



Contents lists available at ScienceDirect

# Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: [www.elsevier.com/locate/jbtep](http://www.elsevier.com/locate/jbtep)



## Decoupling implicit measures of pleasant and unpleasant social attitudes

Alex S. Cohen\*, Melissa R. Beck, Laura A. Brown, Gina M. Najolia

Louisiana State University, Department of Psychology, Baton Rouge, LA, United States

### ARTICLE INFO

#### Article history:

Received 4 November 2008

Received in revised form

24 April 2009

Accepted 27 August 2009

#### Keywords:

Social

Attitudes

Ambivalence

IAT

Implicit

### ABSTRACT

Social attitudes are integral to understanding a wide range of pathological states. The present study adapted the Implicit Association Test, a widely used implicit measure of attitudes, for understanding social attitudes and behavior. In a first study, data from a traditional “Bipolar” IAT and our modified “Unipolar” pleasant and unpleasant IATs were compared in their associations with explicit measures of social attitudes and behavior. In a second study, we examined the relationship between implicit measures of social attitudes and social behavior during a laboratory procedure. Collectively, the present data support the uncoupling of pleasant and unpleasant valences when using implicit measures to understand social attitudes.

© 2009 Elsevier Ltd. All rights reserved.

### 1. Introduction

Social behaviors are central to understanding a broad range of important quality of life variables. Moreover, aberrant social behaviors are diagnostic criteria for most forms of psychopathology, including depression, schizophrenia, anxiety and personality disorders (American Psychiatric Association, 1994). Attempts to understand the underpinnings of social behavior highlight the importance of measuring cognitive processes with traditional questionnaires (Hamilton, Sherman, & Ruvolo, 1990). However, measuring these processes is challenging because assessment is primarily reliant on an individual's explicit subjective evaluations. This is a critical limitation because demand characteristics have a profound influence on individual's responses (Nichols & Maner, 2008; Richman, Kiesler, Weisband, & Drasgow, 1999). Advances from Social and Cognitive Psychology have paved the way for circumventing these limitations by using implicit measures of attitudes. This report presents data from two studies that advance the development of an implicit measure of social attitudes.

The Implicit Association Test (IAT), a measure of attitudes that has garnered considerable empirical support (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, in press; Lane et al., 2007 for recent reviews), assesses the speed of responses as individuals pair

stimuli with “bipolar”, diametrically-opposed attitudes (e.g., “pleasant” versus “unpleasant”) and distinct categories (e.g., “deserts” versus “cars”). Individuals tend to make relatively fast categorizations when categories and attitudes are strongly linked (e.g., “Pleasant” and “Cheesecake”) compared to when they are more weakly linked (e.g., “Unpleasant” and “Cheesecake”). The IAT has been used to understand a broad array of cognitive processes, such as those involved in racism, sexism, depression, substance abuse (see De Houwer, 2002), obesity (Craeynest, Crombez, Koster, Haerens, & De Bourdeaudhuij, 2008) narcissism (Campbell, Bosson, Goheen, Lakey, & Kernis, 2007) and self-esteem (Franck, De Raedt, Dereu, & Van den Abbeele, 2007). Interestingly, IAT performance has been shown to be a predictor of behavior above and beyond explicit evaluations. For example, racist attitudes, as measured by the IAT, have been associated with more negative nonverbal behaviors towards out-group members, while similar association were not found with explicit measures of racism (McConnell & Leibold, 2001).

A major limitation of applying the traditional IAT to the study of social behavior is that it is designed to measure diametrically-opposed attitudes. For example, prior studies have assessed social processes in the context of attitudes towards self versus others using “positive” and “negative” categories arranged as opposing options in a single categorization task (e.g., Tanner, Stopa, & De Houwer, 2006). As noted by Houben and Wiers (2006), this is potentially problematic because the link between the categories and attitudes is generally unclear. That is, the Bipolar IAT does not reveal whether a category is associated with pleasant, unpleasant, or both attitudes. This is a notable problem because attitudes

\* Corresponding author. Louisiana State University, Department of Psychology, 236 Audubon Hall, Baton Rouge, LA 70803, United States. Tel.: +1 225 578 7017; fax: +1 225 578 4125.

E-mail address: [acohen@lsu.edu](mailto:acohen@lsu.edu) (A.S. Cohen).

potentially underlying behavior may be orthogonal or even complimentary rather than oppositional. For example, an individual may have ambivalent attitudes about a category, as simultaneous pleasant (e.g., “Cheesecake tastes good”) and unpleasant (e.g., “Cheesecake is unhealthy”) attitudes are activated. With regard to social relationships, individuals often evidence mixed explicit attitudes towards individuals in their social domains (Cacioppo, Gardner, & Berntson, 1997; Cunningham, Raye, & Johnson, 2004; Priester & Petty, 2001). Moreover, this social ambivalence – the simultaneous activation of disparate attitudes towards people – is an important feature of a range of psychological disorders, including borderline personality disorder (Hopwood & Morey, 2007), schizophrenia (Cohen & Minor, *in press*) and depression (Simpson, Rholes, Campbell, Tran, & Wilson, 2003). In short, empirical studies of abnormal social behavior would benefit from an implicit measure of social ambivalence.

