

3.1. Data reduction and scoring

All questionnaires (DASS-21, EAT-26, and AUDIT) were scored according to their original publications or manuals. Data for the depression, anxiety, and eating disorder symptom measures were excluded if participants did not complete at least 80% of items on the questionnaire (otherwise, mean imputation was used to replace any missing values). Included in the final analyses were: 15,516 participants' DASS-21 anxiety subscale scores (9.1% of the 17,075 participants did not respond to any items, or failed to meet the 80% criterion, on the DASS-21 anxiety subscale[3]), 22,364 participants' DASS-21 depression subscale scores (7.3% of the 24,126 consenting participants did not respond to any of the items or failed to meet 80% criterion), and 9416 participants' EAT-26 scores (6.9% of the 10,115 consenting participants did not respond to any of the items or failed to meet the 80% criterion). Alcohol symptom data were excluded for participants with a history of drinking who did not respond to eight or more items on the AUDIT. Of the 12,387 consenting individuals, 8.3% (1023 participants) were not included in AUDIT analyses based on this criterion. Across measures, individual item scores for all participants were used in the calculation of Cronbach's alpha reported above.

All implicit association measures were scored following recommendations by Greenwald and colleagues (Greenwald, Nosek, & Banaji, 2003; Nosek, Bar-Anan, Sriram, & Greenwald, 2012). The IAT and BIAT scoring algorithms create a *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 current scoring recommendations (e.g., cut points to exclude data with high error rates or too many fast trials), the

[3] 21,304 total individuals consented to this study, however 4229 were administered the DASS -21 stress subscale.

following were excluded: 772 (6.2%) participants' alcohol BIAT data, 1002 (4.7%) participants' anxiety BIAT data, 1058 (4.4%) participants' depression BIAT data, and 1146 (11.3%) participants' eating IAT data.[4]

Prior to performing analyses, all study variables were checked for normality and outliers via visual inspection of Q–Q plots. Due to negative skew, the eating IAT was reflected and transformed using a square-root transformation, then reflected again to maintain the original direction of the variable. Due to positive skew, the AUDIT and EAT-26 were square-root transformed (following the addition of a constant of 1). Age (within each domain) was substantially positively skewed. All other study variables appeared normal.

3.1.1. Removal of outliers

Median absolute deviation (MAD) was used to identify extreme outliers in the implicit association measures, symptom measures, and age variable across domains. MAD is a more robust dispersion statistic than standard deviation, which is unduly influenced by outliers (Leys, Ley, Klein, Bernard, & Licata, 2013). Following Leys et al. (2013), cases with a MAD of 3 or more were removed from the samples. The following cases with a MAD of 3 or more were removed from final analyses: 28 (.2%) of the alcohol BIAT, 78 (.4%) of the anxiety BIAT, 78 (.3%) of the depression BIAT, 53 (.5%) of the eating IAT, 157 (.7%) of the DASS-21 anxiety subscale, 473 (2.0%) of the DASS-21 depression subscale, and 41 (.4%) of the EAT-26. Across domains, between 9.3% and 13.5% of the cases in age across studies were considered outliers. However, given the large proportion of “outliers” in age, all values were left in the dataset for analyses using age.

3.2. Descriptive statistics: implicit associations, explicit associations, and symptom measures

See Table 3 for complete descriptive statistics for the dependent variables.

3.2.1. Alcohol study

On average, individuals did not have a bias to implicitly associate the self with drinking or abstaining (mean not significantly different from zero, $t[9395] = 1.21$, $p = .228$, $d = .01$) on the BIAT. However, they demonstrated a bias to explicitly associate the self (vs. others) with being abstainers vs. drinkers (mean significantly different from zero, $t[10,849] = -27.55$, $p < .001$, $d = -.26$) on the semantic differential items. Scores on the AUDIT before outlier removal ranged from 0 to 40, with 44.5% of the sample scoring 8 or above, which is indicative of hazardous or harmful alcohol use, according to the measure's authors. Using a more conservative criterion (scoring 12 or more), 26.1% of participants indicated potentially hazardous drinking behaviors. This is higher than a past community sample, in which 11–14% of men exhibited scores indicative of abuse or dependence and 5–7% of women exhibited scores indicative of abuse or dependence (Volk et al., 1997). In another online study of mental health symptoms (Shapiro, Chandler, & Mueller, 2013), 37.1% positively endorsed at least one item on the CAGE-AID (Brown & Rounds, 1995), which indicates a positive screen for substance use, including drug use.

3.2.2. Anxiety study

Individuals tended to demonstrate a bias to implicitly associate the self as calm (relative to anxious; $t[17,656] = -113.27$, $p < .001$, $d = -.85$) on the BIAT, while they showed a bias to explicitly associate the self (relative to others) as anxious (relative to calm; mean significantly different from zero, $t[19,644] = 52.59$, $p < .001$, $d = .38$) on the semantic differential items. Scores on the DASS-21 anxiety subscale in the full sample ranged from 0 to 42, with 19.0% of the sample reporting severe to extremely severe symptoms of anxiety (i.e., scored 15 or higher). Our sample's mean on the DASS-21 anxiety subscale is higher than both the mean of a previous community sample (e.g., community mean: 3.76; Henry & Crawford, 2005) and another online study of mental health (e.g., 2.9% of participants endorsed clinical levels of general anxiety; Shapiro et al., 2013).

