

THE IMPLICIT RELATIONAL ASSESSMENT PROCEDURE AS A MEASURE OF SELF-ESTEEM

C. Alix Timko

University of Pennsylvania

Towson University

Erica L. England, James D. Herbert, Evan M. Forman

Drexel University

Two studies were conducted to pilot the Implicit Relational Assessment Procedure (IRAP) in measuring attitudes toward the self: one related to body image specifically and another assessing the broader construct of self-esteem. Study 1 utilized the IRAP with female college students to examine self-referential beliefs regarding body image. Results revealed positive associations between self-referential beliefs on the IRAP and explicit measures of body image satisfaction and acceptance, likelihood of dieting, and internalization of the thin ideal. In Study 2, an IRAP measuring general self-esteem revealed positive correlations between IRAP performance and explicit measures of psychological functioning and negative correlations between the IRAP and psychopathology. Results are discussed in terms of the potential utility of this theoretically grounded implicit measure in assessing self-concept.

Key words: implicit assessment, body image, self-esteem, IRAP, measurement

Implicitly measured self-esteem has been found to be regularly and weakly correlated with explicitly measured self-esteem (Greenwald et al., 2002; Olson, Fazio, & Hermann, 2007). The implicit measurement of self-esteem has drawn some criticism in that it assumes that explicit and implicit measures assess two fundamentally different forms of self-esteem. Thus, it is assumed that individuals either have a distinct, unconscious sense of self-esteem or that they consciously misrepresent how they feel about themselves (Tafarodi & Ho, 2006). More recent research has indicated not that implicit measurement taps into unconscious self-esteem, but rather that individuals may consciously overreport their levels of self-esteem (Olson et al., 2007) or that implicit measurement may simply be assessing a different aspect of conscious self-esteem (Fazio & Olson, 2003; Tafarodi & Ho, 2006).

The authors thank Courtney Darcy, Michael Haas, Amy Neal, and Lily Brown for their assistance in data collection and entry. The authors also wish to thank Dermot Barnes-Holmes and an anonymous reviewer for their comments on an earlier version of this manuscript.

Correspondence concerning this article should be addressed to C. Alix Timko, Department of Behavioral and Social Sciences, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104. E-mail: a.timko@usp.edu

Thus, implicit measurement may be useful as *another* method by which to measure self-esteem, as opposed to a method by which to assess a qualitatively different (unconscious) type of self-esteem. Given the potential relationships between implicitly measured self-esteem, behavior, and therapeutic outcomes, it is necessary to have an implicit measure of self-esteem that is solidly grounded in theory. A firm theoretical grounding provides the basis for specific, testable hypotheses. The Implicit Relational Assessment Procedure (IRAP) may represent such a measure.

Implicit Relational Assessment Procedure

Unlike the Implicit Association Test (IAT; Greenwald, McGhee, & Schwarz, 1998) and other previously constructed implicit measures, the IRAP is grounded in a contemporary behavior-analytic theory of language and cognition: relational frame theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). The IRAP builds upon the IAT by employing relational terms, which specify the relationships between the stimuli to which the participant is responding. According to the relational elaboration and coherence (REC) model (Barnes-Holmes, Murphy, Barnes-Holmes, & Stewart, 2010; Cullen, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009), an RFT-based model designed to clarify the relationship between explicit measures and implicit responding on the IRAP, responses on implicit and explicit measures are *both* due to relational responding that is dependent on an individual's learning history. Reaction times on the IRAP are due to an immediate relational response. That is, given the need to respond to items presented during the IRAP quickly and correctly, there is insufficient time to engage in logical or elaborate processing. Thus, implicit measurement, according to the REC model, reflects the relative strength of immediate relational responding. In completing explicit measures, participants have more time to engage in elaborate relational responding. Furthermore, explicit relational responding is subject to control by certain features of the environmental context that are absent in implicit relational responding (e.g., demand characteristics). Barnes-Holmes, Murphy, and colleagues (2010) note that the REC model may, at first blush, resemble more common dual-system models. In dual-system models, attitudes are assessed and stored either associatively (implicitly) or via a series of propositions or rules (explicitly). These separate systems are thought to be uniquely accessed depending on the type of assessment used. Though rules may play a role in the more elaborate relational responding that can occur when explicit assessment is used, according to the REC model the *process* is the same as the one that occurs when implicit assessment is used.

Research Using the IRAP

In the seminal IRAP publication, Barnes-Holmes and colleagues (2006) presented three IRAP experiments: matching positive or negative target words to the samples "pleasant" and "unpleasant," assessing attitudes toward individuals with autism, and assessing participants' preferences for their own ethnic group. Since these experiments were reported, other studies have used the IRAP to assess beliefs and attitudes in a number of domains, including racial stereotypes (Barnes-Holmes, Murphy, et al., 2010; Drake et al., 2010; Rashid, Haas, & Timko, 2007), urban versus rural life (Barnes-Holmes, Waldron, Barnes-Holmes, & Stewart, 2009), sexual orientation (Cullen & Barnes-Holmes, 2008), and meat eating versus vegetarianism (Barnes-Holmes, Murtagh,

Barnes-Holmes, & Stewart, 2010). One study found that IRAP response latency predicted performance on measures of intelligence (O'Toole & Barnes-Holmes, 2009). Importantly, McKenna, Barnes-Holmes, Barnes-Holmes, and Stewart (2007) reported that participants were unable to "fake" their performance on the IRAP, even when instructed to do so, whereas the IAT has been found to be susceptible to falsified performance (Kim, 2003).

Recently, researchers have begun to use the IRAP to examine self-referential concepts, including self-esteem. The first such study assessed self-esteem in both undergraduates and convicted prisoners. There were two groups of prisoners, one set residing in the main block of a medium-security prison and another set in a privileged, lower security "open area" of the prison (Vahey, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009). Trials consisted of a sample relation ("similar" or "opposite") presented with positive and negative attributes (e.g., "good" and "bad"). Response options were "[*participant's name*]" and "not [*participant's name*]." Compared to main-block prisoners, students and open-area prisoners demonstrated shorter response latencies on consistent (i.e., self-positive) test blocks relative to inconsistent (i.e., self-negative) test blocks. IRAP results were moderately positively correlated with explicit measures of self-esteem, which indicated significantly lower levels of self-esteem in main-block prisoners compared to students and open-area prisoners. In a similar IRAP, Drake (2009) investigated the relationship between latencies on a self-relevant IRAP and explicit measures of self-esteem and general experiential acceptance. Drake used the sample relations "I am" and "I am not" with the response options "true" and "false." Targets were related to overall self-esteem (e.g., "perfect" and "inadequate"). Drake found that participants responded more quickly when affirming self-positive and denying self-negative statements than when affirming self-negative and denying self-positive statements. However, there was no correlation between the IRAP performance and explicitly measured self-esteem. These studies, while preliminary, lend support to further investigation of the IRAP as an implicit tool for examining self-esteem.