Several recent studies, employing a modified “Unipolar” IAT, offer promise for overcoming these limitations. Of note, ambivalent attitudes towards alcohol have been examined with implicit measures (Houben & Wiers, 2006; Jajodia & Earleywine, 2003). The use of implicit measures is particularly important for understanding alcohol because attitudes towards alcohol use are characteristically ambivalent – many individuals experience simultaneous craving and abhorrence (Rollnick & Miller, 1995). The Unipolar IAT features two separate test administrations in which the attitude scale is deconstructed into distinct subcomponents of Pleasant-to-Neutral and Unpleasant-to-Neutral scales. This allows for separate examination of the unipolar attitude dimensions. Results from prior studies are promising in that subjects showed simultaneous pleasant and unpleasant attitudes to alcohol stimuli (Houben & Wiers, 2006; Jajodia & Earleywine, 2003). Moreover, the IAT demonstrated divergent validity in that pleasant and unpleasant attitudes corresponded to a range of distinct explicit measures of alcohol attitudes (Houben & Wiers, 2006; Jajodia & Earleywine, 2003). Although results from prior Unipolar IAT studies are promising for understanding attitudes towards alcohol, the validity of this measure for understanding social attitudes requires further investigation.

This paper presents data from two studies examining the feasibility for understanding social behavior using a modified Unipolar IAT. The first study compares results from a traditional Bipolar IAT and two Unipolar formatted IATs to investigate whether Pleasant and Unpleasant attitudes are better conceptualized as opposing ends on a continuum or as orthogonal/complimentary dimensions. Comparison of effects from the two IAT measures also help gauge the effects of modifying the IAT. Insofar as IATs are dependent on having clear attitude labels that activate the associated attitudes, a potential concern with the Unipolar IAT measure is that the use of an ambiguous “neutral” label might attenuate the IAT effects (Lane et al., 2007). We also examined the relationship between implicit (IAT) and explicit measures of social attitudes and social behavior with the expectation of modest convergence between these measures.

In the second study, we attempted to replicate the IAT effects observed in the first study. We also employed a more sensitive measure of social behavior involving analysis of expressive behavior using a laboratory verbalization task. During this task, subjects were asked to discuss social relationships from their lives. We employed computerized lexical analysis of subjects’ natural speech behavior to gain insight into their social attitudes. Lexical analysis has a rich history in psychological research, for example, in understanding affective processes and attitudes across a range of disorders, such as schizophrenia (Cohen, Alpert, Nienow, Dinzeo, & Docherty, *in press*; Cohen, St-Hilaire, Aakres, & Docherty, *in press*), alexithymia (Kelley, Lumley, & Leisen, 1997; Tull, Medaglia, &

Roemer, 2005), Post-Traumatic Stress Disorder (Bernard, Jackson, & Jones, 2006; Murray & Segal, 1994; Orsillo, Batten, Plumb, Luterek, & Roessner, 2004), depression and anxiety (Gortner, Rude, & Pennebaker, 2006; Lepore, 1997; Watson & Pennebaker, 1989) and eating disorders (Lyons, Mehl, & Pennebaker, 2006). Our use of lexical analysis in the second study offers a potentially more sensitive assessment of social behavior beyond that used in the first study.

## 2. Study 1 methods

### 2.1. Participants

Subjects included 37 men and 129 women recruited from a public university in the Southeastern United States. Ethnic composition was primarily Caucasian ( $n = 137$ ), but also included 18 African-Americans, four Asian-Americans, five Hispanic Americans and one “other”. The average age of the group was 20.60 ( $SD = 5.03$ , Range = 18–53). This study was approved by the appropriate Institutional Review Board and all subjects provided written informed consent prior to beginning the study.