3.2.3. Depression study

On average, individuals had a bias to implicitly associate the self as happy (relative to sad; mean significantly different from zero, $t[20,030] = -127.82$, $p < .001$, $d = -.90$) on the BIAT, while they had a bias to explicitly associate the self (relative to others) as sad (relative to happy; mean significantly different from zero, $t[22,330] = 43.63$, $p < .001$, $d = .29$) on the semantic differential items. Scores on the DASS-21 depression subscale ranged from 0 to 42, with 16.3% of the sample scoring in the severe to extremely severe range (i.e., scored 21 or higher). This sample's mean on the DASS-21 depression subscale is higher than the mean of a previous community sample (e.g., community mean: 5.66; Henry & Crawford, 2005) and another online examination of mental health (e.g., approximately 5% of individuals endorsed clinically significant depressive symptoms; Shapiro et al., 2013).

3.2.4. Eating study

[5] On average, individuals showed a bias to associate shame with high-fat (vs. low-fat) foods, both implicitly (mean significantly different from −.65, which represents the transformed original zero score, $t[8085] = 144.70$, $p < .001$, $d = 1.61$) on the normalized IAT scores, and explicitly (mean significantly different from zero, $t[9405] = 83.66$, $p < .001$, $d = .86$) on the semantic differential items. Scores on the EAT-26 (untransformed, with outliers) ranged from 0 to 78, with 18.6% of the sample scoring 20 or above, which suggests problematic eating behaviors, or concerns about weight and dieting. This percentage is higher than the percentage in a previous college student sample (a group particularly at risk for eating disorders), in which 10.9% of women and 4.0% of men reported scores of 20 or above (Hoerr et al., 2002), however much lower than the percentage in an online eating disorder support group, with 87.7% of participants reporting scores of 20 or above (Darcy & Dooley, 2007).

[4] All consenting participants were included in all analyses, unless otherwise noted (e.g., implicit association or symptom data removed). Additionally, individuals administered the “alcohol-irresistible” BIAT were included in the final demographic and symptom analyses to retain as much data as possible, even though they had missing data for implicit and explicit associations.

[5] A difference in mean implicit association scores based on IAT block order (i.e., whether participants saw *high-fat food + shameful* and *low-fat food + acceptable* paired first versus seeing *high-fat food + acceptable* and *low-fat food + shameful* paired first) was found for the eating IAT, $t(8031) = -18.06$, $p < .001$, Cohen's $d = -.40$. Individuals who saw *high-fat food* paired with *shameful* first exhibited stronger *high-fat food +shameful* implicit associations than individuals who saw *high-fat food* paired with *acceptable* first. However, the implicit-explicit association correlation and the implicit association-symptom measure correlation did not vary based on block order (using Fisher's r-to-z transformation to compare correlation strengths, $ps > .600$). Block order differences did not emerge in the alcohol, anxiety, or depression domains.

Table 3
Effect Sizes and Descriptive Statistics for each Mental Health Study.

