RUNNING HEAD: PREDICTIVE VALIDITY OF THE IAT

Citation:
Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (in press). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*.

Understanding and Using the Implicit Association Test: III.

Meta-analysis of Predictive Validity

Anthony G. Greenwald
University of Washington

T. Andrew Poehlman
Southern Methodist University

Eric Luis Uhlmann
Northwestern University

Mahzarin R. Banaji
Harvard University

CONTACT:
Anthony G. Greenwald
Department of Psychology
University of Washington
Box 351525
Seattle, WA 98195-1525
Tel: (206)-543-7227
E-mail: agg@u.washington.edu

Draft of 30 December 2008

KEYWORDS: Implicit Association Test, implicit measures, validity, implicit attitudes, attitude–behavior relations

Abstract (131 words)

This review of 122 research reports (184 independent samples, 14,900 subjects), found average *r*=.274 for prediction of behavioral, judgment, and physiological measures by Implicit Association Test (IAT) measures. Parallel explicit (i.e., self-report) measures, available in 156 of these samples (13,068 subjects), also predicted effectively (average *r*=.361), but with much greater variability of effect size. Predictive validity of self-report was impaired for socially sensitive topics, for which impression management may distort self-report responses. For 32 samples with criterion measures involving Black–White interracial behavior, predictive validity of IAT measures significantly exceeded that of self-report measures. Both IAT and self-report measures displayed incremental validity, with each measure predicting criterion variance beyond that predicted by the other. The more highly IAT and self-report measures were intercorrelated, the greater was the predictive validity of each.

Understanding and Using the Implicit Association Test: III.

Meta-analysis of Predictive Validity

In the first ever handbook-chapter review of a social psychological construct, Gordon Allport (1935) characterized *attitude* as social psychology's "most distinctive and indispensable concept." That characterization has been accepted by scholars ever since, even during a period in which the attitude construct was enmeshed in a crisis of predictive validity. That crisis was triggered by Wicker (1969), who found very little evidence to support the conclusion that attitudes predicted behavior toward the attitudes' objects (cf. Festinger, 1964). As a result of Wicker's review, during the 1970s social psychologists were obliged to consider that their esteemed attitude construct might not deserve the lofty position that Allport had proposed.

In fairness to the attitude construct, there had been relatively few empirical investigations of the predictive validity of attitude measures prior to Wicker's (1969) review. When social psychologists began to address this empirical lack, they initially found it difficult to obtain the desired evidence for predictive validity of attitudes. However, by the early 1980s several researchers, especially Ajzen and Fishbein (e.g., 1977) and Fazio and Zanna (e.g., 1981), had successfully established the predictive validity of attitude measures, thus restoring the attitude construct to its prior status (see also Kelman, 1974). In 1995, sixty years after Allport had hailed attitude as social psychology's premier construct, Kraus's (1995) meta-analysis of results from 88 attitude–behavior relationship studies yielded an average predictive validity effect size estimate of $r = .38$.

Research on attitude–behavior relations in the 1970s and 1980s established two methods that reliably produced at least moderate effect sizes for attitude–behavior correlations. The first was a refinement of self-report methods for measuring attitudes, to ensure that attitude measures were phrased to correspond closely to the measures of behavior with which their correlations were being examined (Ajzen & Fishbein, 1977). The second was to identify and capitalize on moderator variables that influenced the strength of attitude–behavior correlations, such as the personal importance of the attitude and its stability across time (e.g., Krosnick, 1988).

The attitude construct has developed further since Kraus's (1995) review. Recent findings have revealed attitudinal processes for which their possessors may have limited awareness and which, therefore, may not be well captured by self-report measures (e.g., Bargh, Chaiken, Govender & Pratto, 1992; Bargh & Chartrand, 1999; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald & Banaji, 1995; Hetts, Sakuma, & Pelham, 1999; Jones, Pelham, Mirenberg, & Hetts, 2002; Nisbett & Wilson, 1977; von Hippel, Sekaquaptewa, & Vargas, 1997; Wittenbrink, Judd & Park, 1997). The task of determining whether measures of this *implicit* aspect of attitudes effectively predict behavior has been pursued most extensively with one particular method, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998; see recent overview by Nosek, Greenwald, & Banaji, 2007). This article summarizes research that has been conducted to evaluate the predictive validity of IAT measures. Although the present review is not limited to IAT measures of attitudes, nevertheless attitudes have been the dominant focus in predictive validity research on the IAT. Sixty-nine percent of the presently reviewed IAT studies focused on attitude measures.

The Implicit Association Test (IAT)

IAT measures assess strengths of associations between concepts by observing response latencies in computer-administered categorization tasks. In an initial block of trials, exemplars of two contrasted concepts (e.g., face images for racial *Black* and *White*) appear on a screen and subjects rapidly classify them by pressing one of two keys (for example, an "e" key for Black and "i" for White). Next, exemplars of another pair of contrasted concepts (for example, words representing *positive* and *negative* valence) are also classified, using the same two keys. In a *first combined task*, exemplars of all four categories are classified, with each assigned to the same key as in the initial two blocks (e.g., "e" for *Black* or *positive* and "i" for *White* or *negative*). In a *second combined task*, a complementary pairing is used (i.e., "e" for *White* or *positive* and "i" for *Black* or *negative*).[1] In most implementations, respondents are obliged to correct errors before proceeding, and latencies are measured to the occurrence of the correct response. The difference in average latency between the two combined tasks provides the basis for the IAT measure. For example, faster responses for the {Black+positive|White+negative} task than for {White+positive|Black+negative} indicate a stronger association of *Black* than of *White* with *positive* valence.

Research conducted since the initial 1998 publication of the IAT has provided substantial evidence concerning psychometric properties of IAT measures (Egloff & Schmukle, 2002; Greenwald & Nosek, 2001; Greenwald & Farnham, 2000; Lane, Banaji, Nosek, & Greenwald, 2007; Nosek et al., 2007; Rudman, Greenwald, Mellott, & Schwartz, 1999). IAT measures have typically displayed good internal consistency (Bosson, Swann, & Pennebaker, 2000; Dasgupta & Greenwald, 2001; Greenwald & Nosek, 2001; Greenwald & Farnham, 2001); IAT measures are not influenced by wide variations in subjects' familiarity with IAT stimuli (Dasgupta, McGhee, Greenwald, & Banaji, 2000; Ottaway, Hayden, & Oakes, 2001; Rudman et al., 1999); and IAT measures are relatively insensitive to procedural variations such as the number of trials, the number of exemplars per concept, and the time interval between trials (Nosek, Greenwald, &

Banaji, 2005; Greenwald et al., 1998). Test–retest reliability of IAT measures was recently reported to have a median value of $r = .56$, across nine available reports (Nosek et al., 2007).