### 2.2. Implicit Association Test (IAT)

Two different IATs were used in this study. The first was a “traditional” Bipolar IAT, constructed using specifications in Greenwald, Nosek, and Banaji (2003) and Lane et al. (2007) involving five separate blocks. All IATs were programmed using Eprime 2.0. Words were selected from the ANEW (Bradley & Lang, 1999), a lexicon of English words with extensive valence and arousal ratings. The words were roughly matched across conditions in terms of numbers of syllables and letters. A brief overview of the traditional IAT is provided here (see Greenwald et al., 2003 for further information). The IAT measures reaction times as individuals sort unambiguous stimuli into dichotomous “bins”. For this study, these bins were labeled in terms of Categories (i.e., “People”, “Places”) and attitudes (i.e., “Pleasant”, “Unpleasant”). We were interested in disentangling pleasant from unpleasant social attitudes across a broad spectrum of relationship domains. Thus, methodologies employed in prior IAT studies examining social attitudes using “self” versus “others” (de Jong, 2002; Tanner et al., 2006) or “sociable” versus “unsociable” labels would have been less than ideal for this purpose. The “People” category sampled across immediate family (e.g., “mother”), intimate (e.g., “lover”), friends (“buddy”), peer-based (e.g., “classmate”) and authority (e.g., “teacher”) domains. Labels were paired and presented in the left and right upper corners of the screen and tied to “left key” and “right key” responses. Words, categorized as people, places, pleasant or unpleasant, were presented in boxes in the lower center of the screen. Words pertaining to “People”/“Place” were presented in a separate bin from the “Pleasant”/“Unpleasant” words to improve categorization. The first block is a training condition where subjects categorize 10 unambiguous “Pleasant” or “Unpleasant” words (see Appendix 1;  $n = 20$ ; each stimuli presented twice) into “Pleasant” or “Unpleasant” bins. The second block provides similar training for “People”/“Place” words ( $n = 20$ ). The third block is a “test” block and requires individuals to sort the previously seen attitude ( $n = 20$ ) and category words ( $n = 20$ ) into one of two separate combined categories (e.g., “Pleasant and People” or “Unpleasant and Places”). For the fourth and fifth block, the category labels are reversed so that the categories are associated with the opposing attitude from block three. The fourth block is a training condition involving sorting only people and place stimuli, while the fifth block is a test block similar to block three. The order of presentation of blocks two/three and blocks four/five was

counterbalanced across subjects. Reaction time was recorded for each trial, and response to trials in block three are compared to those in block five using an effect size statistic ( $d$ ; the difference of average reaction time scores divided by the pooled SD) that can be interpreted in terms of small (.20–.49), medium (.50–.79) and large (>.80) effects (see Greenwald et al., 2003). This allows one to evaluate potential differences in reaction times between the two test conditions when categories and attitudes are linked in one manner versus when they are oppositionally linked. Increasing Bipolar IAT scores reflect increasing pleasant/decreasing unpleasant social attitudes. A 600 ms penalty was added to all incorrect trials (Greenwald et al., 2003). There was adequate internal reliability between stimuli within each “person”, “people”, “pleasant” and “unpleasant” word category ( $\alpha > .60$ ). For descriptive purposes, we report accuracy scores (i.e., percent stimuli correctly categorized), raw reaction times and IAT effect scores.

The “Unipolar” IATs were presented in a similar format, with two key changes (see Houben & Wiers, 2006). First, the Unipolar IAT consisted of two separate condition blocks assessing pleasant and unpleasant attitudes separately. Second, during the “pleasant” IAT, the “unpleasant” attitude label was replaced with a “neutral” attitude label (see Appendix 1 for words). Similarly, during the unpleasant IAT, the “pleasant” attitude label was replaced with a “neutral” attitude label. Order of Bipolar versus Unipolar IATs was random, as was the presentation of pleasant versus unpleasant unipolar IAT conditions. Increasing Unipolar IAT scores reflect increasing intensity of the relevant emotion (e.g., increasing “pleasant” IAT scores reflect increasing pleasant social association).

### 2.3. Explicit measures of social attitudes

We employed separate explicit measures of unipolar pleasant, unipolar unpleasant and bipolar social attitudes. The explicit bipolar measure was from the Brief Quality of Life – Interview (Lehman, 1995), a commonly used measure in psychopathology studies. This measure involved a summation of the Social and Family satisfaction subscales. These subscales assess explicit satisfaction regarding family and peer/significant other interactions on a scale from 1 (terrible) to 7 (delighted). We were unable to identify existing explicit unipolar scales, so we employed separate scales assessing pleasant (i.e., “How pleasant do you feel towards \_\_\_?”) and unpleasant (i.e., “How unpleasant do you feel towards \_\_\_?”) attitudes in five social domains (i.e., family, intimate, friends, peer-based and authority) along a six point likert scale from “very pleasant/unpleasant” (coded “5”) to not at all pleasant/unpleasant” (coded “1”). Summary scores for pleasant and unpleasant were computed as average ratings across the five social domain items.

### 2.4. Social behavior

Social functioning was measured using the objective social and family subscales from Brief Quality of Life – Interview (Lehman, 1995). For the present study, we focused on six items measuring objective social behavior in family and social domains (e.g., “How often do you talk to a member of your family on the telephone?” and “How often do you spend time with someone you consider more than a friend, like a boyfriend, girlfriend or spouse?”). These questions were examined separately, as opposed to using a sum score, due to low reliability ( $\alpha < .40$ ). Increasing scores reflect improving quality of life.