Variable	Alcohol (n = 12,387)			Anxiety (n = 21,304)			Depression (n = 24,126)			Eating (n = 10,115)		
	Implicit associations	Explicit associations	Symptoms	Implicit associations	Explicit associations	Symptoms	Implicit associations	Explicit associations	Symptoms	Implicit associations	Explicit associations	Symptoms
Overall mean, full sample (SD)	.00 (.43)	−.62 (2.34)	8.52 (6.80)	−.36 (.44)	.88 (2.36)	10.62 (8.42)	−.39 (.45)	.58 (1.99)	11.76 (9.82)	.68 (.40)	2.26 (2.62)	11.68 (11.05)
Overall mean, normalized variable (SD)	.00 (.43)	–	2.88 (1.12)	−.36 (.43)	–	10.34 (7.99)	−.40 (.44)	–	11.16 (9.00)	−.43 (.13)	–	3.24 (1.40)
Correlations												
Implicit	r = 1	–	–	r = 1	–	–	r = 1	–	–	r = 1	–	–
Explicit	r = .24**	r = 1	–	r = .25**	r = 1	–	r = .32**	r = 1	–	r = .16**	r = 1	–
Symptoms	r = .30**	r = .52**	r = 1	r = .27**	r = .44**	r = 1	r = .35**	r = .57**	r = 1	r = .10**	r = .41**	r = 1
Age	r = −.07**	r = .06**	r = −.06**	r = −.18**	r = −.11**	r = −.23**	r = −.14**	r = −.07**	r = −.10**	r = .14**	r = −.11**	r = −.11**
Education	r_s = −.01	r_s = .15**	r_s = .06**	r_s = −.12**	r_s = −.04**	r_s = −.17**	r_s = −.10**	r_s = −.05**	r_s = −.08**	r_s = .03*	r_s = −.04**	r_s = −.05**
Sex (t-tests)	t = 6.38**, d = .14	t = 9.18**, d = .18	t = 15.34**, d = .29	t = −6.42**, d = −.12	t = −19.27**, d = −.31	t = −10.65**, d = −.19	t = −.34, d = 0	t = 2.64*, d = .03	t = 8.59**, d = .13	t = −2.46, d = −.07	t = −12.97**, d = −.32	t = −21.46**, d = −.50
Male (M, SD)	.04 (.43)	−.37 (2.31)	3.07 (1.11)	−.40 (.42)	.34 (2.53)	9.27 (7.62)	−.40 (.44)	.63 (2.03)	11.94 (9.29)	−.43 (.14)	1.62 (2.53)	2.75 (1.12)
Female (M, SD)	−.02 (.42)	−.79 (2.34)	2.75 (1.10)	−.35 (.43)	1.09 (2.25)	10.75 (8.08)	−.40 (.45)	.56 (1.97)	10.79 (8.85)	−.42 (.13)	2.44 (2.61)	3.39 (1.44)
Ethnicity (t-tests)	t = −.44, d = −.02	t = 1.81, d = .06	t = .79, d = .03	t = −.37, d = 0	t = 1.34, d = .03	t = −3.44*, d = −.09	t = −1.33, d = −.02	t = 3.25*, d = .07	t = 1.58, d = .03	t = 2.18, d = .08	t = −2.90*, d = −.10	t = −2.57*, d = −.09
Non-Hispanic (M, SD)	.00 (.43)	−.59 (2.35)	2.89 (1.11)	−.36 (.43)	.91 (2.37)	10.19 (7.91)	−.40 (.45)	.60 (2.00)	11.11 (8.96)	−.42 (.13)	2.27 (2.60)	3.23 (1.40)
Hispanic (M, SD)	.01 (.41)	−.72 (2.22)	2.86 (1.08)	−.36 (.42)	.84 (2.33)	10.93 (8.24)	−.39 (.43)	.46 (1.94)	10.80 (8.89)	−.43 (.13)	2.53 (2.70)	3.35 (1.40)
Race (ANOVA)[a]	F = 4.43*, η_p^2 = .00	F = 17.88**, η_p^2 = .01	F = 62.08**, η_p^2 = .02	F = 5.45**, η_p^2 = .00	F = 36.94**, η_p^2 = .01	F = 14.93**, η_p^2 = .00	F = 3.17, η_p^2 = .00	F = 18.68**, η_p^2 = .00	F = 53.69**, η_p^2 = .01	F = 14.17**, η_p^2 = .01	F = 7.40**, η_p^2 = .00	F = 6.95**, η_p^2 = .00
Caucasian (M, SE)	.01 (.01)[a]	−.51 (.03)[a]	2.96 (.01)[a]	−.36 (.00)[a]	1.01 (.02)[a]	10.39 (.08)[a]	−.40 (.00)	.62 (.02)[ab]	11.35 (.07)[ab]	−.42 (.00)[a]	2.36 (.03)[a]	3.28 (.02)[a]
Asian (M, SE)	.01 (.02)[ab]	−1.02 (.10)[b]	2.53 (.05)[b]	−.41 (.01)[b]	.60 (.06)[b]	10.03 (.23)[ab]	−.39 (.01)	.72 (.05)[a]	11.89 (.21)[bc]	−.46 (.01)[b]	1.93 (.10)[b]	3.24 (.06)[a]
African American (M, SE)	−.06 (.02)[b]	−1.06 (.09)[b]	2.40 (.04)[b]	−.38 (.01)[ab]	.27 (.07)[c]	8.99 (.26)[b]	−.42 (.01)	.21 (.05)[c]	8.22 (.24)[d]	−.42 (.01)[a]	1.94 (.11)[b]	2.97 (.06)[b]
Multiracial (M, SE)	−.01 (.02)[ab]	−.80 (.10)[ab]	2.84 (.05)[ac]	−.35 (.01)[a]	.77 (.07)[ab]	11.72 (.28)[c]	−.37 (.01)	.66 (.06)[ab]	12.67 (.26)[c]	−.44 (.01)[ab]	2.12 (.12)[ab]	3.30 (.06)[a]
Other (M, SE)	−.01 (.02)[ab]	−.89 (.08)[b]	2.78 (.04)[c]	−.37 (.01)[ab]	.74 (.06)[ab]	10.94 (.23)[ac]	−.38 (.01)	.47 (.05)[b]	10.66 (.21)[a]	−.43 (.01)[a]	2.22 (.09)[ab]	3.26 (.05)[a]

Note. The full sample (using the original, untransformed, with outliers) variables are reported in the first line of the table. The second line includes the means and standard deviations of the transformed (with outliers removed) variables, if applicable. The transformed variables were used in subsequent analyses for this table. Effect sizes for *t*-tests were calculated using the following formula: Cohen's $d = (M_1 - M_2)/SD_{pooled}$.

*$p < .01$, **$p < .001$.

[a] Races that do not significantly differ from each other share the same superscript letter (e.g., all races with an "a" superscript are not significantly different from each other at $p < .01$).

3.3. Relationships among implicit associations, explicit associations, symptom measures, and demographic characteristics

Pearson correlations were computed to test the relationship between implicit and explicit associations and symptoms within studies. To evaluate the relationships between demographic measures and the symptom, explicit association, and implicit association measures, Pearson correlations were conducted with age, Spearman correlations with educational attainment, *t*-tests for the dichotomous gender and ethnicity variables, and Analyses of Variance (ANOVAs) for the categorical race variable. Given the large sample size and the large number of tests conducted, alpha was set at .01 for all analyses to reduce the likelihood of false-positive results.

As seen in Table 3 and in line with hypotheses, across studies, stronger implicit associations reflecting relatively more unhealthy mental health evaluations (e.g., self as a drinker, anxious, or sad, or high-fat food as shameful) were modestly associated with more unhealthy explicit associations. Moreover, stronger unhealthy implicit associations were reliably associated with greater endorsement of symptoms in the same domain. Not surprisingly, the correlation strengths between explicit associations and symptom measures were stronger than the correlations with the implicit measures, likely owing to the shared method variance. This pattern of correlations among dependent variables was consistent across domains.

With respect to the relationships between the demographic variables and the symptom, explicit association, and implicit association measures, few consistent findings emerged. Although some tests reached statistical significance, in general, race and ethnicity did not predict meaningful or large differences on the symptom measures, implicit associations, and explicit associations. There were small relationships with age; greater age was associated with weaker *self*-psychopathology consistent implicit associations in the alcohol, anxiety, and depression domains. Moreover, there were fewer self-reported symptoms associated with greater age across all four domains. Psychopathology-consistent explicit associations in the anxiety, depression, and eating disorders domain were negatively associated with age, however individuals tended to report stronger *self-drinker* associations with age (again, the effects were small). Small effect sizes also emerged within the relationships with education. In the anxiety and depression domains, more education was associated with fewer self-reported symptoms and weaker implicit and explicit *self-anxious* and *sad* associations. Small to moderate gender effects also emerged; being a woman was associated with greater self-reported symptoms and stronger explicit and implicit associations in the anxiety and eating domains, and fewer self-reported symptoms and weaker explicit and implicit associations in the alcohol domain. Surprisingly, no gender differences emerged between men and women in implicit association strength in the depression domain, and men reported more depressive symptoms and stronger self-sad (compared to others) explicit associations than women.