A useful property of IAT measures is their presumed reliance on associative processes that can operate automatically (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Greenwald et al., 2002; see Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005, for an investigation aimed at distinguishing the contributions of automatic and controlled processes to IAT measures). The sensitivity of IAT measures to automatically activated associations is sometimes credited with making IAT scores resistant (even if not immune) to faking. For example, subjects instructed to fake positive attitudes towards gay men were able to do so on a self-report questionnaire but not on a homosexual–heterosexual attitude IAT (Banse, Seise, & Zerbes, 2001). Asendorpf, Banse, and Mücke (2002) obtained similar findings with a shyness self-concept IAT, as did Kim (2003) with a race attitude IAT measure. Similarly, subjects instructed to make a good impression in a job application scenario easily altered their self-report responses to appear low in anxiety, but their scores on an anxiety self-concept IAT were relatively unaffected (Egloff & Schmukle, 2002). Subjects who are explicitly instructed to slow their responding in one of the IAT's two combined tasks can use that instruction to produce faked scores. At the same time, most naïve subjects do not spontaneously discover this strategy (Cvencek, Greenwald, Brown, Gray, & Snowden, 2008; Kim, 2003; Steffens, 2004; but cf. Fiedler & Bluemke, 2005).

Widespread use of the IAT to investigate attitudes has produced a situation like that which existed for self-report measures of attitudes at the time of Wicker's (1969) review. It is time to evaluate the IAT's ability to predict relevant social behavior (cf. Banaji, 2001; Fazio & Olson, 2003; Karpinski & Hilton, 2001; Olson & Fazio, 2004). The need for this evaluation of predictive validity is heightened by expressions of interest in using IAT measures for applications in law, policy, and business (e.g., Ayres, 2001; Banaji & Bhaskar, 2000; Banaji & Dasgupta, 1998; Chugh, 2004). Evaluating the predictive validity of IAT measures can also help achieve a goal that several commentators on IAT measures have urged — appraising the construct validity of IAT measures (Arkes & Tetlock, 2004; Blanton & Jaccard, 2006; Fiedler, De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Messner, & Bluemke, 2006; Karpinski & Hilton, 2001; Rothermund & Wentura, 2004).

In recent investigations, IAT measures have been found to correlate with many measures of interest, such as anxious behaviors (Asendorpf et al., 2002), preference for a partner to perform an intellectual task (Ashburn-Nardo, Knowles, & Monteith, 2003), math SAT scores (Nosek, Banaji, & Greenwald, 2002a), and alcohol consumption over the course of a month (Wiers, Woerden, Smulders, & de Jong, 2002). In other studies, IAT measures did not predict measures with which a relation was expected (e.g., food choice in Karpinski & Hilton, 2001). The present research assessed the predictive validity of IAT measures quantitatively, while also comparing the predictive validity of IAT measures with that of parallel explicit (self-report) measures, which were available for almost 90% of the studies included in this review.

Method

Criteria for Study Inclusion

The authors sought to include all studies that reported predictive validity correlations involving four types of IAT measures of association strengths — attitudes (concept–valence associations), stereotypes (group–trait associations), self-concepts or identities (self–trait or self–group associations), and self-esteem (self–valence associations). A requirement for inclusion was that the predicted (i.e., criterion) measure was itself neither an implicit measure nor an alternative-format measure of the same construct being measured by the IAT predictor. Excluded, therefore, were studies focusing on correlations among IAT measures of different constructs (e.g., Greenwald et al., 2002) or studies in which use of the IAT was limited to investigating correlations between IAT and parallel self-report measures. Numerous studies of the latter type were recently reviewed meta-analytically by Hofmann, Gawronski, Gschwendner, Le, and Schmitt (2005) and this was also the subject of a 57-topic study by Nosek (2005). Also excluded were studies in which an IAT measure of self–group association (implicit identity) or group–valence association (implicit attitude) was correlated with membership in that group. An additional category of exclusions consisted of studies in which an IAT measure was used as a moderator variable, because these studies had no expectation of observing a direct correlation between the IAT measure and a criterion measure of behavior. The criterion measures that remained available for meta-analysis included a wide variety of measures of physical actions, judgments, preferences expressed as choices, and physiological reactions.

To illustrate the exclusions and inclusions: A study of correlations between an IAT measure of attitude toward mathematics and self-report measures of math attitudes (e.g., Nosek, Banaji, & Greenwald, 2002b) was excluded because the observed relationship was between IAT and self-report measures of the same construct (i.e., attitude towards mathematics). In contrast, studies reporting correlations between IAT race attitude measures and nonverbal actions toward persons of that race (e.g., McConnell & Leibold, 2001) were included. Known-groups studies that compared (for example) whether Japanese Americans and Korean Americans differed in an IAT measure of associations of positive or negative valence with the concepts Japanese and Korean (Greenwald et al., 1998, Experiment 2) were excluded because the self-identification (e.g., as Japanese American) was regarded as being too similar to a self-report of attitude.

However, a study examining correlations between an IAT measure of attitude toward smoking and self-reported smoking status (Swanson, Rudman, & Greenwald, 2001) was included because the self-identification (as smoker or non-smoker) could be understood as the measure of a relevant behavior. An experiment by Greenwald & Farnham (2000, Experiment 3) was excluded by the IAT-as-moderator exclusion because their hypothesis was that IAT-measured self-esteem might moderate attributions in response to success versus failure, rather than predicting a direct relation between the self-esteem and attribution measures.

Search Method

Studies were initially sought using three methods: (a) PsycINFO search (using the keywords "IAT", "implicit association test", "implicit measure", "implicit attitudes", "automatic attitudes" or "implicit social cognition"), (b) Internet search (using google.com, keywords: "IAT" or "implicit association test"); and (c) email to the Society of Personality and Social Psychology's mailing list, requesting any in press or unpublished research using IAT measures. The reference sections of the articles thus obtained were further searched for relevant studies. When this article was accepted for publication in December 2007, the database for its meta-analysis included 103 reports. At that point, a search to determine the availability of more recent versions of included reports led to dropping four reports that were superseded by more recent versions that reported more data, and one other report for which insufficient documentation was available. The search for more recent versions produced an additional 20 reports for which manuscripts had not previously come to the authors' attention but were determined to have existed in some usable preliminary form prior to the February 1, 2007 cutoff date. Reports that could not be established as having been distributed in some form prior to the cutoff date were not included.

Many reports did not contain effect sizes either in the desired form of zero-order correlations (*r*s) or as other statistics that could be converted to zero-order correlations. Additionally, many desired effect sizes — especially ones involving self-report measures — were not included in the available reports. The first author corresponded with authors in search of these potentially useful additional effect sizes. These were obtained in the great majority of cases. Of the 1,461 effect sizes that comprise the database for this report, 426 (29.2%) were obtained as a result of such further correspondence with authors.