### 2.5. Analyses

The analyses were conducted in three phases. First, we examined the potential influence of IAT administration order effects.

Second, we examined IAT effects by evaluating whether the effects for each of the Bipolar, Unipolar Pleasant and the Unipolar Unpleasant conditions were statistically different from “0” using one-sample  $t$ -tests. We next employed repeated measures ANOVA examining differences within these conditions. Finally, we computed correlations between implicit, explicit and social behavior measures. Unless otherwise noted, each of the variables examined in this study were normally distributed. All tests reported here are two-tailed.

## 3. Results

### 3.1. Data reduction and order-effects

There were two order effects worth noting. First, subjects receiving the Unipolar IAT first tended to have more robust Bipolar IAT effects than other subjects ( $d$ 's = .66 ± .43 and .42 ± .41 respectively;  $t[164] = 3.66$ ,  $p < .01$ ). Second, subjects tended to show slightly stronger Unipolar Unpleasant IAT effects when they received the Unipolar Pleasant IAT first ( $d$ 's = -.11 ± .38 and .03 ± .35 respectively;  $t[164] = 2.45$ ,  $p < .05$ ). Presumably, these differences reflect priming effects, although it is difficult to systematically examine this issue in our dataset. There were no significant order effects for the Unipolar pleasant ( $d$ 's = .30 ± .45 and .26 ± .47 for Unipolar and Bipolar first respectively;  $t[164] = .66$ ,  $p = .51$ ) or unpleasant ( $d$ 's = -.03 ± .40 and -.04 ± .34 respectively;  $t[164] = .12$ ,  $p = .90$ ) IATs. To address the order effects, we controlled for bipolar versus unipolar order effects when possible.

### 3.2. The bipolar versus the unipolar IAT

Reaction time, accuracy (i.e., percent correct) and IAT effect scores are presented in Table 1 as are the results of the repeated measures ANCOVA's. IAT effect scores for the Bipolar and Unipolar Pleasant, but not the Unipolar Unpleasant, conditions were significantly greater than zero, providing evidence of an IAT effect for these conditions. Finally, the IAT effect scores from the Bipolar, Unipolar Pleasant and Unipolar Unpleasant conditions each significantly differed from each other such that the Bipolar score had the largest magnitude, and the Unipolar Unpleasant had the weakest effects (all  $p$ 's < .05). In sum, IAT effects were observed for the Bipolar and Unipolar Pleasant, but not the Unipolar Unpleasant conditions.

### 3.3. Implicit, explicit and social behavior measures

Bivariate correlations are presented in Table 2. The IAT scores from the Bipolar condition were modestly but significantly

**Table 1**

Descriptive data and test statistics for the bipolar and unipolar implicit measures of social attitudes.

	Reaction time	% Errors	$d$	$t$ (Diff from 0)
<i>Bipolar</i>				
Pleasant condition	891.04 ± 200.81	.04 ± .04	.54 ± .42	16.05**
Unpleasant condition	1130.66 ± 223.00	.09 ± .06		
<i>Unipolar Pleasant</i>				
Pleasant condition	893.25 ± 175.21	.05 ± 0.05	.28 ± .46	7.95**
Neutral condition	1009.30 ± 218.94	.07 ± .06		
<i>Unipolar Unpleasant</i>				
Unpleasant condition	979.39 ± 214.95	.05 ± .05	-.03 ± 0.37	-1.21
Neutral condition	963.95 ± 195.35	0.05 ± 0.05		
Omnibus $F$				20.73***

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

correlated to those in the Unipolar Pleasant condition but not to those in the Unipolar Unpleasant condition. These correlations were significantly different from each other (Fisher  $r$ -to- $z = 3.40$ ,  $p < .001$ ). Moreover, the IAT scores from the Unipolar Pleasant and Unpleasant conditions were positively and significantly correlated with each other ( $r[164] = .16$ ,  $p < .05$ ), suggesting that they are not inversely but complimentary to each other (albeit very weakly).

Interestingly, the bipolar implicit and explicit measures were significantly inter-correlated at a small effect size. The unipolar implicit and explicit unpleasant measures were similarly inter-correlated. However, the explicit pleasant measure was associated (inversely) with the implicit unpleasant measure but not the pleasant measure. Generally, there were few significant correlations between the social attitude measures and social behavior.