3.4. Incremental predictive validity of implicit mental health evaluations

To address the question of whether implicit evaluations add to the prediction of self-reported symptoms, over and above explicit mental health evaluations, hierarchical linear regressions were conducted within each study. For each study, implicit associations uniquely predicted symptoms, beyond explicit associations (alcohol: $F[2, 9007] = 2094.42$, $p < .001$; anxiety: $F[2, 13{,}340] = 1953.98$, $p < .001$; depression: $F[2, 19{,}155] = 5510.50$, $p < .001$; eating: $F[2, 7850] = 849.26$, $p < .001$; see Table 4 for details, including effect sizes).

3.5. Secondary analyses using single target explicit associations

To further evaluate the relationship between implicit associations and explicit reports, we ran secondary analyses using just one of the explicit association semantic differential items (as opposed to taking the relative difference between the two items) to examine whether the implicit/explicit relationship differed when just examining the explicit association as a single target (e.g., association with the self, as opposed to the self relative to others). In these analyses, we used the “yourself” explicit association questions (in the alcohol use, anxiety, and depression studies) and the “high-fat” question (in the eating study), with higher values representing stronger *self +psychopathology* or *high-fat food +shameful* associations. Results revealed that individuals were more likely to self-identify with being a drinker (versus abstainer, $M = .35$, $SD = 2.44$), feeling anxious (versus calm, $M = .74$, $SD = 1.96$), and feeling happy (versus sad, $M = -.90$, $SD = 1.84$). Individuals were also more likely to rate high-fat food as shameful (versus acceptable, $M = .13$, $SD = 1.91$). Note, when comparing these descriptive statistics to the relative dual target version of the explicit association measure (reported in the main analyses), some interesting differences emerge. Within the alcohol study, although individuals were more likely to associate themselves with being a drinker (compared to abstainer) on the single target version, when they compared themselves to others, they more strongly associated with abstaining. The opposite pattern emerged in the depression study. Although participants more strongly associated themselves with being happy (as compared to sad) on the single target version, when comparing themselves to others, they considered themselves more strongly associated with sadness. This points to the value of examining both single and dual target indicators.

Other analyses were mostly similar using the single target explicit association measures. Correlation strengths for implicit with single target explicit associations (alcohol use: $r = .31$; anxiety: $r = .27$; depression: $r = .37$; eating: $r = .18$; all $ps < .001$) and for symptom with single target explicit associations (alcohol use: $r = .61$; anxiety: $r = .52$; depression: $r = .69$; eating: $r = .45$; all

Table 4
Linear regression models pred ons and implicit associations.

	B	*β*	*t*	f^2
Alcohol				
Step 1			$R^2 = .285$	.40
Explicit associations	.25	.53	59.87**	
Step 2			$\Delta R^2 = .033$	.46
Explicit associations	.23	.49	54.37**	
Implicit associations	.48	.19	20.80**	
Anxiety				
Step 1			$R^2 = .199$	.25
Explicit associations	1.48	.45	57.65**	
Step 2			$\Delta R^2 = .027$	.29
Explicit associations	1.35	.40	51.41**	
Implicit associations	3.16	.17	21.64**	
Depression				
Step 1			$R^2 = .332$	.50
Explicit associations	2.66	.58	97.51**	
Step 2			$\Delta R^2 = .034$	.57
Explicit associations	2.39	.52	85.59**	
Implicit associations	3.90	.19	31.80**	
Eating				
Step 1			$R^2 = .177$	.22
Explicit associations	.23	.42	41.03**	
Step 2			$\Delta R^2 = .001$	.22
Explicit associations	.22	.41	39.97**	
Implicit associations	.39	.04	3.56**	

**$p < .001$.

$ps < .001$) were either similar or stronger than the dual target explicit association correlations reported in Table 3. Further, rerunning the incremental predictive validity regression models with the single target explicit associations yielded similar results to analyses with the dual target measures (i.e., the implicit measures continued to show significant incremental predictive validity in each case, though the magnitude of variance accounted for was slightly smaller; details available from the first author).

4. Discussion

The present study examined implicit associations across four clinical domains in a large, diverse sample, and their association with explicit associations, symptom measures, and key demographic characteristics. Within the alcohol use, anxiety, depression, and disordered eating domains, modest, positive correlations emerged between implicit and explicit associations, and between implicit associations and symptom measures. Not surprisingly, correlations between explicit associations and symptom measures were stronger than either of their respective correlations with implicit associations, perhaps owing to shared method variance given both explicit associations and symptoms were assessed using explicit self-report. Notably, any statistically significant relationship that emerged between implicit associations and demographic variables exhibited small effect sizes, perhaps suggesting that the BIAT and IAT are tools that can be used across diverse groups of individuals. Implicit associations were also found to be incrementally predictive of symptom measures, above and beyond explicit associations, across all four disorder domains (albeit with small effect sizes).