Purpose of Current Studies

The two studies presented in this article investigated the use of the IRAP as an implicit measure of self-esteem. The studies used the sample phrases "I am" and "I am not," as these reflect the natural viewpoint of the participant and also made the IRAP easier to administer, given that individual participants' names did not need to be programmed into the task. Much of the research conducted on the implicit measurement of self-esteem includes two categories (or samples): one that reflects the self and one that reflects the other (e.g., Vahey et al., 2009). These implicit measures, therefore, assess both attitudes toward the self and attitudes toward others (Karpinski, 2004; Yamaguchi et al., 2007) or, in the case of the IRAP, relational responding in regard to the self and with respect to someone (or some concept) outside of the self. The current studies used two samples that exclusively reflect responding in regard to the self, thereby eliminating the "other" comparison. Furthermore, much of the IRAP research published to date (including those studies focused on self-esteem) has originated from a single research laboratory. In order to establish the utility of the IRAP, data from other laboratories must be able to replicate published findings.

Domains of self-esteem chosen for investigation were based on self-generated, self-referring statements by participants in a study by Masuda, Hayes, Sackett, and Twohig (2004). The most commonly generated self-statements concerned physical appearance (e.g., "fat") and intelligence (e.g., "dumb"). Thus, these domains seemed particularly relevant for an examination of various aspects of self-esteem. The current studies utilized the IRAP to examine self-esteem in the areas of weight, intelligence, and general self-efficacy. It was hypothesized that participants would produce significantly shorter response latencies for self-positive than for self-negative statements. Consistent with an RFT understanding of self-concept, the IRAP was conceptualized as an implicit measure of self-esteem as opposed to a measure of a distinct construct of implicit self-esteem. Therefore, it was hypothesized that the IRAP would be correlated with measures of pathology and beliefs about the self such that higher self-esteem (measured implicitly) would be associated with greater overall health and satisfaction (measured explicitly).

Study 1

The first study utilized the IRAP to assess self-esteem in the domain of body image. Body image was chosen as the specific domain to study given that a woman's perception of her body can be central to her sense of self. Furthermore, body-image satisfaction has been shown to be related to overall self-esteem and self-efficacy (Ip & Jarry, 2008; Mercurio & Landry, 2008; Verplanken & Velsvik, 2008). Given that previous research on the measurement of implicit beliefs about the self has found a weak correlation between implicit and explicit measures of self-esteem (Greenwald et al., 2002; Olson et al., 2007; Yamaguchi et al., 2007), it was hypothesized that women would have positive D_{IRAP} scores and that the size of the D_{IRAP} score would be significantly (but weakly) associated with greater body-image satisfaction and body-image acceptance and less internalization of the thin ideal. Furthermore, it was hypothesized that there might be differences in body satisfaction (measured implicitly) when comparing dieters to non-dieters. Though not a clinical population, those who identify as dieting to lose weight often have greater body dissatisfaction and tend to internalize the thin ideal more than those who do not self-identify as dieting (Faherty, Timko, & Kalodner, 2009; Timko, Martin, & Darcy, 2010). If the IRAP is able to measure body satisfaction implicitly, there should be significant differences between these groups in terms of their overall D_{IRAP} score.

Participants

A total of 68 women were recruited as part of a larger study on the impact of the media on body image (Timko et al., 2010). Of these, 54 completed the IRAP (50 of whom provided useable data). All participants were recruited from undergraduate psychology courses and ranged in age from 18 to 24 years ($M = 19.10$, $SD = 1.53$). Given that body image is grounded, to some degree, in culture and that rates and degree of dissatisfaction have been shown to vary across ethnic groups (e.g., Roberts, Cash, Feingold, & Johnson, 2006; Vaughan, Sacco, & Beckstead, 2008), only Caucasian women were included in the study. Body mass index [BMI; weight in kg/(height in cm²)] ranged from 17.37 to 26.63 ($M = 21.85$, $SD = 2.39$). The sample

included 20 women who indicated they were not dieting (40%), 21 who self-identified as dieting to lose weight (42%), and nine who self-identified as dieting to maintain weight (18%).

Measures

Implicit Relational Assessment Procedure (IRAP). The 2006 version of the IRAP program developed by Barnes-Holmes was used in this study (the program can be downloaded from http://psychology.nuim.ie/IRAP/IRAP_1.shtml). In the IRAP, 24 trials are presented in a series of alternating blocks of consistent or inconsistent trials. Participants are required to respond with the experimenter-designated “correct” relation as quickly and accurately as possible for each trial in a test block. If an incorrect response option is chosen, a large red “X” appears on the screen and the participant is prompted to choose the correct response option. In the current study, the “d” key corresponded to “true” and the “k” key to “false.” These response options remained static throughout the entire IRAP. See Figure 1 for a pictorial representation. The IRAP program automatically records the response latency (the time it takes for the participant to press the correct key) for each trial, and an overall difference score between the mean response latency on consistent versus inconsistent test blocks is calculated. The conclusion drawn from the IRAP is that one’s immediate relational responding is represented by the type of trial on which one shows significantly shorter response latencies.

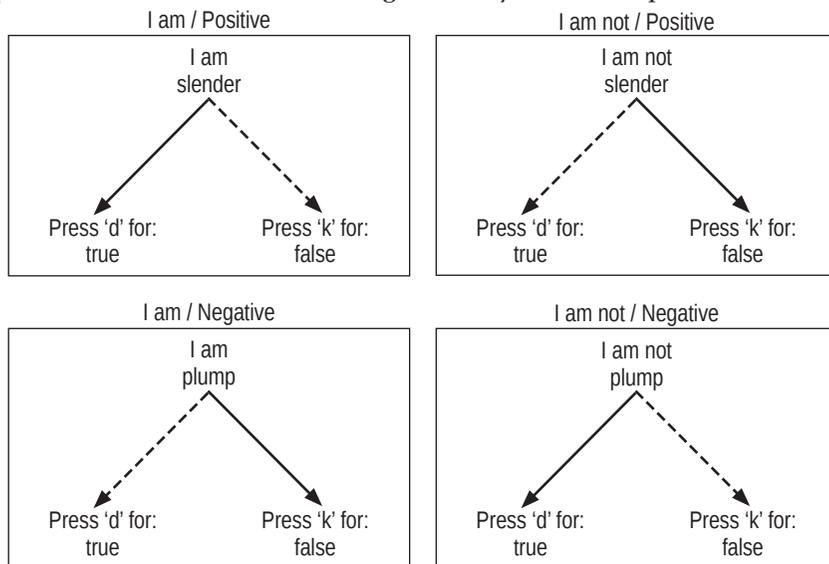


Figure 1. Example screenshots of the four IRAP trial types used in the current studies. The sample phrase (“I am” or “I am not”) appeared at the top of the screen, while a target word (e.g., “slender,” “plump”) appeared in the center. At the bottom corners of the screen were the response options “true” and “false.” All elements appeared simultaneously at the beginning of each trial. Solid arrows indicate consistent responses for each trial type; dashed arrows indicate inconsistent responses for each trial type (arrows did not appear on screen). A correct response prompted the next trial; an incorrect response resulted in a red “X” on the screen and required the participant to choose the correct response before progressing.