#### 4. Discussion

Converging evidence from this study supports the use of the Unipolar IAT for understanding social attitudes. First, subjects, as a group, showed a robust IAT effect using the Bipolar IAT, although the Bipolar IAT was ineffective for clarifying whether the IAT effect reflected strong pleasant attitudes, unpleasant attitudes or both. These results indicate that the Unipolar IATs were helpful for resolving this issue because there were demonstrable IAT effects for the Unipolar Pleasant but not the Unipolar Unpleasant IAT condition. Thus, it stands to reason that the Bipolar IAT effects are driven by Pleasant rather than Unpleasant attitudes. Second, the Unipolar Pleasant and Unipolar Unpleasant IAT effects showed modest positive intercorrelation with each other, suggesting that these attitudes are not diametrically opposed but (very weakly) complimentary. Although we cannot rule out the possibility that this negligible inter-correlation reflects low reliability between the two conditions, the results highlight the potential importance of considering Pleasant and Unpleasant attitudes separately. Third, there was modest convergence between explicit and implicit measures of social attitudes. Notably, the implicit Unipolar Unpleasant IAT seemed to explain responses on both pleasant and unpleasant explicit attitude measures. The associations between IAT performance and social behavior were generally negligible.

**Table 2**  
Partial correlations between implicit and explicit measures of social attitudes and social behavior.

	Pleasant social attitudes	Unpleasant social attitudes	Bipolar attitudes
<i>Implicit social attitudes</i>			
Pleasant social IAT	1.00	.18*	.23**
Unpleasant social IAT	–	1.00	–.09
Bipolar IAT	–	–	1.00
<i>Explicit social attitudes</i>			
Unipolar pleasant social	.02	–.18*	.10
Unipolar unpleasant social	–.03	.16*	–.11
Bipolar social	.03	.00	.18*
<i>Social behavior</i>			
Spend time on the phone with family <sup>a</sup>	.18*	.09	.19*
Spend time with family <sup>a</sup>	–.08	.01	–.03
Spend time with a friend <sup>a</sup>	.05	–.11	.08
Spend time on the phone with a friend <sup>a</sup>	–.07	.07	–.11
Make plans with a friend <sup>a</sup>	.07	.05	–.15
Spend time with a boyfriend/girlfriend <sup>a</sup>	–.01	–.01	–.07

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

<sup>a</sup> Increasing scores reflect increased social behavior. Spearman's correlations were used to compensate for ordinal scale data.

Two concerns are raised from this study. First, the use of three IATs may have produced odd priming or habituation effects within subjects. As evidence of this, there was a fairly robust order effect for unipolar versus bipolar administration. Second, the measure of social behavior was relatively insensitive in that it was based solely on self-report assessments of global social behaviors. This may explain why the correlations between the IATs and social behavior were, for the most part, negligible. In the second study, we addressed these concerns by employing only the Unipolar IATs and examining the relationship between IAT performance and social behavior during a controlled laboratory task. Our goal was to replicate some of the findings from the first study (e.g., pleasant unipolar IAT effect) as well as to examine the link between IAT performance in the context of a more sensitive assessment of social behavior.

#### 5. Study 2 methods

##### 5.1. Participants

Data for 213 university students (61 men and 152 women) recruited from the same public university were examined in this study. This group included 172 Caucasians, 19 African-Americans, 11 Asian-Americans, four Hispanic Americans and seven "others". The average age of the group was 20.73 (SD = 3.78. Range = 18–60). Data for two subjects were not included in this report (see below). This study was approved by the appropriate Institutional Review Board and all subjects provided written informed consent prior to beginning the study.

##### 5.2. IAT

Separate Pleasant and Unpleasant Unipolar IATs were employed using the same specifications and stimuli in Study 1.

##### 5.3. Laboratory-based measure of social behavior

Natural speech was procured during a laboratory verbalization procedure. Subjects were seated in front of a computer monitor and could not see the interviewer. A headmount microphone was affixed to the subject and they were asked to talk for 90 s. They were allowed to talk about anything that came to mind so long as it involved social relationships in some capacity. This procedure has been used extensively in our lab to understand affective processes (e.g., Cohen, Minor, Najolia & Hong, in press).

The digitally-recorded narratives were carefully transcribed by trained research assistants, double-checked for accuracy, and then analyzed using the Lexical Inquiry and Word Count software (LIWC; Pennebaker, 2001). The LIWC program processes text files one word at a time, matching the base form of words to a "dictionary" of over 2290 words stems. Word stems are organized into 83 categories. A frequency count of the total instances of target words from each category is yielded. This count is divided by the total number of words to control for individual differences in verbosity. Thus, scores reflect a percentage of words in that category. For this study, we evaluated the narratives in terms of three broad categories, including "social references" – direct reference to self, others and social processes, "narrative form" – use of words indicative of collective or exclusive relationships, and "affective tone" – words with overt pleasant or unpleasant valence. Social references involved the frequency of self-references, either in first person singular (e.g., me, I) or plural (e.g., we, us) and also total social processes (e.g., talk, they). We examined narrative form in terms of negating (e.g., no, never), inclusive (e.g., with, include) and exclusive (e.g., without, exclude) words. Affective tone refers to either

pleasant (e.g., happy, joy) or unpleasant (e.g., hate, fear) words. With respect to the psychometric properties of the LIWC, each of these categories has shown modest convergence with trained undergraduate raters as well as good internal consistency (see Pennebaker, 2001). Moreover, these variables have been important for understanding a wide range of psychopathological disorders. Data for two subjects were excluded for failing to produce the recommended minimum 100 words for analysis (as in Pennebaker, 2001).