We are reticent to make overly broad claims that extend beyond these particular implicit tasks and this sample; however, it is notable that individuals exhibited a bias for psychopathology-related (vs. health-related) implicit mental health evaluations, on average, only in the context of eating disorder-relevant concerns (i.e., showing an implicit bias for *high-fat foods-shameful*). This finding may reflect the ubiquity with which North American culture assigns a moral value to food (as the majority of participants were from the United States), and the pervasiveness of the current anti-fat culture (Crandall, Merman, & Hebl, 2009). In contrast, individuals exhibited positive implicit associations (i.e., *self-calm* or *self-happy*) in the anxiety and depression domains, on average, and showed no implicit *self* associations between *abstaining* or *drinking* in the alcohol use domain. These three latter tasks have in common a central self-evaluation component, and suggest that among this unselected sample, implicit self-evaluations may be somewhat protected. This is in line with other *self*-related implicit evaluation research; for example, among American, Japanese, and Chinese cultures, implicit self-esteem appears to be universally positive (Yamaguchi et al., 2007). These relatively positive-self implicit associations are noteworthy, given that these participants reported relatively high mean levels of symptoms for an unselected sample, and even biases for self-anxious and self-sad explicit self-associations in the anxiety and depression domains.

Despite the discrepant implicit-explicit mean-level associations in the alcohol use, anxiety, and depression domains, there remains a small, positive correlation between implicit and explicit associations across all four domains in the present investigation. Thus, it appears that implicit and explicit reports are measuring at least some related content, though there are likely different factors influencing the implicit and explicit responses. Previous work has examined moderators of implicit–explicit relationships in the social psychology domain, and has found that greater concern regarding self-presentation predicts weaker implicit-explicit correspondence (Nosek, 2005). Future work could assess whether self-presentation concerns predict differing implicit-explicit responses in clinical applications; owing to the stigma of mental health difficulties, it is likely that self-presentation affects how one explicitly responds to questions regarding mental health, though it is intriguing that it is the explicit responses that tend to be more negative in the current study (counter to typical self-presentation explanations).

Critically, the importance of these particular implicit associations tested in the present study should not be overstated, and there are of course other mental health-related implicit associations that are also valuable predictors of alcohol use, anxiety, depression, and eating disorder symptoms. However, we believe the specific associations tested here are informative. For the alcohol, anxiety, and depression domains, we opted to test implicit associations that had clear theoretical ties to core aspects of the psychopathology and had been measured in previous research (alcohol: Lindgren, Neighbors, et al., 2013; anxiety and depression: Glashouwer et al., 2012). Within the eating disorder domain, we tested a novel IAT examining *high fat-food* + *shameful* implicit associations. Despite its modest correlation strength with eating disorder symptoms, the IAT was not as strong of a predictor of symptoms when controlling for explicit associations. This suggests that there may be better concepts to measure to capture automatic thoughts among individuals with eating disorder symptoms.

It is also notable that when examining the relationship between demographic variables and implicit associations, explicit associations, and symptoms, the effect sizes are very similar across tests. With our extremely large sample size, we were able to detect significant effects of demographic characteristics, even when effects were very small. Thus, the general lack of effects of ethnicity and extremely small effect sizes of race suggest that use of the IAT within these domains may be relatively uninfluenced by differences among individuals in these respects. There were also small effects of education, when measuring by number of years of education. These results were in line with the effect sizes and direction of effects for age, so it is important to note that these two measures may be related (i.e., greater age related to more education). Together, these findings suggest the IATs can likely be used across relatively diverse samples without needing to account for different mean values across groups, which establishes one element of the tasks' measurement invariance.

Gender had variable effects on the implicit associations, explicit associations, and symptom measures. Within each domain, gender had the smallest effect on the implicit association (vs. other) measures, suggesting that perhaps social desirability or gender roles may more strongly influence explicit reports of mental health symptoms. Surprising results emerged in the depression domain; men reported higher depressive symptoms than women, while *self-depressed* implicit associations did not differ between men and women. The gender difference in symptoms is contrary to the majority of previous research suggesting that women experience depression at higher rates than men (Kessler et al., 2003). However, more recent evidence suggests that men and women may experience depression at more similar rates when the measure includes male-oriented symptom items (e.g., substance abuse and aggression; Martin, Neighbors, & Griffith, 2013). In the present study, it is also possible that men may be more comfortable disclosing depressive symptoms on the web, though this is clearly speculative. Further evaluating gender differences in reporting on the web in this domain will be important for future work. Also, women in the disordered eating domain explicitly expressed stronger associations between *high-fat food* and *shameful* and self-reported more eating-disorder symptoms than men, though gender differences did not emerge on the implicit association measure. These findings are consistent with previous research in which men and women did not significantly differ on a weight identity IAT (Grover, Keel, & Mitchell, 2003), and may suggest that cultural messages about the 'morality' of high-fat foods have been internalized comparably by both men and women, even if they are reported differently explicitly.

Finally, understanding the implicit association measures' incremental predictive validity allows us to examine if additional clinical information can be predicted by the IATs above and beyond that of known risk factors. For example, previous research in this area suggests that stronger implicit associations linking the *self* to *death* (as opposed to *life*) are able to uniquely predict suicide attempts, even when known risk factors are accounted for (Nock et al., 2010). In the current study, all four implicit association measures accounted for additional variance in symptom scores after controlling for explicit associations. Although the effect sizes of implicit associations' prediction were small, these results suggest that implicit association measures may be beneficial to examine further as potential clinical tools, especially to assess thoughts or behaviors for which self-report may be unreliable (e.g., due to motivation to conceal, lack of conscious awareness, or difficulties with prediction and introspection).