Calculation of Effect Sizes

Each of the 122 published or unpublished reports that met criteria for inclusion was separated into statistically independent samples (Lipsey & Wilson, 2001, p. 112). For each of these samples, a mean IAT–criterion measure correlation (ICC) was computed. Whenever possible, mean explicit (i.e., self-report) measure–criterion measure correlations (ECCs) and mean IAT–explicit correlations (IECs) were also computed. All of these mean effect sizes were computed using Fisher's *r*-to-*Z* transformation to average all correlations of the same type that were available in each independent sample. Each mean *Z* was associated with an inverse variance weight, which was computed as (n – 3) where n is the number of subjects in the independent sample (Hedges & Olkin, 1985, p. 333). The 122 reports thus provided 526 ICCs from 184 independent samples, based on 14,900 subjects.[2] There were 557 ECCs available from 156 of these independent samples (based on 13,068 subjects), and 378 IECs available from 155 of the samples (based on 13,120 subjects).

Description and Coding of Moderators

Variables identified as moderators that might explain across-sample variance in effect sizes fell into three categories: *conceptual*, *methodological*, and *publication*. Conceptual moderators were variables suggested either by previous reviews of attitude–behavior relations (e.g., Kraus, 1995) or by findings of the developing literature with IAT measures. Methodological moderators included procedural variations that occur frequently in laboratory studies as well as other routine procedural variations of IAT studies. Two publication characteristics were used as moderators: (a) publication year and (b) status of report as published or unpublished.

Coding of several of the moderators required judgments based on reading of reports' *Methods* sections. For the studies in the original (103-report) data analysis, three raters judged each study independently. One of those three raters was blind to results of all studies. The other two were aware of the results of different portions of the studies. For all study characteristics that required such judgments, satisfactory inter-rater reliability was observed (Cronbach's $\alpha \geq .70$) and the three raters' judgments were averaged for use in analyses. Such reliable ratings of study characteristics have been used successfully in previous meta-analyses (e.g., Eagly, Johannesen-Schmidt, & van Engen, 2003). For methodological and other predictors, the few disagreements among the three judges were resolved by discussion.

While the meta-analysis was under review for publication, additional studies that qualified for inclusion were identified. For these studies, moderators were coded by one of the raters who had judged all of the previous studies (the other two were unavailable for this purpose). At the same time, that rater reviewed all previous ratings to ensure that the full set of studies was coded in consistent fashion.[3]

Conceptual Moderators

Descriptive statistics for the study characteristics coded as conceptual moderators are summarized in Table 1.

Social sensitivity. Subjects' desire to be perceived positively is widely assumed to be a potential source of distortion of self-report measures (e.g., Crosby, Bromley, & Saxe, 1980; Crowne & Marlowe, 1960; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Nosek & Banaji, 2002). Consequently, self-report measures in socially sensitive domains — such as self-reported attitudes and beliefs about racial or ethnic groups — might suffer impression-management distortions that could reduce their predictive validity. If, as is also widely assumed, IAT measures are relatively resistant to impression management, their predictive validity may show relatively little influence by social sensitivity of the study topic (cf. Asendorpf et al., 2002; Banse et al., 2001; Egloff & Schmukle, 2002; Kim, 2003).

Raters were instructed to make separate judgments for each self-report and IAT measure in a report, judging the extent to which self-reporting the construct assessed by the measure might activate concerns about the impression that their response would make on others. For example, self-reporting attitudes towards Black Americans is something that raters might judge to be considerably more socially sensitive than is self-reporting attitudes toward brands of yogurt. Judgments were made on a scale of 1–7 (1 = not at all likely to be affected by social desirability concerns; 7 = extremely likely to be affected by social desirability concerns). To repeat for clarity, the social sensitivity measure for IAT measures was judged to be the sensitivity associated with *self-reporting* the same attitude, belief, self-concept, or self-esteem measure. Inter-rater reliability for social sensitivity was acceptable ($\alpha = .74$). Because self-report scales often assessed constructs very similar to those assessed by IAT measures in the same study, the social sensitivity ratings for IAT and explicit measures that predicted the same criterion were very highly correlated, $r(462) = .99$. They lack of perfect correlation occurred because the IAT and self-report measure in a study did not always measuring the same construct.

Controllability of responses to the criterion measure. Dual-process models of social cognition suppose that introspectively accessible attitudes and beliefs effectively guide deliberate actions, but play weaker roles in determining spontaneous actions. Therefore, implicit measures of attitudes and beliefs may predict spontaneous actions more effectively than do explicit measures (Asendorpf et al., 2002; Egloff & Schmukle, 2002; Devine, 1989; Dovidio et al., 1997; Fazio, 1990; Wilson, Lindsey, & Schooler, 2000). Some research with implicit measures other than the IAT has supported this supposition (e.g., Dovidio et al., 1997; Fazio et al., 1995).

However, not all automaticity theorists suppose that automatic attitudes relate more to spontaneous than to deliberately controlled responses. For example, Rudman (2004) pointed out that implicit measures sometimes correlate substantially with the (highly controllable) responses to parallel explicit measures of attitude, such as political candidates (e.g., Greenwald et al., 2003; Nosek, 2005; Nosek & Banaji, 2002), suggesting that they may also effectively predict other controllable behaviors (see also Bargh & Chartrand, 1999; Haidt, 2001; Wegner & Bargh, 1998).

Each criterion measure was rated for the extent to which the responses that it required were judged easy to consciously control. For example, choice of vote for a presidential candidate might be easy to control, whereas nonverbal behaviors such as eye blinks, speech hesitations, or body orientation might be difficult to control. Judgments were made on a scale of 0–10 (0 = no component of the response is consciously controllable, 10 = all components of the response are consciously controllable). Inter-rater reliability for controllability was satisfactory ($\alpha = .80$).

Complementarity. For some preferences, liking one alternative implies disliking a complementary alternative. For example, having a positive attitude toward a candidate of one political party might imply having a negative attitude towards a political competitor from another party, but it might not imply having a negative attitude toward another candidate from the same party. In contrast, having a positive attitude toward one brand of yogurt might not imply having a negative attitude toward other brands of yogurt.

To rate complementarity, judges estimated the extent to which liking one of the two IAT target categories in a measure implied disliking the other. Judgments of complementarity used a 9-point scale (1 = extremely non-complementary, 9 = extremely complementary). Inter-rater reliability was satisfactory (α = .84). Complementarity was not coded for explicit measures because contrast categories were much less frequently included in the construction of explicit measures.

Correspondence. Ajzen and Fishbein (1977) identified a moderating role of similarity between verbal descriptions of attitude and behavior measures (i.e., correspondence between the measures) on magnitude of attitude–behavior correlations. They found greater attitude–behavior correlations the more the attitude measure shared features with the behavior measure. For example, church attendance was predicted more strongly by measures of an attitude toward the church being attended than by measures of an attitude toward religion in general. Kraus's (1995) meta-analysis confirmed this hypothesized moderating role of correspondence.