The IRAP begins with a series of at least two (and no more than eight) practice blocks (one consistent, one inconsistent). During this practice phase of the IRAP in the current study, participants were expected to achieve a minimum correct response rate of 65% and median response latency less than 3,000 ms in each block.¹ At the end of each block, participants were given feedback regarding their speed and accuracy. If participants were able to achieve these criteria, they proceeded to the test phase of the IRAP. The test phase was essentially identical to the practice blocks, though feedback regarding speed and accuracy was missing. If participants were unable to achieve the minimum criteria, they were given the opportunity to continue for up to eight exposures to the practice blocks. Participants who were unable to attain the required accuracy and speed level within the eight allotted practice blocks did not continue to the test blocks, and the IRAP task automatically ended after the eighth practice block.

In the IRAPs used in the current study, the task required the participant to confirm or deny her belief in self-referential statements on each trial. There were four trial types: "I am [*positive word*]," "I am not [*positive word*]," "I am [*negative word*]," and "I am not [*negative word*]." See Figure 1 for a pictorial representation. The presence of these four trial types in each test block of the IRAP is one of its primary strengths, as it allows for investigation of specific types of relations and can provide a more in-depth analysis of participants' responding.

At the beginning of the IRAP, standardized instructions (available upon request from the authors) appeared on the screen, describing the response procedure and giving examples of each of the four different types of trials. The instructions also directed the participant to consider the phrases "I am" and "I am not" in reference to herself. An experimenter sat with each participant during review of the instructions and answered any questions the participant had before the IRAP began. It was made clear to participants that sometimes they would be required to affirm statements with which they might not agree and at other times they would be asked to affirm statements with which they did agree. This was presented as part of the computer task.

D_{IRAP} algorithm. The primary raw datum for the IRAP is response latency, defined as the amount of time (in milliseconds) from the beginning of a trial until a correct response is made. The D_{IRAP} score reflects the difference in response latency between consistent and inconsistent blocks; therefore a D_{IRAP} score that is significantly different from zero indicates that there was, in fact, a significant difference between response latencies in consistent versus inconsistent blocks. A positive D_{IRAP} score indicates that respondents took significantly longer to answer correctly during inconsistent blocks, whereas a negative D_{IRAP} indicates

1 Recent research using the IRAP has employed a more stringent latency criterion, requiring participants to answer correctly within 2,000 ms. Furthermore, the minimum response criterion required in the current studies (65% correct) is lower than most recent IRAP studies. The typical percentage correct required is more commonly 80–85%. More stringent criteria may result in a larger D_{IRAP} effect and a greater reliability of the IRAP (Barnes-Holmes, Murphy, et al., 2010).

that respondents took significantly longer to answer correctly in consistent test blocks.²

Stimuli. The words chosen for this study had been rated previously according to their pleasantness on a 7-point scale (1 = *highly unpleasant*; 7 = *highly pleasant*) by a sample of 117 undergraduates. The words chosen represent a range of scores on the pleasant-unpleasant scale. Four words were chosen to be consistent with each sample (see Figure 2). An attempt was made to avoid including only words that were strongly negatively or positively biased in order to reduce the valence of the words impacting the effect. To that end, one word each was chosen to represent “mildly unpleasant/pleasant” and “very unpleasant/pleasant.” Two words each were chosen to represent “unpleasant/pleasant.” The words and their mean pleasantness ratings were as follows: *slender* (5.19), *toned* (5.86), *skinny* (4.13), *gorgeous* (6.69), *plump* (2.63), *bony* (2.40), *chunky* (2.11), and *obese* (1.32). Each target word was entered into the IRAP program twice, creating a total of eight target words consistent with each sample (16 total stimuli). An attempt was made to choose words specifically related to body shape as opposed to facial or overall appearance. In the case of *gorgeous*, there were no words that specifically addressed the body that were rated as very pleasant; hence, it was chosen as it was the only viable candidate.

Sample 1 I am		Sample 2 I am not	
Response Option 1 True		Response Option 2 False	
Targets deemed consistent with Sample 1		Targets deemed consistent with Sample 2	
Study 1	Study 2	Study 1	Study 2
Slender	Smart	Plump	Stupid
Toned	Attractive	Bony	Ugly
Skinny	Friendly	Chunky	Shy
Gorgeous		Obese	

Figure 2. Sample and target stimuli and response options for Studies 1 and 2. Each target word was entered into the IRAP program two times, resulting in a total of eight targets consistent with each sample in Study 1 and six consistent with each sample in Study 2.

2 The D_{IRAP} algorithm is calculated as follows: (1) Only response latency data from test blocks are used; (2) latencies above 10,000 ms are eliminated from the data set; (3) all data for a participant are removed if he or she produces more than 10% of test-block trials with latencies less than 300 ms; (4) 12 standard deviations for the four trial types are computed: four for the response latencies from Test Blocks 1 and 2, four from the latencies from Test Blocks 3 and 4, and a further four from Test Blocks 5 and 6; (5) 24 mean latencies for the four trial types in each test block are calculated; (6) difference scores are calculated for each of the four trial types, for each pair of test blocks, by subtracting the mean latency of the consistent block from the mean latency of the corresponding inconsistent block; (7) each difference score is divided by its corresponding standard deviation from step 4, yielding 12 D_{IRAP} scores, one score for each trial type for each pair of test blocks; (8) four overall trial-type D_{IRAP} scores, or IRAP effects, are calculated by averaging the scores for each trial type across the three pairs of test blocks; and (9) a single overall D_{IRAP} score is calculated by averaging the 12 trial-type D_{IRAP} scores from step 7.

Demographic information. This information was gathered for descriptive purposes and included age, gender, height, self-reported weight, and current dieting status.

Body Image Acceptance and Action Questionnaire (BIAAQ; Sandoz, Wilson, & Merwin, 2008). The BIAAQ is a 12-item self-report measure designed to assess an individual's level of experiential acceptance regarding her own body. Participants rate their agreement with various statements on a 7-point scale ranging from *never true* to *always true*. This scale exhibits strong internal consistency (Cronbach's alpha = 0.92) as well as construct validity (Sandoz et al., 2008; Faherty et al., 2009) and is negatively correlated with measures of body dissatisfaction and disordered eating (Faherty et al., 2009), such that high acceptance is associated with greater body satisfaction and less disturbed eating patterns.

Sociocultural Attitudes Towards Appearance-3 (SATAQ; Thompson, van den Berg, Roehrig, Guarda, & Heinberg, 2004). This scale is a 30-item self-report measure that assesses internalization of the "thin ideal" on a 5-point scale, ranging from *definitely disagree* to *definitely agree*. It reflects the media's influence on body image via television, magazines, and sports figures and has four subscales assessing general internalization, internalization of the athletic ideal, pressure to obtain the ideal, and information gleaned from the media in regard to the ideal. Cronbach's alpha is high across all subscales and ranges from 0.92 to 0.96. For the total measure, internal consistency is 0.96.

Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995). This 21-item scale assesses the states of depression, anxiety, and chronic overarousal. This measure has been shown to have high internal consistency (Cronbach's alpha ranging from 0.87 to 0.94; Antony, Bieling, Cox, Enns, & Swinson, 1998) and discriminant validity. It is correlated with more traditional measures of anxiety and depression. The 21 items are ranked on a 0 to 3 point scale, ranging from *Did not apply to me at all* to *Applied to me very much, or most of the time*.