#### 5.4. Analyses

The data were analyzed in two steps. First, we sought to replicate the findings from the first study of a Unipolar Pleasant IAT effect. Second, we examined the relationship between the pleasant and unpleasant IAT scores and the laboratory lexical expression variables. We predicted that Unpleasant IAT scores would be associated with more self-references, negating words, exclusion and negative emotion words and fewer group-references, social processes, exclusion and positive emotion words. Conversely, we predicted that the Pleasant IAT scores would show the opposite pattern of associates. All variables examined here were normally distributed (skew values < 1.0).

## 6. Results

### 6.1. Unipolar IAT effects: replication

Error rates within the Unipolar Pleasant (.06 ± .06 and .07 ± .07 for the Pleasant and Neutral categorizations respectively) and Unipolar Unpleasant (.06 ± .07 and .07 ± .07 for the Unpleasant and Neutral categorizations respectively) were similar to Study 1, as were the reaction times within the Unipolar Pleasant (944.38 ± 387.36 and 1042.48 ± 457.06 for the Pleasant and Neutral categorizations respectively) and Unipolar Unpleasant (981.68 ± 397.46 and 1006.62 ± 434.15 for the Unpleasant and Neutral categorizations respectively). The IAT effect scores were also relatively similar to those computed in Study 1 and both the Unipolar Pleasant ( $d = .23$ ,  $t[209] = 7.55$ ,  $p < .001$ ) and the Unipolar Unpleasant ( $d = .06$ ,  $t[209] = 2.10$ ,  $p < .05$ ) scores were significantly greater than zero. These values were significantly different from each other (omnibus  $F = 14.79$ ,  $p < .001$ ). There were no order effects ( $t[209]$ 's < 1.29,  $p$ 's > .19).

### 6.2. Verbal expression during the laboratory procedure

Subjects produced, on average, 228.94 (SD = 50.45) words. As hypothesized, increasing Unipolar Unpleasant effect scores corresponded to fewer first-person plural and inclusive words and more negating words. In contrast to our expectations, pleasant Unipolar Pleasant IAT scores were not significantly associated with any of the expression variables. However, none of these correlations were statistically different between unipolar pleasant and unpleasant conditions using Fisher  $r$ -to- $z$  transformations (all  $p$ 's > .05) (Table 3).

## 7. Discussion

There were two important findings from study 2. First, the effect for the Unipolar Pleasant IAT was replicated. Second, unpleasant but not pleasant social attitudes corresponded to modest, but detectable communicative idiosyncrasies when discussing social relationships. Of note, individuals high in implicit unpleasant social attitudes tended to make fewer collective social references and employ fewer inclusive words. They also employed more negating

**Table 3**

Correlations between implicit unipolar IAT measures and lexical analysis measures.

	Pleasant social associations	Unpleasant social associations
Pleasant social assns	1.00	.10
Unpleasant social assns	–	1.00
Laboratory lexical analysis variables		
<i>Social references</i>		
1st Person singular (e.g., I)	.10	.07
1st Person plural (e.g., we)	–.07	–.14*
Total social Processes (e.g., talk)	.02	.05
<i>Narrative form</i>		
Negating Words (e.g., no, never)	.03	.14*
Inclusive words (e.g., with)	–.08	–.14*
Exclusive words (e.g., without)	.02	.09
<i>Affective tone</i>		
Positively valenced words (e.g., joy)	.01	–.03
Negatively valence words (e.g., hate)	.03	.13

\* $p < .05$ .

words. However, the overall use of pleasant or unpleasant valenced words within the sample was not associated with any IAT effects. It is important to acknowledge that these correlations are relatively weak, and correspond to the lower limits of a small effect size defined by Cohen (1988). Nonetheless, these data provide qualified convergent validity for the Unipolar IAT as a measure of individual differences in social attitudes/behavior.

## 8. General discussion

There are several important advances from the current set of studies. The first involves the promise of an improved technology for understanding implicit attitudes. Traditional IATs are dependent on measuring constructs in the context of diametrically opposed attitudes (e.g., good versus bad), which seems to be an overly simplistic assessment of many individuals' attitudes. When the present study employed a traditionally formatted IAT to examine attitudes towards social relationships, subjects demonstrated a detectable IAT effect. However, it was unclear whether subjects were associating social relationships with "pleasant" attitudes, with "unpleasant" attitudes, or some combination of the two. When the pleasant and unpleasant attitudinal labels were examined independently, it was revealed that pleasant categorizations were driving this relationship in that subjects, as a group, associated social relationships with pleasant but not unpleasant attitudinal labels. The present findings are consistent with those of Houben and Wiers (2006) and Jajodia and Earleywine (2003) suggesting that the Unipolar IAT can be helpful for deconstructing attitudes.