Correspondence was judged on a 7-point scale (1 = extremely low correspondence, 7 = extremely high correspondence). Inter-rater reliability for correspondence was acceptable (α = .76). Mean correspondence ratings between IAT and self-report measures that predicted the same criterion were highly correlated, $r(464) = .80$.

The highest levels of correspondence observed in the data set for both IAT and self-report measures (rated for both at 5 on the 7-point scale) occurred with criterion measures involving political or consumer preferences. For example, in Karpinski, Steinman, and Hilton's (2005) study of intention to use Coke or Pepsi products, their IAT measure used these two brands as the contrasted categories, while their self-report measures included feeling thermometer, semantic differential, and 6-point Likert ratings of the two brands. An example of very low correspondence (rated 1 on the 7-point scale) was use of a race attitude IAT measure and self-reported racial attitudes in a study in which the criterion measures consisted of subtle nonverbal indicators of discomfort in interaction, such as speech dysfluency or bodily position (e.g., McConnell & Leibold, 2001).

Type of predictor construct: attitude vs. belief, self-concept, or self-esteem. Predictor IAT and explicit measures were easily categorizable as corresponding to constructs of attitude, belief (most often a stereotypic group–trait association), self-concept (including group identity), or self-esteem. Partly because the majority of studies used attitude measures and also because attitude has been such an important focus of previous predictive validity research, the measure of type of predictor was reduced to a dichotomy for moderator analyses, separating attitude measures (coded 1) from the other three types. Use of this moderator allowed determination of whether predictive validity for attitude measures was possibly greater than that for the other three types of measures. This binary moderator was coded separately for IAT and self-report measures. Seventeen percent of independent samples had mixtures of types of predictors, leading to independent samples having values of this predictor between 0 and 1. The correlation of values of this moderator between IAT and self-report measures in independent samples that had both types of measure was $r(157) = .67$.

IAT–explicit correlation (IEC). Research investigations have found that correlations between implicit and explicit measures vary widely (Hofmann, et al., 2005; Nosek, 2005). Several theorists have proposed that weak relationships between implicit and self-report attitude responses may indicate intrapsychic conflict (Epstein, 1994; Fazio, 1990; Gaertner & Dovidio, 1986; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; McGregor & Marigold, 2003; Nosek, 2005; Petty, Tormala, Briñol, & Jarvis, 2006; Wilson et al., 2000). Empirical research has demonstrated discrepancies between implicit (or automatic) and explicit (or deliberative) measures in the domains of problem solving (Epstein, 1994), race prejudice (Gaertner & Dovidio, 1986; Fazio & Olson, 2003), and attitude change (Wilson et al., 2000). The frequent observation of weak correlations between implicit and explicit measures suggests that inconsistency between them is relatively common (cf. Hofmann et al., 2005; Nosek, 2005). If high IECs indicate that automatic and controlled influences on behavior support one another, then high IECs may be associated with high predictive validity for both IAT and self-report measures. As already described, IECs were available for 155 of the 184 independent samples.

Methodological Moderators

Procedural and method variations that were coded for use as potential moderators are summarized in Table 2.

Numbers of effect sizes and numbers of IAT measures. It was possible that, when studies included multiple effect sizes, these might include measures expected to show weak effects along with ones expected to show strong effects. Consequently, average effect sizes might be weaker in studies that had larger numbers of effect sizes. Number of IAT measures in the study was a related predictor that was used as a potential moderator of ICCs.

Numbers of subjects. Sample sizes averaged n = 81.0, but there was wide variation (SD = 141.5). There are two diverging expectations for a moderating role of sample size. If large sample sizes are used to provide added power when expected effect sizes are small, large sample sizes should be associated with relatively small ICCs or ECCs. However, sample size variations may also result from variations in cost or convenience of obtaining subjects, in which case there is little basis for expecting a relationship between sample size and predictive validity effect size.

Subject response vs. experimenter-observed criterion measure. Each criterion measure was coded dichotomously as to whether it was observed (i.e., unobtrusively recorded by the experimenter, which was coded 1) or, alternately, based on subjects providing information via either paper–pencil responses or computer entry (coded 0). For IAT and self-report predictors, respectively, 33% and 35% of independent samples had unobtrusively observed criterion measures. There was no advance expectation about how this might relate to observed effect sizes.

IAT scoring method. Each study was coded as to whether its IAT measure was computed using averaged combined-task latencies in millisecond units, averaged combined-task latencies in log-transformed latencies (both coded 0) or the *D* scoring algorithm (Greenwald, Nosek, & Banaji, 2003; coded 1). There was a weak expectation of effects being stronger for the *D* algorithm, because of its somewhat superior psychometric properties.

Order and proximity of measures. Variation of timing of IAT or self-report predictors in relation to the criterion measure was a potentially interesting moderator. When an IAT measure precedes the criterion measure, accessibility of the associations measured by the IAT may be enhanced or primed, thereby possibly inflating predictive validity correlations. In support of this possibility, Monteith, Voils, and Ashburn-Nardo (2001) reported that IAT effects can be "palpable" to subjects, who may be able to discern their possession of the associations measured by the IAT. On the other hand, and as suggested by Bem's (1972) self-perception theory, completing a criterion measure may temporarily modify the associations measured by the IAT, which might increase correlation between IAT and criterion measure. Each study was coded to indicate whether IAT or self-report predictors preceded or followed their associated criterion measures. Studies that counterbalanced predictor–criterion order (only 5% or the total for IAT measures) received an intermediate code.

Administering criterion measures in sessions separate from assessment of IAT or self-report predictors might minimize mutual influences between predictor and criterion measures (Fazio & Olson, 2003; Kraus, 1995). For both ICCs and ECCs, 18% of independent samples had criterion measures in a separate session from predictors.[4]

Publication Moderators

Standard practice for reporting meta-analyses includes coding year of publication, type of research participant (student or non-student), and site of study (field or laboratory). Most of the studies included in this meta-analysis were laboratory studies with undergraduate students as subjects. Only 14 reports (11%) used non-student samples. Ten of these used clinical populations, for some of which data collection was in a laboratory setting. Because of the small numbers of non-student and non-laboratory samples, neither the type of subject nor the site of study was used as a moderator. However, studies were coded for year of publication and for publication status (unpublished = 0 versus published or in press = 1). For both ICCs and ECCs, only 17% of independent samples were from unpublished reports. This small proportion of unpublished studies may in part be a consequence of the approximate 18-month interval between the cutoff date for inclusion in the meta-analysis (early 2007) and completion of this report. In the interim, several reports that had entered the meta-analysis in unpublished form transitioned to published or in press status.

Criterion Measure Domain

Effect sizes were sorted into nine domains based on similarities among criterion measures. These nine criterion categories, which are listed in Table 3 and will be considered in more detail later, served primarily to distinguish well-recognized topical groupings of effect sizes, primarily for use in presenting descriptive summaries (see Figures 1 and 2).