Body Shape Questionnaire (BSQ; Cooper, Taylor, Cooper, & Fairburn, 1986). This scale measures a person's concern with his or her weight or shape. The 34-item scale exhibits good test-retest reliability (0.88; Rosen, Jones, Ramirez, & Waxman, 1996) as well as high internal consistency (Cronbach's alpha = 0.97). The current study employed an 8-item version of the scale (BSQ-8B; Evans & Dolan, 1993), which asks participants to rate how they have been feeling about their appearance over the past 4 weeks on a 6-point scale ranging from *never* to *always*. The version of the BSQ used in this study has a Cronbach's alpha of 0.87 (Evans & Dolan, 1993).

Feeling thermometer. Participants completed a feeling thermometer (adapted from Greenwald & Farnham, 2000) for each of the eight words per sample used on the IRAP. The feeling thermometer asks participants to rate how warmly they feel toward each word on a continuous vertical scale anchored at 0, with an upper bound of 99.

Procedure

Participants reported to the laboratory individually, completed self-report measures, and then completed the IRAP. A recent meta-analysis

indicated that the order of presentation of implicit and explicit measurement did not affect the correlations between those measures (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005).

Results

IRAP analyses. Of the 54 participants, four were excluded from the IRAP analyses. In these cases, one person was unable to complete it due to a computer malfunction (the computer experienced an unexpected reboot), and three failed to meet criteria for continuing past the practice blocks (greater than 65% correct and median latency less than 3,000 ms).

Feeling thermometer ratings. To ascertain whether participants viewed the eight IRAP target words as having the intended valence (positive or negative), the feeling thermometer ratings were examined. There was a significant main effect for target word, $F(7, 343) = 145.54$, $p < 0.001$, $\eta_p^2 = .75$. Post-hoc least significant difference (LSD) pairwise comparisons revealed that *gorgeous* and *toned* were rated equally pleasant, followed by *slender* and *skinny*. *Plump* and *bony* were rated equally negatively (and lower than all positive words); the lowest rated words were *chunky* and *obese*. Thus, participants interpreted the positive target words as positive and the negative words as negative. Pearson correlations showed no significant relationships between thermometer ratings and D_{IRAP} scores.

D_{IRAP} score analyses. The overall D_{IRAP} score ranged from -0.1609 to 0.77 , with a positive mean ($M = 0.2316$, $SD = 0.2255$), indicating that the average participant's immediate relational response expressed body satisfaction. Though the criterion for continuation in the study was a minimum of 65% correct across trials, participants ranged from 78% correct to 100% correct ($M = 93.49$, $SD = 5.56$), indicating that they were able to successfully complete the task. Consistent with previous IRAP research (Cullen & Barnes-Holmes, 2008; Drake et al., 2010), the current studies examined the IRAP at the level of individual trial types (i.e., the four possible combinations of the two sample phrases with the two types of target stimuli), allowing for examination of the data in greater detail. Four mean $D_{\text{IRAP-TRIAL TYPE}}$ scores were calculated. One-sample t tests indicated that the $D_{\text{IRAP-TRIAL TYPE}}$ scores for three out of the four trial types differed significantly from zero (see Figure 3). A repeated-measures ANOVA examining differences between the four trial types was conducted. The hypothesis of sphericity was not rejected ($p = 0.34$), and significant differences were found between trial types, $F(3, 147) = 13.84$, $p < .001$, $\eta_p^2 = 0.22$. Post-hoc LSD pairwise comparisons indicated several significant differences, namely, the $D_{\text{IRAP-TRIAL TYPE}}$ score for "I am [positive word]" was significantly higher than those for "I am [negative word]" and "I am not [negative word]." The $D_{\text{IRAP-TRIAL TYPE}}$ score for "I am [negative word]" was significantly lower than the score for "I am not [positive word]." The $D_{\text{IRAP-TRIAL TYPE}}$ score for "I am not [positive word]" was also significantly higher than that for "I am not [negative word]."

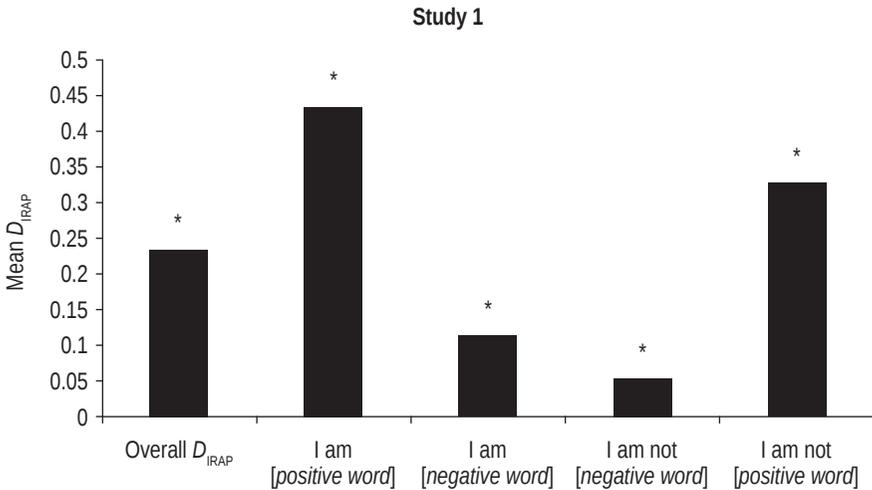


Figure 3. Overall D_{IRAP} and trial-type D_{IRAP} scores for Study 1. *Significantly different from zero ($p < .001$).

Relationship of D_{IRAP} to dieting status. Individuals who self-identify as dieting to lose weight tend to have higher BMIs than those who are not dieting (Timko & Perone, 2006; Timko, Perone, & Crossfield, 2006). Dieters also tend to have higher levels of body dissatisfaction and greater internalization of the thin ideal (Timko et al., 2009; Faherty et al., 2009). If the IRAP does, indeed, measure implicit beliefs about the body, it would be expected that those dieting to lose weight would be less satisfied with their body when body satisfaction is measured implicitly as well as explicitly. In a comparison of the dieters to nondieters, all had overall D_{IRAP} scores greater than zero [nondieters: $t(19) = 6.64$, $p < .001$; dieting to lose: $t(20) = 2.92$, $p = .008$; dieting to maintain: $t(8) = 3.73$, $p = .006$]. Those who identified as dieting to lose weight had significantly lower D_{IRAP} scores ($M = 0.1277$, $SD = .2003$) than those who identified as not dieting [$M = 0.3236$, $SD = 0.2181$; $F(2, 47) = 4.62$, $p = .02$]. Those who identified as dieting to maintain weight ($M = 0.2687$, $SD = 0.2172$) did not differ from either of the other groups.