The overlap between implicit and explicit measures of social attitudes, although statistically significant was modest. Thus, these measures appear to be capturing fundamentally distinct phenomenon. This finding is consistent with what has been observed in the larger IAT literature – that explicit and implicit measures show limited convergence (e.g., McConnell & Leibold, 2001).

Implicit unipolar measures showed modest correlations to some speech characteristics during a laboratory condition but not to more global measure of social behavior. These speech characteristics were associated with Unpleasant but not Pleasant IAT scores, and, although these differences were not statistically significant, they further support the notion that these valences should be considered orthogonal rather than as opposing ends of a continuum (Cacioppo et al., 1997; Davidson, 2003). The weak magnitude of these correlations limits conclusions that can be

drawn about the importance of the Unipolar Unpleasant IAT. However, these data provide some evidence that it reflects a potentially important individual difference measure.

While the present data are promising, it is worth considering improvements that could potentially improve the sensitivity of the Unipolar IAT. The chief limitation revealed in this study is that the effects for subjects as a group were quite small. This is of particular concern given that we did not control for Type I errors in the correlational analyses. It is worth noting that the Unipolar IAT assessed a broad range of social domains that may potentially be heterogeneous in nature. For example, individuals may harbor unpleasant attitudes in some social domains (e.g., parents, teachers) but not all (e.g., friends, significant others). Social domains may be better considered separately. Unfortunately, the present studies did not employ enough stimuli to conduct these analyses. Changing the stimuli so that they more profoundly activate social attitudes may also improve sensitivity. Some of the stimuli used in this study may have been perceived as ambiguous or inapplicable by some subjects (e.g., “fiancé”, “roommate”). A more sophisticated methodology might employ individually-tailored stimuli such as names of significant others or pictures from a subjects personal photo-albums. Third, the use of a “neutral” category appears to attenuate the IAT effect. Although this format has shown promise in prior studies (Houben & Wiers, 2006; Jajodia & Earleywine 2003), it is worth considering the “single category” IAT (SC-IAT) in future research. The SC-IAT Karpinski, A., and Steinman, R. B. (2006) is a recent innovation that would eliminate the need for the neutral category.

In sum, the present data support the uncoupling of pleasant and unpleasant valences to understand social attitudes. The Unipolar IAT is a promising implicit methodology for this endeavor. Application of this measure may yield important insights about the nature of social impairments in mental disorders and about social attitudes in the general population.

#### Appendix 1. Stimuli used in this study.

Category1	Category2	Attitude1	Attitude2	Attitude3
People	Places	Pleasant	Neutral	Unpleasant
Friend	Texas	Paradise	Cabinet	Lonely
Mother	Idaho	Lovely	Elbow	Cruel
Teacher	Oregon	Triumphant	Scissors	Anguish
Roommate	Utah	Miracle	Appliance	Cancer
Lover	Missouri	Joyous	Sphere	Torture
Buddy	Georgia		Kettle	
Father	Kansas		Fork	
Instructor	Maine		Month	
Classmate	Ohio		Market	
Fiancé	Virginia		Column	