Results

Analysis Overview

This article's first goal was to estimate an average effect size for IAT–criterion correlations (ICCs). Shortly after the authors started to analyze the ICC data it became obvious that, because parallel self-report measures were used in many of the studies, it would be possible to provide comparative predictive validity estimates for explicit–criterion correlations (ECCs). A further consequence of the frequent use of self-report measures in the meta-analyzed studies was that IAT–explicit correlations (IECs) were available for 155 (84%) of the 184 independent samples. This made it possible to report analyses that partly overlapped with Hofmann et al.'s (2005) recent meta-analysis of IECs and Nosek's (2005) extensive study of IAT–self-report correlations. Availability of a complete trio of effect sizes — ICC, ECC, and IEC — in 152 samples permitted estimation of partial correlations of IAT and self-report measures with criterion measures. With the other type of predictor partialed, it was possible to assess incremental validity of each type of predictor. Although these estimated partial correlations provided less-than-perfect estimates of incremental validity (for reasons to be explained when presenting them), they are nevertheless informative.

Mean sample-size-weighted effect sizes for ICCs, ECCs, and IECs were examined both as aggregates across all independent samples and as aggregates within each of the nine criterion category domains. Potential moderators of magnitude for each type of effect size (ICC, ECC, and IEC) were also examined in two series of sample-size-weighted regression analyses — the first for conceptual moderators and the second for methodological and publication moderators. Except where noted otherwise, these analyses used mixed statistical models in which a random component of between-study variance was fit by maximum likelihood estimation.[5]

Aggregate Effect Sizes and Homogeneity Tests

Table 3's top row of data reports weighted average effect sizes for ICCs, ECCs, and IECs, aggregated across all available independent samples (k = 184 for ICCs; k = 156 for ECCs; k = 155 for IECs). The aggregate weighted average effect sizes were $\bar{r}_{ICC}$ = .274, $\bar{r}_{ECC}$ = .361, and $\bar{r}_{IEC}$ = .214. All three types of effect sizes were significantly heterogeneous when tested with fixed effects models. The Q statistics (with their associated *df*) were Q_{ICC} (183) = 576.7, Q_{ECC} (155) = 1914.5, and Q_{IEC} (154) = 731.2. Substantially greater heterogeneity in ECCs than ICCs was revealed not only by the very different values of these Q statistics, but also by standard deviations reported in Table 3. The weighted standard deviation for all ECCs (SD = .391) was almost double that for ICCs (SD = .215), indicating considerably greater variability of effect sizes for ECCs than ICCs.

Table 3 also reveals variations in effect sizes across the nine criterion domains. For ICCs, mean effect sizes ranged from .171 to .483 for the nine domains. For ECCs, the range was almost double that for ICCs, from .118 to .709. IECs were overall slightly lower than ICCs and more substantially lower than ECCs, with a range from .091 to .537. All three types of aggregate effect size were largest for political preferences ($\bar{r}_{ICC}$ = .48; $\bar{r}_{ECC}$ = .71; $\bar{r}_{IEC}$ = .54). ICCs were smallest for close relationships ($\bar{r}_{ICC}$ = .17) and for gender/sexual orientation ($\bar{r}_{ICC}$ = .18), while ECCs were smallest for race ($\bar{r}_{ECC}$ = .12) and other intergroup behavior ($\bar{r}_{ECC}$ = .12). IECs were smallest for close relationships ($\bar{r}_{IEC}$ = .09) and race ($\bar{r}_{IEC}$ = .12). Except for the aggregate IEC for the close relationship category (bottom right of Table 3), all reported aggregate effect sizes for ICCs, ECCs, and IECs differed significantly from zero in the positive direction by random-effects test, with 2-tailed α = .05.

The criterion domain of White–Black interracial behavior (*k* = 32) and the "other intergroup" category (*k* = 15), which included behavior toward groups defined by ethnicity, age, or weight, were the only two domains in which average magnitudes of ICCs significantly exceeded those of ECCs. (These two domains were not grouped into a single category mainly because of a priori separate interest in the category of White–Black interracial behavior.) For interracial behavior, aggregate ICC ($\bar{r}_{ICC}$ = .24) was significantly greater than aggregate ECC ($\bar{r}_{ECC}$ = .12), $z = 4.27$, $p = 10^{-5}$. The domain of interracial behavior was the only domain within which there was statistical homogeneity for all three types of effect sizes (i.e., all *p*s > .05 for fixed-effect homogeneity tests).

Regression Analyses with Conceptual Moderators

In the attempt to identify sources of variation in magnitudes of all three types of effect sizes, weighted regression analyses were conducted (Hedges & Olkin, 1985; cf. Lipsey & Wilson, 2001, p. 122f.), using the previously described conceptual, method, and publication moderators. Criterion domain was not used as a predictor in any analyses of conceptual moderators because (as will be described more fully later) criterion domain variations were extensively confounded with the several conceptual moderators.

ICCs. Table 4 summarizes the weighted regression analyses involving conceptual moderators of ICCs. Moderator effects are shown both for analyses using each moderator as the sole regression predictor (univariate analysis, left side of Table 4), and for a multiple weighted regression format that entered all moderators simultaneously (right side of Table 4).

When used as a univariate predictor in a mixed model (fixed slopes, random intercepts), magnitude of IECs explained 29.9% of ICC variance, $p = 10^{-15}$. Predictive validity of IAT measures was greater when self-report and IAT measures were more strongly correlated. This finding is consistent with the reasoning that both ICCs and ECCs should be relatively strong when there is little conflict or dissociation between these two measures. A large value of IEC (correlation between self-report and IAT) indicates the lack of dissociation. Three other conceptual moderators were also significant in univariate analyses (see left side of Table 4). Predictive validity of IAT measures was greater with (a) greater complementarity of the two categories contrasted in IAT measures (explaining 9.8% of variance, $p = 10^{-5}$), (b) greater correspondence between the IAT and the criterion measure (6.5% of variance, $p = .001$), and (c) lower social sensitivity of the implicit construct being measured (3.4% of variance, $p = .02$). Complementarity and social sensitivity were also significant predictors in the simultaneous analysis (right side of Table 4), but correspondence was not. This difference between univariate and simultaneous regression results is expected when there is collinearity (correlation) among predictors — this will be considered more fully in the Discussion.

ECCs. Table 5 presents the analysis of conceptual moderators of predictive validity for ECCs. Note that even though complementarity was a property of each study's IAT measures it was used as a predictor in the analysis of ECCs. This was because of the suspicion that complementarity might capture properties of study independent of the structure of the IAT measure.