Correlations between overall D_{IRAP} scores and self-report measures. For the purpose of elucidating the relationships between the overall D_{IRAP} and measures of body satisfaction, the overall and trial-type D_{IRAP} scores were entered into a correlation matrix with the self-report measures. The overall D_{IRAP} scores were significantly correlated with a number of the explicit variables (see Table 1). The D_{IRAP} and three of the $D_{IRAP-TRIAL\ TYPE}$ scores indicated positive body image; these scores correlated with explicit measures in the anticipated directions.

Table 1
Correlations Between Overall D_{IRAP} and $D_{IRAP-TRIAL\ TYPE}$ Scores and Self-Report Measures: Study 1

Self-report measure	Overall D_{IRAP} score	I am [positive word]	I am [negative word]	I am not [positive word]	I am not [negative word]
Body mass index	-.14	.09	.03	-.20	-.08
BIAAQ	.32*	.15	.16	.26	.19
SATAQ-Info	-.09	-.15	-.11	-.02	.10
SATAQ-Pressure	-.30*	-.16	-.15	-.23	-.14
SATAQ-General internalization	-.25	-.24	-.16	-.10	-.10
SATAQ-Athletic ideal internalization	-.12	-.29*	.01	-.05	.02
SATAQ total	-.23	-.24	-.14	-.12	-.04
DASS-Stress	-.10	-.19	.15	-.11	-.04
DASS-Anxiety	-.17	-.20	.17	.21	-.17
DASS-Depression	-.29*	-.29*	.01	-.16	-.24
BSQ	-.48**	-.32*	-.15	-.43**	-.24

Note. BIAAQ = Body Image Acceptance and Action Questionnaire (Sandoz et al., 2008); SATAQ = Sociocultural Attitudes Towards Appearance-3 (Thompson et al., 2004); DASS = Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995); BSQ = Body Shape Questionnaire (Cooper et al., 1986).

* $p < .05$. ** $p < .01$.

Discussion

This study assessed the use of the IRAP as an implicit measure of self-esteem in the domain of body image. The samples “I am” and “I am not” were chosen. In effect, this created a single-category implicit measure that assessed women’s opinions of their bodies in relation to what they believed themselves to be (or not to be). This is in opposition to standard implicit measures of self-esteem that typically measure an individual’s perception of self in relation to his or her perception of others (Karpinski, 2004), or in comparison to an overall anti-fat bias (e.g., Schwartz, Vartanian, Nosek, & Brownell, 2006). As predicted, the overall D_{IRAP} was positive (indicating overall body satisfaction). The trial-type analyses revealed that, overall, the tendency to affirm self-positive statements and deny self-negative statements was fairly robust, with only one trial type (“I am not [negative word]”) not significantly different from zero.

Examination of the correlations between the overall and trial-type difference scores revealed modest correlations with body dissatisfaction and higher levels of body acceptance, as well as lower depression and anxiety. Differences in the D_{IRAP} were also seen when comparing dieters to nondieters, in that dieters had lower D_{IRAP} scores than nondieters (i.e., they were less satisfied with their bodies). Given that dieters have higher levels of explicitly measured body dissatisfaction, the differences in D_{IRAP} scores between dieters and nondieters seem to support the IRAP as a measure of individual body image. Although the findings from this study support the IRAP as a measure of self-referential relational responding, the sample terms chosen

for this IRAP should be tested again in a different self-related domain (such as general self-esteem) in order to further validate the use of the IRAP as an implicit measure.

Study 2

Study 2 sought to broaden the domain of self-esteem beyond body image, to further assess the IRAP as an implicit measure of self-esteem. This study utilized the IRAP to assess self-concept in the areas of intelligence, physical appearance, and friendliness/shyness. It was hypothesized that response latencies affirming self-positive and denying self-negative statements would be shorter than latencies associated with denying self-positive and affirming self-negative statements, resulting in an overall positive D_{IRAP} score.

Participants

One hundred participants were recruited from undergraduate psychology courses. The mean age of the participants was 20.2 years (range = 18 to 36 years). The ethnic breakdown of the sample was as follows: 67% Caucasian, 21% Asian, 5% African-American, 2% Hispanic, and 5% other. The sample was evenly divided between men and women.

Measures

Implicit Relational Assessment Procedure (IRAP). The IRAP was presented as described in Study 1. The words chosen and their mean pleasantness ratings were as follows: *friendly* (6.57), *smart* (6.13), *shy* (3.64), *stupid* (1.70), and *ugly* (1.57). Although the word *attractive* did not appear on the list of words to be rated, the word *unattractive* received a mean rating of 2.34. Assuming that the word *attractive* would receive an approximately equivalent rating at the opposite end of the scale, the corresponding rating for *attractive* was estimated to be 5.66 [$7 - (2.34 - 1) = 5.66$], making it highly pleasant but less pleasant than *smart* and *friendly*. The word *attractive* was chosen despite its absence from the ratings list because it was deemed to be the most appropriate, gender-neutral word denoting positive physical appearance.

Quality of Life Index (QLI): Generic Version-III (Ferrans & Powers, 1985). The QLI measures life satisfaction in a variety of areas, such as health, family, emotional support, and education. Participants rate their satisfaction with and the perceived importance of each domain on a 6-point scale ranging from *very dissatisfied* to *very satisfied*. Internal consistency reliability for the QLI is high; Cronbach's alpha ranges from 0.84 to 0.98 across 26 studies (Ferrans & Powers, n.d.). Construct validity for the QLI is also high; there are four factors explaining 91% of the variance (Ferrans & Powers, 1992).

Brief Symptom Inventory (BSI; Derogatis and Melisaratos, 1983). The BSI is a 53-item self-report measure assessing symptoms of psychopathology. Participants rate, on a 5-point scale, the extent to which each item distresses them, from *not at all* to *extremely*. The BSI is comprised of nine subscales (Cronbach's alpha ranges from 0.75 to 0.89) and has high convergent validity with the Minnesota Multiphasic Personality Inventory (MMPI) subscales (Boulet & Boss, 1991).

Feeling thermometer. Participants completed a feeling thermometer (adapted from Greenwald & Farnham, 2000, and described previously) for each of the six words (*stupid*, *smart*, *ugly*, *attractive*, *shy*, *friendly*) used on the IRAP.

Procedure

Participants presented to the research laboratory individually. Each participant first completed the self-report measures (presented in the standardized order listed above) and then the IRAP. To be consistent with prior studies, the trials were counterbalanced, such that some participants completed the consistent trials first, and others completed the inconsistent trials first.

Results

IRAP analyses. Seven participants were excluded from the IRAP analyses. One was excluded due to computer malfunction during the IRAP (the computer experienced an unexpected reboot and the participant's data was incomplete); the other six were excluded because they failed to pass the IRAP practice trial blocks with 65% accuracy (final $n = 93$).

Feeling thermometer ratings. To ascertain whether participants viewed the six IRAP target words as having the intended valence (positive or negative), the feeling thermometer ratings were examined. There was a significant main effect for target word, $F(5, 485) = 256.24$, $p < .001$, $\eta_p^2 = 0.73$. Post-hoc LSD pairwise comparisons revealed that *friendly* was rated significantly higher than both *smart* and *attractive*, which did not differ significantly. These terms were followed by *shy*, *stupid*, and *ugly*. The latter two were rated equally unpleasant. In sum, the target words, ranked in order from highest to lowest mean thermometer rating, were as follows: *friendly*, *smart/attractive*, *shy*, *stupid/ugly*. Thus, participants interpreted the positive target words as positive and the negative words as negative.