It is also important to note that this is the first large-scale study that examines implicit associations relating to mental health among a diverse sample. Although laboratory studies are very valuable, they are typically limited to small (and often homogenous) sample sizes and often are not powerful enough to detect differences in implicit associations based on critical individual differences. This is the first large-scale study that examines how mental health implicit associations vary based on a number of critical demographic characteristics, such as race and ethnicity, which is extremely important information for both researchers and clinicians. Without these extremely large and diverse samples, we would not have been able to detect the relatively small effects of demographic variables across implicit associations, pointing to the utility of our study samples. Thus, while we do not claim our sample is representative of the general population, it is by far the largest and most representative (based on its heterogeneity) sample in which implicit mental health associations have been examined, and gives researchers the best information to date about norms for these measures, and mean differences across numerous demographic groups.

4.1. Limitations and conclusion

Results should be interpreted in light of several limitations. First, although diverse and more heterogeneous than typical college samples, our sample is not representative of any particular population. Second, participants were assessed for symptoms but not clinically diagnosed. However, use of an unselected sample allowed us to examine implicit associations across a wide range of symptomatology, and our sample was more highly symptomatic than typical community samples. It should also be noted that the self-reported symptom measures in the current study have limitations (e.g., reliance on self report and they do not establish a diagnosis), despite their good validity with clinical and non-clinical samples. Third, given that the data were collected via the PIMH web site, participants were able to select the specific studies they were interested in taking and participate in as many studies as desired. While having multiple data points for a single participant is not ideal, the large sample size decreased the likelihood that multiple administrations would significantly affect the results (see discussion in Nosek, Banaji, & Greenwald, 2002). In web studies, it is also possible for participants to end the study early; however, 86% of the participants that consented did complete all study tasks. We also had no way to assess the setting in which the participants completed the studies and whether or not the participants were fully engaged. Notwithstanding, web-based computer data also have several advantages, such as rapid data collection and large sample sizes. Further, the prevalence of general anxiety and depression symptoms disclosed online is equal to or greater than prevalence in the general population (Shapiro et al., 2013), supporting the use of a web-based methodology for the current questions. Additionally, across each disorder domain, only one disorder-relevant concept was measured implicitly. Previous research suggests that different variants of the IAT show unique predictive validity (e.g., drinking identity and drinking avoidance associations predict differently; Lindgren, Foster, Westgate, & Neighbors, 2013; see also Teachman & Woody, 2003, in the anxiety domain), so it will be important to assess a broader range of implicit mental health associations to determine the strongest predictors for different types of clinical outcomes. Finally, the (B) IAT needs to be understood as a relative measure and it is not a measure of "lie detection," so needs to be interpreted carefully and cannot be said to capture an individual's "true" feelings about a topic (Nosek et al., 2007).

Possible order effects are also an important consideration. In the current study, participants completed the explicit associations and symptom measures in a fixed order, so this may have potentially influenced our results. Additionally, we found that there was an effect of IAT block order for the eating disorder IAT. This was surprising, given that there have been improvements to the IAT design (i.e., increasing number of practice trials mid-way through the task), that have almost eliminated block order effects (Nosek, Greenwald, & Banaji, 2005). Despite the block order effect observed, it is noteworthy that the symptom and explicit association correlations with the eating disorder IAT did not differ as a function of block order, suggesting that order effects are not likely to alter the main conclusions drawn in this study.

The current study is the first to describe implicit mental health associations across four highly prevalent mental disorder domains in very large and heterogeneous samples. This work extends laboratory-based findings by examining how implicit alcohol, anxiety, depression, and eating-related associations vary (or do not vary) based on age, gender, race, ethnicity, and education, given that the samples allow for the fist examination of individual differences in implicit mental health associations that can detect even small effects. The results suggest that there are very minimal effects of race and ethnicity on implicit mental health associations, which provides important information for future research studies or clinical applications that want to apply these implicit cognition tools across racially diverse groups. Moreover, the results indicate that implicit mental health associations incrementally predict symptoms, over and above individuals' explicit associations with these mental health concepts, which suggests that automatic self- and psychopathology-concepts are important in understanding and predicting alcohol use, anxiety, depression, and eating disorders. The results also suggest that implicit associations are modestly positively correlated with both explicit mental health associations and self-reported symptoms. This research also emphasizes the ability for researchers to assess sensitive and personal topics (that may be difficult to discuss openly in laboratory-based, in-person settings) on the web. It will be important in future work to assess how the implicit–explicit relationships differ among diagnosed vs. unselected samples, and the implicit associations' sensitivity to clinical interventions. Overall, this research highlights the importance and feasibility of examining both implicit and explicit cognition in mental health domains.

Acknowledgments

We would like to thank the Project Implicit Mental Health collaborators, including Brian Nosek, PhD, Mahzarin Banaji, PhD, and Tony Greenwald, PhD. We would also like to thank Frederick Smyth, PhD, and Emily Umansky for their technical assistance. We would like to thank Nauder Namaky for his assistance on PIMH data management. Finally, we would like to thank the Teachman Program for Anxiety, Cognition, and Treatment (PACT) laboratory for their feedback on this manuscript. This research was made possible by NIH grants R01AG033033 and R34MH106770 to Bethany Teachman. B. 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 work was also supported by the Norlien Foundation.

References

Allen, J. P., Litten, R. Z., Fertig, J. B., & Babor, T. (1997). A review of research on the alcohol use disorders identification test (audit). *Alcoholism: Clinical and Experimental Research, 21*(4), 613–619. http://dx.doi.org/10.1111/j.1530-0277.1997.tb03811.x.

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176–181. http://dx.doi.org/10.1037/1040-3590.10.2.176.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (1992). *Audit: The alcohol use disorders identification test, guidelines for use in primary care* (2nd ed.). World Health Organization.