As was true for ICCs, IEC magnitude was the strongest individual predictor of ECCs in univariate analyses (left side of Table 5), where it accounted for 34.3% of ECC variance in a univariate analysis, $p = 10^{-20}$. All five other moderators also had significant univariate effects in the analysis of ECCs. Predictive validity of self-report measures was (a) strongly reduced by social sensitivity of the construct being measured (24.4% of variance, $p = 10^{-10}$), (b) increased by correspondence between predictor and criterion measures (35.3% of variance, $p = 10^{-17}$), (c) increased by complementarity (9.5% of variance, $p = .0001$), (d) increased by controllability of the criterion measure (8.8% of variance, $p = .0009$), and (e) was also greater for attitude predictors than other types (9.8% of variance, $p = .0002$). The univariate effect of social sensitivity on predictive validity of ECCs was an order of magnitude greater (24.4% of variance in the univariate analysis) than was its effect on predictive validity of ICCs (3.4% of variance). This large difference was consistent with the expectation that predictive validity of self-report (but not of IAT) measures might be impaired in socially sensitive domains.

Social sensitivity, correspondence, and complementarity remained significant as predictors in the simultaneous regression analysis, but controllability and predictor type did not (right side of Table 5). All regression coefficients were substantially reduced from the univariate analyses, a consequence of correlations among the predictors. These correlations have implications for theoretical interpretation, a topic to be treated in Discussion.

IECs. Table 6 presents the analysis of IECs that is parallel to those just described for ICCs and ECCs. Two type-of-predictor (attitude vs. other types) moderators were used, one each for IAT and self-report measures. The social sensitivity and correspondence moderators were very highly correlated for IAT and self-report measures and were therefore averaged for use as predictors in this analysis. Although all six of the conceptual moderators were significant in the univariate regression analyses of IECs, only two remained significant in the simultaneous regression analysis: IECs were greater with (a) greater complementary (p = .0001) and (b) lower social sensitivity (p = .02).

Correlations among conceptual predictors. Table 7 presents unweighted correlations among variables in the regression analyses of Tables 4–6. The lower left of Table 7 describes correlations involving the conceptual moderators of ICCs and the upper right describes those for ECCs. Most notable in Table 7 are (a) the large magnitudes of many of these correlations — more than half of them corresponded to moderate or large effect sizes, and (b) the considerably larger correlations of moderators with ECCs (top row of Table 7) than with ICCs (first column of Table 7). This contrast of correlation magnitudes fits with the observations of stronger moderation effects for ECCs (Table 5) than for ICCs (Table 4). The intercorrelations among the moderators will play a role in the Discussion section's analysis of differences between results of univariate and simultaneous regressions for conceptual moderators.

Weighted Regressions with Methodological and Publication Moderators

Tables 8 and 9 summarize regression analyses involving method and publication moderators. (Table 2 provides summary descriptions of these moderator variables.) The simultaneous multiple regression analyses in the right sides of both tables showed very weak overall results (for ICCs, R^2 = .104, p = .12; for ECCs, R^2 = .110, p = .03). Only the effect of method of data collection for the criterion measure on ECCs (Table 9) was statistically significant in both univariate and simultaneous multiple regression tests. Effect sizes were smaller for criterion measures that required an observer's coding of behavior than for criterion measures that resulted directly from subjects' responses. A similar, but weaker, effect was observed in the multiple regression analysis for ICCs (Table 8).

Another effect in the analysis of method moderators that was consistent in direction for ICCs and ECCs, although not in statistical significance, was for number of effect sizes. For both ICCs and ECCs, effect sizes were weaker as more effect sizes were averaged together for a sample. This suggests some support for the speculation that studies with more predictive validity effect sizes were more likely to include effect sizes that were expected to show little or no predictive validity.

Tables 8 and 9 are most interesting for what they did not reveal. There were no effects of order-of-measurement moderators for either ICCs or ECCs. This included no significant effects due either to (a) timing of administration of criterion measures in relation to IAT or self-report predictors or (b) use of criterion and predictor measures in the same versus separate sessions. A second interesting non-significant result was that effect sizes were unrelated to publication status. Although effect sizes are often expected to be smaller for unpublished than published studies, the present findings showed no support for that expectation. Lastly, the effect of IAT scoring method on ICC magnitudes (Table 8) was weakly in the expected direction of showing stronger effect sizes for use of the *D* scoring algorithm (Greenwald et al., 2003) than for other scoring methods. However, this effect was not statistically significant.

Differences among Criterion Measure Domains

Figure 1 summarizes aggregate effect sizes of ICCs and ECCs for the nine categories of criterion measures. Three findings are visible in the figure. First, effect sizes for both ICCs and ECCs varied widely across domains. Second, this across-domain variation in effect sizes was much greater for ECCs than for ICCs (i.e., lengths of the black bars in Figure 1 vary much more than do those of the gray bars); this greater heterogeneity of ECCs than ICCs can also be seen in the wider 95% confidence intervals for black than gray bars in Figure 1. Third — and possibly most important — although average ECCs were significantly greater than ICCs in six criterion domains, the reverse was true for the two domains that involved intergroup behavior (the top two pairs of bars in Figure 1).

Incremental Validity — Partial Correlation Method

The aim of determining whether IAT and self-report measures independently explained variance in criterion measures was at first frustrated because so few of the reports examined for this meta-analysis included regression analyses in which IAT and self-report measures were used as simultaneous predictors. An alternative approach was available for the 152 samples that permitted estimates of all three types of effect size — ICC, ECC, and IEC. Availability of these three effect sizes permitted estimate of two partial correlations: (a) correlation of IAT with criterion, partialing self-report ($r_{IC.E}$) and (b) correlation of self-report with criterion, partialing IAT ($r_{EC.I}$).

Three cautions must be considered in using the trios of effect sizes to estimate partial correlations that might be interpreted as indicators of incremental validity. First, the *r* values that were used to compute the partial *r*s were often averages over available *r*s of the same type within each independent sample. Second, the three *r* values for each sample were not always based on data provided by exactly the same subjects. Third, although a significant partial correlation indicates that the effect of one predictor is *statistically* independent of the partialed predictor, it does not necessarily indicate that the two predictors are *conceptually* independent.

The possibility of significant partial correlations for conceptually similar predictors arises when the two predictors contain measurement error — as was certainly the case for all IAT and self-report measures in the reports gathered for the present meta-analysis. Consider the hypothetical case of a criterion measure (Y) being predicted by two predictors (X_1 and X_2), both of which are assumed to measure exactly the same construct (X). If X_1 and X_2 have uncorrelated measurement errors, each can have some incremental validity in predicting Y — that is, each will have a positive partial correlation with Y, controlling for the other. Even though these three considerations complicate interpretation of partial correlations as indicators of incremental validity, some patterns of results can permit unequivocal conclusions.