Order effects and partial D_{IRAP} analyses. During the IRAP task, participants were exposed to six successive test blocks, alternating between consistent and inconsistent blocks. For each of the three consistent-inconsistent pairs of test blocks, the IRAP program calculates a partial D_{IRAP} score. The presentation order of test blocks was counterbalanced. A 2 (consistent first versus inconsistent first) \times 3 (test block pair) mixed repeated-measures ANOVA was conducted to assess whether either presentation order or the partial D_{IRAP} scores interacted with the overall IRAP effect. There were no significant main effects of either presentation order, $F(1, 91) = 1.57$, $p = .21$, $\eta_p^2 = 0.02$, or partial D_{IRAP} score, $F(2, 182) = 1.36$, $p = .26$, $\eta_p^2 = 0.02$. The interaction was also not significant, $F(2, 182) = 0.07$, $p = .94$, $\eta_p^2 = 0.001$. Therefore, presentation order and partial D_{IRAP} scores were excluded from further analyses.

IRAP trial-type analyses. As in Study 1, the IRAP data were next examined at the level of individual trial types. Thus, four $D_{\text{IRAP-TRIAL TYPE}}$ scores were calculated for each participant as described above. One-sample t tests indicated that the $D_{\text{IRAP-TRIAL TYPE}}$ scores for each of the four trial types differed significantly ($p < .001$) from zero (see Figure 4).

Study 2

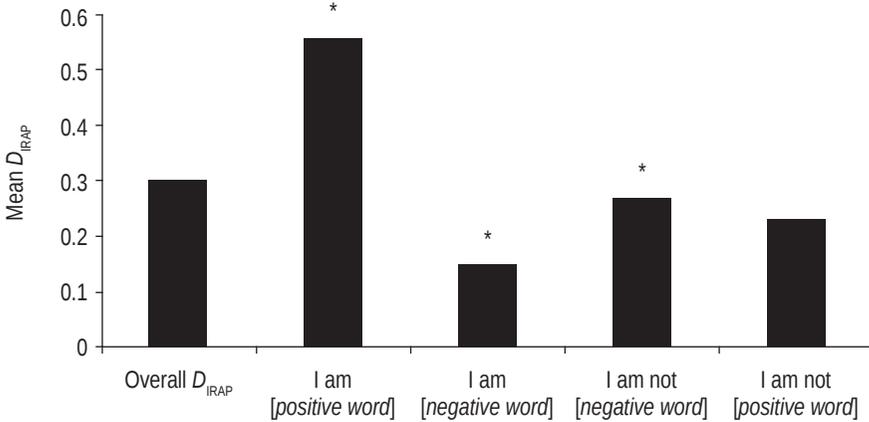


Figure 4. Overall D_{IRAP} and trial-type D_{IRAP} scores for Study 2. *Significantly different from zero ($p < .001$).

A repeated-measures ANOVA comparing the four trial types was significant, $F(3, 276) = 22.82, p < .001$. LSD post-hoc pairwise comparisons indicated several significant differences between trial types (see Figure 3 for a pictorial representation of the results). The $D_{IRAP-TRIAL\ TYPE}$ score for “I am [positive word]” was significantly higher than those for “I am [negative word]” and “I am not [positive word].” The score for “I am not [negative word]” was significantly lower than that for “I am [positive word]” and significantly higher than that for “I am [negative word].” The score for “I am not [positive word]” did not differ significantly from those for “I am not [negative word]” and “I am [negative word].”

Correlations between overall D_{IRAP} scores and self-report measures.

For the purpose of elucidating the relationships between the IRAP and the various explicit measures, the overall D_{IRAP} scores and the $D_{IRAP-TRIAL\ TYPE}$ scores were entered into a correlation matrix with the self-report measures. Correlations were small in magnitude, and only the “I am [positive word]” $D_{IRAP-TRIAL\ TYPE}$ score was significantly correlated with explicit measures. Specifically, the trial-type score was positively and weakly correlated with quality of life and weakly and negatively correlated with measures of pathology (see Table 2).

Table 2
Correlations Between Overall D_{IRAP} and $D_{IRAP-TRIAL\ TYPE}$ Scores and Self-Report Measures: Study 2

Self-report measure	Overall D_{IRAP} score	I am [positive word]	I am [negative word]	I am not [positive word]	I am not [negative word]
QLI total score	.13	.26*	-.01	.06	-.01
QLI-Health and Functioning	.14	.28**	.03	.05	-.01
QLI-Social and Economic	.03	.12	-.06	.05	-.04
QLI-Psychological/Spiritual	.13	.29**	.01	.04	-.02
QLI-Family	.07	.12	-.09	.07	.07
BSI-Anxiety	-.06	-.12	.04	.02	-.08
BSI-Depression	-.03	-.22*	-.08	.14	.08
BSI-Interpersonal Sensitivity	-.14	-.25*	-.17	.09	-.04
BSI-Global Symptom Severity	-.07	-.21*	-.06	.04	.05

Note. QLI = Quality of Life Index: Generic Version-III (Ferrans & Powers, 1985); BSI = Brief Symptom Inventory (Derogatis and Melisaratos, 1983). * $p < .05$. ** $p < .01$.

Discussion

Study 2 attempted to assess the validity of the IRAP as an implicit measure of general self-esteem, using the sample relations “I am” and “I am not.” As in Study 1, participants were able to affirm positive self-statements and deny negative self-statements more quickly than they could affirm negative self-statements and deny positive self-statements. Overall, participants responded in a way consistent with positive self-esteem or self-concept as measured implicitly via the IRAP.

Though there were no significant correlations between explicit measures and the overall D_{IRAP} score, examination of the $D_{IRAP-TRIAL\ TYPE}$ correlations with explicit measures revealed some significant relationships. Specifically, a greater IRAP effect for the trial type “I am [positive word]” was associated with increased overall quality of life, greater health and psychological/spiritual functioning, and decreased depression, interpersonal sensitivity, and psychopathology. Overall, this provides further evidence that the IRAP can implicitly assess self-esteem.

General Discussion

Implicit measures have become popular due to their apparent ability to assess attitudes or beliefs about concepts while eliminating variables that could influence responding on explicit measures. The purpose of the present studies was to evaluate the usefulness of the IRAP as an implicit measure of self-esteem. In Study 1, self-esteem was assessed by investigating overall body satisfaction in Caucasian women; in Study 2, self-esteem was broadly assessed (friendliness, intelligence, and appearance) in both

men and women. Participants in both studies generated positive D_{IRAP} scores, indicating overall positive self-image. Therefore, the results of these studies are in line with both previous research using the IAT to assess self-esteem and research using the IRAP to measure self-referential relational responding. Given that both studies used different target words to tap into self-esteem, the IRAP appears to be a fairly robust and reliable measure of this construct.