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: automatic and strategic processes. *Behaviour Research and Therapy, 35*(1), 49–58. http://dx.doi.org/10.1016/S0005-7967(96)00069-1.

Beevers, C. G. (2005). Cognitive vulnerability to depression: a dual process model. *Clinical Psychology Review, 25*(7), 975–1002. http://dx.doi.org/10.1016/j.cpr.2005.03.003.

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the depression anxiety stress scales (dass) in clinical samples. *Behaviour Research and Therapy, 35*(1), 79–89. http://dx.doi.org/10.1016/S0005-7967(96)00068-X.

Brown, R. L., & Rounds, L. A. (1995). Conjoint screening questionnaires for alcohol and other drug abuse: criterion validity in a primary care practice. *Wisconsin Medical Journal, 94*(3), 135–140.

Crandall, C. S., Merman, A., & Hebl, M. (2009). *Anti-fat prejudice. Handbook of prejudice, stereotyping, and discrimination* (pp. 469–487). New York, NY, US: Psychology Press.

Crawford, J. R., & Henry, J. D. (2003). The depression anxiety stress scales (dass): normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology, 42*(2), 111–131. http://dx.doi.org/10.1348/014466503321903544.

Darcy, A. M., & Dooley, B. (2007). A clinical profile of participants in an online support group. *European Eating Disorders Review, 15*(3), 185–195. http://dx.doi.org/10.1002/erv.775.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83*(6), 1441–1455. http://dx.doi.org/10.1037/0022-3514.83.6.1441.

Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitudes test: psychometric features and clinical correlates. *Psychological Medicine, 12*(4), 871–878. http://dx.doi.org/10.1017/S0033291700049163.

Glashouwer, K. A., & de Jong, P. J. (2010). Disorder-specific automatic self-associations in depression and anxiety: results of the netherlands study of depression and anxiety. *Psychological Medicine, 40*(7), 1101–1111. http://dx.doi.org/10.1017/S0033291709991371.

Glashouwer, K. A., de Jong, P. J., & Penninx, B. W. J. H. (2012). Prognostic value of implicit and explicit self-associations for the course of depressive and anxiety disorders. *Behaviour Research and Therapy, 50*(7–8), 479–486. http://dx.doi.org/10.1016/j.brat.2012.05.002.

Glashouwer, K., de Jong, P., Penninx, B. J. H., Kerkhof, A. F. M., van Dyck, R., & Ormel, J. (2010). Do automatic self-associations relate to suicidal ideation? *Journal of Psychopathology and Behavioral Assessment, 32*(3), 428–437. http://dx.doi.org/10.1007/s10862-009-9156-y.

Glenn, J. J., Werntz, A. J., Slama, K., Steinman, S. A., Teachman, B. A., & Nock, M. K. (2016). *Suicide and self-injury-related implicit cognition: A large-scale examination and replication* (Manuscript under review).

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4–27. http://dx.doi.org/10.1037/0033-295X.102.1.4.

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*(6), 1464–1480. http://dx.doi.org/10.1037/0022-3514.74.6.1464.

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*(2), 197–216. http://dx.doi.org/10.1037/0022-3514.85.2.197.

Grover, V. P., Keel, P. K., & Mitchell, J. P. (2003). Gender differences in implicit weight identity. *International Journal of Eating Disorders, 34*(1), 125–135. http://dx.doi.org/10.1002/eat.10167.

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the depression anxiety stress scales (dass-21): construct validity and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology, 44*(Pt 2), 227–239. http://dx.doi.org/10.1348/014466505x29657.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*(10), 1369–1385. http://dx.doi.org/10.1177/0146167205275613.

Hollon, S. D., Kendall, P. C., & Lumry, A. (1986). Specificity of depressotypic cognitions in clinical depression. *Journal of Abnormal Psychology, 95*(1), 52–59. http://dx.doi.org/10.1037/0021-843X.95.1.52.

Houben, K., & Wiers, R. W. (2008). Measuring implicit alcohol associations via the internet: validation of web-based implicit association tests. *Behavior Research Methods, 40*(4), 1134–1143. http://dx.doi.org/10.3758/BRM.40.4.1134.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K. … Wang, P. (2010). Research domain criteria (rdoc): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748–751.

de Jong, P. J. (2002). Implicit self-esteem and social anxiety: differential self-favouring effects in high and low anxious individuals. *Behaviour Research and Therapy, 40*(5), 501–508. http://dx.doi.org/10.1016/S0005-7967(01)00022-5.

de Jong, P. J., & Veenstra, E. (2007). *Automatic approach–avoidance tendencies and dysfunctional eating patterns* (Paper presented at the Fifth World Congress of Behavioural and Cognitive Therapy, Barcelona, Spain).

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R. … Wang, P. S. (2003). The epidemiology of major depressive disorder: results from the national comorbidity survey replication (ncs-r). *JAMA: Journal of the American Medical Association, 289*(23), 3095–3105. http://dx.doi.org/10.1001/jama.289.23.3095.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the implicit association test: Iv. What we know (so far). In B. Wittenbrink, & N. S. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford Press.

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*(4), 764–766. http://dx.doi.org/10.1016/j.jesp.2013.03.013.

Lindgren, K. P., Foster, D. W., Westgate, E. C., & Neighbors, C. (2013). Implicit drinking identity: drinker+me associations predict college student drinking consistently. *Addictive Behaviors, 38*(5), 2163–2166. http://dx.doi.org/10.1016/j.addbeh.2013.01.026.

Lindgren, K. P., Neighbors, C., Teachman, B. A., Wiers, R. W., Westgate, E., & Greenwald, A. G. (2013). I drink therefore i am: validating alcohol-related implicit association tests. *Psychology of Addictive Behaviors, 27*(1), 1–13. http://dx.doi.org/10.1037/a0027640.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney: Psychology Foundation.