The desired partial correlations, $r_{IC.E}$ and $r_{EC.I}$, were computable for 152 of the 184 independent samples in the meta-analysis. For both $r_{IC.E}$ and $r_{EC.I}$, overall weighted mean values were significantly greater than zero. For the partial correlations of IAT measures with criterion measures, $\bar{r}_{IC.E} = .179$, $z = 14.18$, $p = 10^{-45}$ (random effects model). These $\bar{r}_{IC.E}$ values were significantly heterogeneous by fixed-effects test, $Q = 238.8$, $df = 151$, $p = 10^{-5}$. For correlations of self-report measures with criterion measures, $\bar{r}_{EC.I} = .321$, $z = 11.73$, $p = 10^{-31}$. These $\bar{r}_{EC.I}$ values were also heterogeneous, $Q = 1356.8$, $df = 151$, $p = 10^{-192}$.

Figure 2 shows the weighted average effect sizes for both $r_{IC.E}$ and $r_{EC.I}$, separately for the nine criterion-measure domains. For two domains (White vs. Black race and other intergroup), $\bar{r}_{IC.E}$ significantly exceeded $\bar{r}_{EC.I}$. In all seven other domains, $\bar{r}_{EC.I}$ significantly exceeded $\bar{r}_{IC.E}$. The White vs. Black race category was the only criterion domain for which the difference between partial correlations was statistically homogeneous, Q (26 df) = 32.70, $p = .17$, fixed-effects test.

Discussion

The first goal of this meta-analysis was to estimate the average predictive validity effect size (r) of IAT measures. The weighted average of these IAT–criterion correlations (ICCs), based on 122 reports that contained 184 independent samples, was $\bar{r}_{ICC}$ = .274, a level conventionally characterized as "moderate" (Cohen, 1977, p. 80). On average, correlations of self-report measures with criterion measures (ECCs) were larger: $\bar{r}_{ECC}$ = .361. Other important findings were that (a) predictive validity of self-report measures (but not of IAT measures) was sharply reduced when research topics were socially sensitive, (b) IAT measures had greater predictive validity than did self-report measures for criterion measures involving interracial behavior and other intergroup behavior, and (c) both IAT and self-report measures showed incremental predictive validity with respect to each other.

The average ICCs and ECCs observed in this research were inevitably attenuated in magnitude due to unreliability of both predictor and criterion measures. Some methodologists (e.g., Schmidt, Pearlman, Hunter, & Hirsch, 1985) advocate conducting meta-analyses on effect sizes that have had preliminary corrections for unreliability. Such disattenuated effect sizes are necessarily larger than published effect sizes, because they are computed by dividing the published effect size by the product of square roots of reliabilities of the two component measures — a quantity that is necessarily less than 1.0. This strategy was not used because the authors had, at best, imprecise knowledge of reliabilities for most of the measures used in this meta-analysis.

Disattenuated correlations can nevertheless be crudely approximately estimated by making assumptions about reliabilities of the measures composing the correlations. Using assumed reliabilities of r = .56 for IAT predictors (based on Nosek et al., 2007) and r = .80 for criterion measures (an estimate for which there can be no strong basis), the estimated average predictive validity of IAT measures in the present research would increase from observed $\bar{r}_{ICC}$ = .274 to disatennuated $\hat{\bar{r}}_{ICC}$ = .409. A similar computation for ECCs, using reliability estimates of

r = .85 for self-report predictors and r = .80 for criterion measures, yields a disattenuated estimate of $\hat{\bar{r}}_{ECC}$ = .438, compared to the observed $\bar{r}_{ECC}$ = .361.

Predictive Validities Vary Across Domains of Criterion Behavior

Average ECCs were greater than average ICCs for seven of the nine criterion domains (see Table 3 and Figure 1). Both ICCs and ECCs were greatest in magnitude for political preferences ($\bar{r}_{ICC}$ = .483; $\bar{r}_{ECC}$ = .709). The relatively high ICC effect sizes in the political domain and in the consumer preferences domain may indicate why these two, in combination, accounted for 28% of the meta-analyzed independent samples. Figures 1 and 2 showed that, in the domains of Black–White interracial behavior and other intergroup behavior (and only in these two domains) IAT measures had greater predictive validity than did self-report measures. The relative success of IAT measures for these two topics may explain why these two have been so prominent in research using IAT measures — together they comprise 26% of the meta-analysis.

Social Sensitivity of Topic Impairs Predictive Validity of Self-Report Measures

As a single predictor, social sensitivity of topic explained 24.4% of variance in ECC effect sizes (see Table 5). Social sensitivity was a much weaker moderator of ICC effect sizes (3.4% of variance, see Table 4). To interpret this contrast, mean levels of social sensitivity were examined for the nine criterion domains. Rated social sensitivity ranged from 1 (not at all likely to be affected by social desirability concerns) to 7 (extremely likely to be affected by social desirability concerns). For the two domains with highest average ECCs (political preferences and consumer preferences), 100% of samples had social sensitivity ratings of 3 or below. In contrast, for the two domains with lowest average ECCs (White–Black race and other intergroup behavior) 100% of the samples had social sensitivity scores above 3. There was thus no overlap in rated social sensitivity of the study topics in these two sets of domains. While indicating that social sensitivity plausibly played a causal role in moderating predictive validity of self-report measures, it is also clear that its effect was confounded with differences in topic domains.

Comparison of social sensitivity's large effect on ECCs, compared with its much weaker effect on ICCs, fits with previous conclusions that impression management can undermine validity of self-report measures in socially sensitive domains (e.g., Greenwald et al., 2002; Nosek et al., 2007). An estimate of the magnitude of this interfering effect can be obtained by applying the unstandardized regression parameter estimates from the univariate regression of ECCs on social sensitivity.[6] Using those estimates, the expected predictive validity of self-report measures for a topic rated 1 (lowest) in social sensitivity is $\hat{r}_{ECC}$ = .60, while that for a topic rated 7 (highest) in social sensitivity is $\hat{r}_{ECC}$ = .10.

Mutual Incremental Validity of IAT and Self-Report Measures

For 152 samples, availability of a trio of effect sizes — ICC, ECC, and IEC — permitted estimation of partial correlations. These partial correlations indicated that IAT and self-report measures had mutual incremental validity in predicting criterion measures. As was emphasized in presenting results, these partial-correlation analyses have potential problems associated with averaging correlations within independent samples as well as from unreliability of measures. Those limitations notwithstanding, the partial correlation findings indicated clearly that IAT and self-report measures each predicted criterion variance that was not predicted by the other. For IAT measures, this was clearest for the White–Black race and other intergroup behavior topics. In these topic domains, evidence for incremental validity of IAT measures was accompanied by evidence for very low predictive validity of self-report measures (see Figures 1 and 2). In several other topic domains — especially consumer preferences, political preferences, and clinical phenomena — it was strongly evident that self-report measures predicted criterion variance not predicted by IAT measures.