Implicit positive self-attitudes as measured by the IRAP in these studies were correlated with explicit measures. In Study 1, overall implicitly measured body satisfaction was correlated with greater body satisfaction and less pressure to conform to the thin ideal. These correlations support the contention that the IRAP, as used here, more appropriately taps into body satisfaction than studies that implicitly assess an anti-fat bias. Furthermore, implicitly measured body satisfaction was associated with greater body image acceptance and less depression, both of which may be indicators of greater self-esteem.

An advantage of the IRAP, namely, the ability to analyze the data via specific trial-type scores, became apparent in Study 2. The overall D_{IRAP} score was not correlated with any explicit measures. This lack of association could be interpreted as a lack of relationship between implicit and explicit measures or as evidence that the IRAP was not eliciting immediate self-referential relational responding. As the IRAP is more flexible than traditional implicit measures, the pattern of relational responses and how they correlated with explicit measures can be examined in greater detail. In Study 2, the pattern of correlations between the $D_{\text{IRAP-TRIAL TYPE}}$ for the "I am [positive word]" trials was in the expected direction, with greater implicitly measured self-esteem associated with greater quality of life and less overall pathology. This pattern of correlations is notably different from Study 1, wherein a number of significant correlations were found across trial types. In Study 1, the explicit measures used were directly related to the body, as were the stimuli used in the IRAP. Thus, it is unsurprising that correlations were present across trial types. In Study 2, the explicit measures assessed much broader constructs; consequently, the stimuli used in the IRAP were less domain specific. Thus, the relationship between the explicit and implicit measure in Study 2 was only apparent in one of the four trial types, a relationship that would not have been observed if only the D_{IRAP} score was used.

Both studies employed the terms "I am" and "I am not" for the samples. This, in effect, created a measure akin to a single-concept IAT. Participants were asked to affirm or deny statements about what they believed themselves to be or not to be. Though the samples "I am" and "I am not" were chosen in order to recreate natural speech, it may be that referring to oneself in terms of what one *is* versus what one *is not* in fact better reflects natural language. This is perhaps due to humans being taught to refer to themselves in terms of what they are versus what they are not, as well as the finite number of possibilities of who or what one *is* versus the infinite number of possibilities of who or what one *is not*.

It has been argued that in order for an implicit measure to demonstrate construct validity it must be able to differentiate between known groups (De Houwer, 2002; De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). The IRAP employed by Vahey and colleagues (2009) was able to do this. The current Study 1 broadly replicated this finding by demonstrating

differences between the self-identified groups of nondieters, those dieting to lose weight, and those dieting to maintain their weight. Women who self-identified as dieting to lose weight had lower overall D_{IRAP} scores than those not dieting to lose weight, indicating not only that the IRAP was able to measure implicit attitudes about one's own body but also that it was also sensitive to differences between groups.

The IRAP was introduced less than 5 years ago as a more flexible and theoretically grounded implicit measure. As such, it may be better suited to address some of the criticisms of the IAT. To date, over a dozen studies have been conducted using the IRAP. The method has been used to assess a variety of attitudes, and data increasingly support its validity and robustness (e.g., in terms of fakeability and malleability). The current studies add to this literature by providing further evidence that the IRAP can serve as a measure of self esteem and that known groups do differ in terms of their IRAP scores. Research is needed to establish further the validity of the IRAP, particularly according to the criteria described by De Houwer and colleagues (2009). Thus far, IRAP studies (including the ones presented herein) have primarily focused on establishing validity via correlations and quasi-experimental research. The experimental validation of the IRAP is an exciting area of research yet to be fully explored. In terms of self-esteem, this would include determining whether or not the implicit measurement of self-esteem predicts behavior differentially across contexts. Future research should also investigate whether or not the relations being measured with the IRAP can be experimentally manipulated or changed. For example, if body image is targeted for improvement, will relevant IRAP effects change following intervention? Given that the IRAP is theoretically grounded in RFT, specific predictions could be made regarding how and if this immediate relational responding will change with intervention.

References

- 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*, 176-181. doi:10.1037/1040-3590.10.2.176
- BARNES-HOLMES, D., BARNES-HOLMES, Y., POWER, P., HAYDEN, E., MILNE, R., & STEWART, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*, 169-177. Retrieved from http://www.psihq.ie/comms_index.asp
- BARNES-HOLMES, D., MURPHY, A., BARNES-HOLMES, Y., & STEWART, I. (2010). The Implicit Relational Assessment Procedure (IRAP): Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record, 60*, 57-80.
- BARNES-HOLMES, D., MURTAGH, L., BARNES-HOLMES, Y., & STEWART, I. (2010). Using the Implicit Association Test and the Implicit Relational Assessment Procedure to measure attitudes towards meat and vegetables in vegetarians and meat-eaters. *The Psychological Record, 60*, 287-306.

- BARNES-HOLMES, D., WALDRON, D., BARNES-HOLMES, Y., & STEWART, I. (2009). Testing the validity of the Implicit Relational Assessment Procedure and the Implicit Association Test: Measuring attitudes towards Dublin and country life in Ireland. *The Psychological Record*, *59*, 389–406.
- BOULET, J., & BOSS, M. W. (1991). Reliability and validity of the Brief Symptom Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 433–437. doi:10.1037/1040-3590.3.3.433
- COOPER, P. J., TAYLOR, M. J., COOPER, Z., & FAIRBURN, C. G. (1986). The development and validation of the Body Shape Questionnaire. *International Journal of Eating Disorders*, *6*, 485–494. doi:10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O
- CULLEN, C., & BARNES-HOLMES, D. (2008). Implicit pride and prejudice: A heterosexual phenomenon? In T. G. Morrison & M. A. Morrison (Eds.), *Modern prejudice* (pp. 195–223). New York: Nova Science.
- CULLEN, C., BARNES-HOLMES, D., BARNES-HOLMES, Y., & STEWART, I. (2009). The Implicit Relational Assessment Procedure (IRAP) and the malleability of ageist attitudes. *The Psychological Record*, *59*, 591–620.
- 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. doi:10.1016/S0005-7916(02)00024-1
- DE HOUWER, J., TEIGE-MOCIGEMBA, S., SPRUYT, A., & MOORS, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi:10.1037/a0014211
- DEROGATIS, L. R., & MELISARATOS, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, *13*, 595–605. doi:10.1017/S0033291700048017
- DRAKE, C. E. (2009). Evaluating a behavioral measure of psychological flexibility. *Dissertation Abstracts International: Section B. Sciences and Engineering*, *69*(9-B), 5764.
- DRAKE, C. E., KELLUM, K. K., WILSON, K. G., LUOMA, J. B., WEINSTEIN, J. H., & ADAMS, C. H. (2010). Examining the Implicit Relational Assessment Procedure: Four pilot studies. *The Psychological Record*, *60*, 81–100.
- EVANS, C., & DOLAN, B. (1993). Body Shape Questionnaire: Derivation of shortened “alternate forms.” *International Journal of Eating Disorders*, *13*(3), 315–321. doi:10.1002/1098-108X(199304)13:3<315::AID-EAT2260130310>3.0.CO;2-3
- FAHERTY, A., TIMKO, C. A., & KALODNER, C. (2009). *Body image acceptance moderates the relationship between dissatisfaction and eating disorder symptomatology*. Unpublished manuscript.
- FAZIO, R. H., & OLSON, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327. doi:10.1146/annurev.psych.54.101601.145225
- FERRANS, C. E., & POWERS, M. J. (n.d.). *Description of the Quality of Life Index (QLI)*. Retrieved from <http://www.uic.edu/orgs/qli/reliability/reliabilityhome.htm>
- FERRANS, C. E., & POWERS, M. J. (1985). Quality of Life Index: Development and psychometric properties. *Advances in Nursing Science*, *8*, 15–24. Retrieved from <http://journals.lww.com/advancesinnursingscience/pages/default.aspx>
- FERRANS, C. E., & POWERS, M. J. (1992). Psychometric assessment of the Quality of Life Index. *Research in Nursing & Health*, *15*, 29–38. doi:10.1002/nur.4770150106