Martin, L. A., Neighbors, H. W., & Griffith, D. M. (2013). The experience of symptoms of depression in men vs women: analysis of the national comorbidity survey replication. *JAMA Psychiatry*. http://dx.doi.org/10.1001/jamapsychiatry.2013.1985.

Mintz, L. B., & O'Halloran, M. S. (2000). The eating attitudes test: validation with dsm-iv eating disorder criteria. *Journal of Personality Assessment, 74*(3), 489–503. http://dx.doi.org/10.1207/S15327752JPA7403-11.

Nock, M. K., & Banaji, M. R. (2007). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of Consulting and Clinical Psychology, 75*(5), 707–715. http://dx.doi.org/10.1037/0022-006X.75.5.707.

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.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*(4), 565–584. http://dx.doi.org/10.1037/0096-3445.134.4.565.

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6*(1), 101–115. http://dx.doi.org/10.1037/1089-2699.6.1.101.

Nosek, B. A., Bar-Anan, Y., Sriram, N., & Greenwald, A. G. (2012). *Understanding and using the brief implicit association test: I. Recommended scoring procedures*. Unpublished manuscript. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2196002.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the implicit association test: Ii. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*(2), 166–180. http://dx.doi.org/10.1177/0146167204271418.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). *The implicit association test at* *age 7: A methodological and conceptual review. Social psychology and the unconscious: The automaticity of higher mental processes (pp. 265–292*. New York, NY, US: Psychology Press.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the implicit association test: implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54*(1), 14–29. http://dx.doi.org/10.1027/1618-3169.54.1.14.

Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The depression anxiety stress scales—21 (dass-21): further examination of dimensions, scale reliability, and correlates. *Journal of Clinical Psychology, 68*(12), 1322–1338. http://dx.doi.org/10.1002/jclp.21908.

Peris, T. S., Teachman, B. A., & Nosek, B. A. (2008). Implicit and explicit stigma of mental illness: links to clinical care. *Journal of Nervous and Mental Disease, 196*(10), 752–760. http://dx.doi.org/10.1097/NMD.0b013e3181879dfd.

Ranganath, K. A., & Nosek, B. A. (2008). Implicit attitude generalization occurs immediately; explicit attitude generalization takes time. *Psychological Science, 19*(3), 249–254. http://dx.doi.org/10.1111/j.1467-9280.2008.02076.x.

Reinert, D. F., & Allen, J. P. (2002). The alcohol use disorders identification test (audit): a review of recent research. *Alcoholism: Clinical and Experimental Research, 26*(2), 272–279. http://dx.doi.org/10.1097/00000374-200202000-00016.

Roefs, A., Huijding, J., Smulders, F. T. Y., MacLeod, C. M., de Jong, P. J., et al. (2011). Implicit measures of association in psychopathology research. *Psychological Bulletin, 137*(1), 149–193. http://dx.doi.org/10.1037/a0021729.

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (audit): who collaborative project on early detection of persons with harmful alcohol consumption: Ii. *Addiction, 88*(6), 791–804. http://dx.doi.org/10.1111/j.1360-0443.1993.tb02093.x.

Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using mechanical turk to study clinical populations. *Clinical Psychological Science*. http://dx.doi.org/10.1177/2167702612469015.

Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental Psychology, 56*(4), 283–294. http://dx.doi.org/10.1027/1618-3169.56.4.283.

Stoner, S. A., Fedoroff, I. C., Andersen, A. E., & Rolls, B. J. (1996). Food preferences and desire to eat in anorexia and bulimia nervosa. *International Journal of Eating Disorders, 19*(1), 13–22. doi:10.1002/(SICI)1098-108X(199601)19:1<13::AID-EAT3>3.0.CO;2-Z.

Teachman, B. A., Gapinski, K. D., Brownell, K. D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: the impact of providing causal information and evoking empathy. *Health Psychology, 22*(1), 68–78. http://dx.doi.org/10.1037/0278-6133.22.1.68.

Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology, 110*(2), 226–235. http://dx.doi.org/10.1037/0021-843X.110.2.226.

Teachman, B. A., Joormann, J., Steinman, S. A., & Gotlib, I. H. (2012). Automaticity in anxiety disorders and major depressive disorder. *Clinical Psychology Review, 32*(6), 575–603. http://dx.doi.org/10.1016/j.cpr.2012.06.004.

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*(6), 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*(1), 100–109. http://dx.doi.org/10.1037/0021-843X.112.1.100.

United States Census Bureau. (2015a). *Educational attainment in the united states: 2014-detailed tables*. Retrieved from http://www.census.gov/hhes/socdemo/education/data/cps/2014/tables.html.

United States Census Bureau. (2015b). State & county quickfacts: USA. Retrieved from http://quickfacts.census.gov/qfd/states/00000.html.

Vitousek, K. B., & Hollon, S. D. (1990). The investigation of schematic content and processing in eating disorders. *Cognitive Therapy and Research, 14*(2), 191–214. http://dx.doi.org/10.1007/BF01176209.

Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive bias modification and cognitive control training in addiction and related psychopathology: mechanisms, clinical perspectives, and ways forward. *Clinical Psychological Science, 1*(2), 192–212. http://dx.doi.org/10.1177/2167702612466547.

Yamaguchi, S., Greenwald, A. G., Banaji, M. R., Murakami, F., Chen, D., Shiomura, K. … Krendl, A. (2007). Apparent universality of positive implicit self-esteem. *Psychological Science, 18*(6), 498–500. http://dx.doi.org/10.1111/j.1467-9280.2007.01928.x.