#### References

- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of mental disorders (DSM-IV)* (4th ed.). Washington, DC: American Psychiatric Press.
- Bernard, M., Jackson, C., & Jones, C. (2006). Written emotional disclosure following first-episode psychosis: effects on symptoms of post-traumatic stress disorder. *British Journal of Clinical Psychology, 45*, 403–415.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli instruction manual and affective ratings*. Technical report C-1. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: the case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*, 3–25.
- Campbell, W. K., Bosson, J. K., Goheen, T. W., Lakey, C. E., & Kernis, M. H. (2007). Do narcissists dislike themselves ‘deep down inside?’. *Psychological Science, 18*, 227–229.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, A. S., Alpert, M., Nienow, T. M., Dinzeo, T. J., & Docherty, N. M. (2008). Computerized analysis of negative symptoms in schizophrenia. *Journal of Psychiatric Research, 42*, 827–836.
- Cohen, A. S., & Minor, K. Emotional experience in schizophrenia revisited: meta-analysis of laboratory studies. *Schizophrenia Bulletin*, in press.
- Cohen, A. S., Minor, K. S., Najolia, G., & Hong, S. L. (2009). A laboratory-based procedure for measuring emotional expression from natural speech. *Behavior Research Methods, 41*, 192–203.
- Cohen, A. S., St-Hilaire, A., Aakres, J. M., & Docherty, N. M. (2009). Understanding anhedonia in schizophrenia through lexical analysis of natural speech. *Cognition & Emotion, 569*–586.
- Craeynest, M., Crombez, G., Koster, E. H. W., Haerens, L., & De Bourdeaudhuij, I. (2008). Cognitive-motivational determinants of fat food consumption in overweight and obese youngsters: the implicit association between fat food and arousal. *Journal of Behavior Therapy and Experimental Psychiatry, 39*, 354–368.
- Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience, 16*, 1717–1729.
- Davidson, R. J. (2003). Affective neuroscience and psychophysiology: toward a synthesis. *Psychophysiology, 40*, 655–665.
- De Houwer, J. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: strengths and limitations. *Journal of Behavior Therapy and Experimental Psychiatry, 33*, 115–133.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. Implicit measures: a normative analysis and review. *Psychological Bulletin*, in press.
- Franck, E., De Raedt, R., Dereu, M., & Van den Abbeele, D. (2007). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 75–85.
- Gortner, E. M., Rude, S. S., & Pennebaker, J. W. (2006). Benefits of expressive writing in lowering rumination and depressive symptoms. *Behavior Therapy, 37*, 292–303.
- 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.
- Hamilton, D. L., Sherman, S. J., & Ruvalo, C. M. (1990). Stereotype-based expectancies: effects on information processing and social behavior. *Journal of Social Issues, 46*, 35–60.
- Hopwood, C. J., & Morey, L. C. (2007). Psychological conflict in borderline personality as represented by inconsistent self-report item responding. *Journal of Social & Clinical Psychology, 26*, 1065–1075.
- Houben, K., & Wiers, R. W. (2006). Assessing implicit alcohol associations with the Implicit Association Test: fact or artifact? *Addictive Behaviors, 31*, 1346–1362.
- Jajodia, A., & Earleywine, M. (2003). Measuring alcohol expectancies with the Implicit Association Test. *Psychology of Addictive Behaviors, 17*, 126–133.
- de Jong, (2002). Implicit self-esteem and social anxiety: differential self-favouring effects in high and low anxious individuals. *Behaviour Research and Therapy, 40*, 501–508.
- Karpinski, A., & Steinman, R.B. (Jul 2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*, 16–32.
- Kelley, J. E., Lumley, M. A., & Leisen, J. C. (1997). Health effects of emotional disclosure in rheumatoid arthritis patients. *Health Psychology, 16*, 331–340.
- Lane, K. A., Banaji, M. R., Nosek, B. A., Greenwald, A. G., Wittenbrink, B., & Schwarz, N. (2007). Understanding and using the Implicit Association Test: IV: what we know (so far) about the method. In B. Wittenbrink, & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 59–102). New York, NY, US: Guilford Press.
- Lehman, A. (1995). *Evaluating quality of life for persons with severe mental illness: Assessment toolkit*. Cambridge, MA: The Evaluation Center at Health Services Research Institute.
- Lepore, S. J. (1997). Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. *Journal of Personality and Social Psychology, 73*, 1030–1037.
- Lyons, E. J., Mehl, M. R., & Pennebaker, J. W. (2006). Pro-anorexics and recovering anorexics differ in their linguistic Internet self-presentation. *Journal of Psychosomatic Research, 60*, 253–256.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435–442.
- Murray, E. J., & Segal, D. L. (1994). Emotional processing in vocal and written expression of feelings about traumatic experiences. *Journal of Traumatic Stress, 7*, 391–405.
- Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: investigating participant demand characteristics. *Journal of General Psychology, 135*, 151–165.
- Orsillo, S. M., Batten, S. V., Plumb, J. C., Luterek, J. A., & Roessner, B. M. (2004). An experimental study of emotional responding in women with posttraumatic stress disorder related to interpersonal violence. *Journal of Traumatic Stress, 17*, 241–248.
- Pennebaker, J. W. (2001). *Linguistic inquiry and word count*. Lawrence Erlbaum Associates.
- Priester, J. R., & Petty, R. E. (2001). Extending the bases of subjective attitudinal ambivalence: interpersonal and intrapersonal antecedents of evaluative tension. *Journal of Personality and Social Psychology, 80*, 19–34.

- Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84*, 754–775.
- Rollnick, S., & Miller, W. R. (1995). What is motivational interviewing? *Behavioural and Cognitive Psychotherapy, 23*, 325–334.
- Simpson, J. A., Rholes, W. S., Campbell, L., Tran, S., & Wilson, C. L. (2003). Adult attachment, the transition to parenthood, and depressive symptoms. *Journal of Personality and Social Psychology, 84*, 1172–1187.
- Tanner, R. J., Stopa, L., & De Houwer, J. (2006). Implicit views of the self in social anxiety. *Behaviour Research and Therapy, 44*, 1397–1409.
- Tull, M. T., Medaglia, E., & Roemer, L. (2005). An investigation of the construct validity of the 20-Item Toronto Alexithymia Scale through the use of a verbalization task. *Journal of Psychosomatic Research, 59*, 77–84.
- Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychological Review, 96*, 234–254.