- GREENWALD, A. G., BANAJI, M. R., RUDMAN, L. A., FARNHAM, S. D., NOSEK, B. A., & MELLOTT, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25. doi:10.1037/0033-295X.109.1.3
- GREENWALD, A. G., & FARNHAM, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022–1038. doi:10.1037/0022-3514.79.6.1022
- GREENWALD, A. G., MCGHEE, D. E., & SCHWARZ, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi:10.1037/0022-3514.74.6.1464
- HAYES, S. C., BARNES-HOLMES, D., & ROCHE, B. (Eds.). (2001). *Relational frame theory: A Post-Skinnerian account of human language and cognition*. New York: Plenum Press.
- 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. *Personality and Social Psychology Bulletin*, *31*, 1369–1385. doi:10.1177/0146167205275613
- IP, K., & JARRY, J. L. (2008). Investment in body image for self-definition results in greater vulnerability to the thin media than does investment in appearance management. *Body Image*, *5*, 59–69. doi:10.1016/j.bodyim.2007.08.002
- KARPINSKI, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin*, *30*, 22–34. doi:10.1177/0146167203258835
- KIM, D.-Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly*, *66*, 83–96. doi:10.2307/3090143
- LOVIBOND, S. H., & LOVIBOND, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. (2nd ed.). Sydney: Psychology Foundation.
- MASUDA, A., HAYES, S. C., SACKETT, C. F., & TWOHIG, M. P. (2004). Cognitive defusion and self-relevant negative thoughts: Examining the impact of a ninety-year-old technique. *Behaviour Research and Therapy*, *42*, 477–485. doi:10.1016/j.brat.2003.10.008
- MCKENNA, I. M., BARNES-HOLMES, D., BARNES-HOLMES, Y., & STEWART, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology and Psychological Therapy*, *7*, 253–268. Retrieved from <http://www.ijpsy.com>
- MERCURIO, A. E., & LANDRY, L. J. (2008). Self-objectification and well-being: The impact of self-objectification on women's overall sense of self-worth and life satisfaction. *Sex Roles*, *58*, 458–466. doi:10.1007/s11199-007-9357-3
- OLSON, M. A., FAZIO, R. H., & HERMANN, A. D. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological Science*, *18*, 287–291. doi:10.1111/j.1467-9280.2007.01890.x
- O'TOOLE, C., & BARNES-HOLMES, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record*, *59*, 119–132.
- RASHID, N., HAAS, M., & TIMKO, C. A. (2007, November). *The Implicit Relational Assessment Procedure as a measure of implicit modern racism*. Poster presented at the 41st annual meeting of the Association for Behavioral and Cognitive Therapy, Philadelphia, PA.

- ROBERTS, A., CASH, T. F., FEINGOLD, A., & JOHNSON, B. T. (2006). Are black-white differences in females' body dissatisfaction decreasing? A meta-analytic review. *Journal of Consulting and Clinical Psychology, 74*(6), 1121-1131. doi:10.1037/0022-006X.74.6.1121
- ROSEN, J., JONES, A., RAMIREZ, E., & WAXMAN, S. (1996). Body shape questionnaire: Studies of validity and reliability. *International Journal of Eating Disorders, 20*, 315-319. doi:10.1002/(SICI)1098-108X(199611)20:3<315::AID-EAT11>3.0.CO;2-Z
- SANDOZ, E. K., WILSON, K. G., & MERWIN, R. M. (2008). *Assessment of body image acceptance: The Body Image-Acceptance and Action Questionnaire*. Manuscript submitted for publication.
- SCHWARTZ, M., VARTANIAN, L., NOSEK, B., & BROWNELL, K. (2006). The influence of one's own body weight on implicit and explicit anti-fat bias. *Obesity, 14*, 440-447. doi:10.1038/oby.2006.58
- TAFARODI, R. W., & HO, C. (2006). Implicit and explicit self-esteem: What are we measuring? *Canadian Psychology, 47*, 195-202. doi:10.1037/cp2006009
- THOMPSON, J. K., VAN DEN BERG, P., ROEHRIG, M., GUARDA, A. S., & HEINBERG, L. J. (2004). The Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3): Development and validation. *International Journal of Eating Disorders, 35*, 293-304. doi:10.1002/eat.10257
- TIMKO, C. A., MARTIN, L. M., & DARCY, C. (2010). *The impact of the media on body satisfaction: The role of experiential avoidance and dieting*. Manuscript submitted for publication.
- TIMKO, C. A., & PERONE, J. (2006). Rigid and flexible control of eating behavior and their relationship to dieting status. *Eating and Weight Disorders/Studies on Anorexia, Bulimia, and Obesity, 11*, e90-e95. Retrieved from <http://www.kurtis.it/ewd/en/>
- TIMKO, C. A., PERONE, J., & CROSSFIELD, A. (2006). Are you currently on a diet? What respondents mean when they say "yes." *Eating Disorders: The Journal of Treatment and Prevention, 14*, 157-166. doi:10.1080/10640260500536300
- VAHEY, N. A., BARNES-HOLMES, D., BARNES-HOLMES, Y., & STEWART, I. (2009). A first test of the Implicit Relational Assessment Procedure (IRAP) as a measure of self-esteem: Irish prisoner groups and university students. *The Psychological Record, 59*, 371-388.
- VAUGHAN, C. A., SACCO, W. P., & BECKSTEAD, J. W. (2008). Racial/ethnic differences in body mass index: The roles of beliefs about thinness and dietary restriction. *Body Image, 5*(3), 291-298. doi:10.1016/j.bodyim.2008.02.004
- VERPLANKEN, B., & VELSVIK, R. (2008). Habitual negative body image thinking as psychological risk factor in adolescents. *Body Image, 5*(2), 133-140. doi:10.1016/j.bodyim.2007.11.001
- YAMAGUCHI, S., GREENWALD, A., BANAJI, M., MURAKAMI, F., CHEN, D., SHIOMURA, K., Kobayashi, C., Cai, H., & KRENDL, A. (2007). Apparent universality of positive implicit self-esteem. *Psychological Science, 18*, 498-500. doi:10.1111/j.1467-9280.2007.01928.x