Understanding the Strong Moderating Role of IAT–Explicit Correlation (IEC) Magnitude

The finding that IEC magnitude was positively associated with predictive validity for both ICCs and ECCs was expected from reasoning that, when IAT and self-report measures agree, the constructs that they measure will likely reinforce each other in determining behavior. This, in turn, should produce relatively large predictive validity correlations for both types of measure. Confirming this expectation, IEC magnitude positively predicted 29.9% of the variability of ICC effect sizes and 34.3% of the variability of ECCs (see Tables 4 and 5).

It has been theorized that *response factors* (including demand characteristics, evaluation-apprehension, and subject role-playing) and *introspective limits* will cause self-report measures to diverge from IAT measures, resulting in relatively low IECs (Greenwald et al., 2002, p. 17). In support of findings by Nosek (2005), the present research found strong evidence for effects of a response factor that plausibly reduced correlations between self-report and IAT measures — social sensitivity of the research topic. The present research found IECs to be markedly lower for highly sensitive topics than for topics rated low in social sensitivity. Supporting that observation, the unweighted correlation of IEC magnitude with rated social sensitivity of topic was $r = -.35$ for the samples used to analyze conceptual moderators of both ICCs and ECCs (see Table 7).[7]

Previous evidence for the role of introspective limits in affecting IEC magnitudes appeared in Hofmann et al.'s (2005) finding that IECs were larger when self-report measures were judged to be high in spontaneity (i.e., to be based on little introspection). Reinforcing that observation, Ranganath, Smith, and Nosek (2008) reported that self-report measures had larger correlations with IAT measures when self-report procedures were modified to invite greater spontaneity. They achieved this either by asking subjects to describe "gut reactions" in their self-reports or by obliging subjects to give self-report responses under time pressure, thereby reducing opportunity to think about how to respond.

Other Conceptual Moderators

Correspondence. Correspondence between criterion measures and attitude predictors was first identified as a moderator of attitude–behavior relations by Ajzen and Fishbein (1977) and later confirmed as a moderator in Kraus's (1995) meta-analytic review. Correspondence was likewise found to be a significant moderator of ECCs in the present meta-analysis, both in univariate and simultaneous regression analyses (see Table 5). However, correspondence was a significant moderator of ICCs only in the univariate regression analysis (see Table 4).

Complementarity. As a characteristic of IAT measures, complementarity is high when liking one target category implies disliking its contrasted category. For example, in the United States liking the Republican Party implies disliking the Democratic Party. As previously described, complementarity of IAT measures was suspected to be as much a characteristic linked to the topic of a research study as it was a characteristic of the study's IAT measures. This suspicion was confirmed by observing that, on its 1–9 scale, complementarity was much higher for political preference topics (mean = 6.2) than for all other topics (mean = 2.0). The moderating effects of complementarity on ICC and ECC effect sizes might therefore be in part a consequence of the relatively large number of subjects in studies of political preferences (see Table 3), giving that category relatively large weight in the meta-analysis.

Controllability. The introductory discussion of controllability as a potential moderator described dual-process theories that credit implicit measures with a stronger role in predicting spontaneous than controlled behavior. According to these dual-process views (e.g., Strack & Deutsch, 2004; Wilson et al., 2000), the more controlled the behavior, the less well it should be predicted by implicit measures. The introductory description also described the opposed view that implicit measures should be capable of predicting both controlled actions and spontaneous ones (e.g., Nosek & Banaji, 2002; Rudman, 2004). The present findings strongly supported the latter position. Controllability showed no significant moderating effects in the regression analyses for conceptual moderators of ICCs (Tables 4). Perhaps this should not be surprising, considering that some of the largest ICC effect sizes occurred in the political and consumer preferences domains, in which criterion measures were often coded as being highly controllable.

Intercorrelations Among Conceptual Moderators of ECCs

The analysis of conceptual moderators of ECCs (see Table 5) showed six unequivocally significant predictors in univariate regressions. For these, absolute beta values ranged from .296 to .594. In the corresponding simultaneous regression, two of these predictors were no longer significant and the set of absolute beta values was noticeably lower, ranging from .015 to .292. These reduced regression coefficients can be understood by considering the high correlations among the conceptual moderators (see Table 7).

Correlations among moderators notwithstanding, the effects of three moderators on ECCs seem well established. First, the effect of correspondence in increasing ECCs, which was initially proposed and confirmed about three decades ago and again in Kraus's (1995) meta-analysis (see Introduction), was effectively confirmed in this meta-analysis. Second, the effect of social sensitivity, which was expected on the basis of widespread understanding of impression management as an interfering factor in self-reports, was evident in both the univariate and simultaneous regression analyses of conceptual moderators of ECCs. Third, the effect of IEC magnitude on increasing predictive validity was very strongly evident in the present analyses of both ECCs and ICCs. The commonsense interpretation of this effect is that, when IAT and self-report measures are highly correlated, their respective bases for predicting behavior should be mutually reinforcing, which should result in relatively high predictive validity correlations for both types of measure.

More difficult to interpret are the significant effects of predictor type (attitude vs. other types of measure), controllability of the criterion response, and complementarity of the contrasting IAT categories. Of these three, only complementarity had a significant effect in the simultaneous regression analysis. However, this property of IAT measures was not expected to have any role in predictive validity of self-report measures. Further, this property was not coded for self-report measures because it could be a characteristic only of measures that contrasted two categories. Although a substantial fraction of self-report measures did make use of contrasted categories, there were too few of these for that coding to be useful in a regression analysis — a large fraction of the meta-analyzed samples would have been dropped for lack of coding.

To test whether the effect of complementarity was a consequence of the high value of this moderator in studies of political preferences, the simultaneous regression of Table 5 was rerun omitting the nine political preference samples for which both ECCs and IECs were available. The effect of complementarity in moderating predictive validity of self-report measures was not at all diminished in the reduced-sample analysis. This observation, along with the relatively low correlations of complementarity with other moderators (see Table 7), indicates that, at least in the analysis of predictive validity of self-report, the effect of complementarity was not reflecting effects that might reasonably be credited to other moderators. At the same time, a similar reduced-sample rerun of the simultaneous regression for ICCs showed that the significant effect of complementarity (Table 4: $\beta = .259$, $p = .001$) became non-significant ($\beta = .148$, $p = .07$). The only justifiable present conclusion is that the moderating effect of complementarity on predictive validity of self-report measures remains a puzzle yet to be solved.

The lack of significant effect of controllability in the simultaneous regression analysis of ECCs (Table 5) may be a consequence of its substantial correlations with social sensitivity and correspondence. In practice it may be difficult to separate controllability of self-report measures from their social sensitivity and correspondence properties. Correspondence was highest in the same topic groupings in which controllability was high, and these were also topics for which social sensitivity was low.

Predictive validity effect sizes of self-report measures were larger in studies involving attitude measures than in studies using the three other types of self-report measures. The conversion of this effect to non-significance in the simultaneous regression (right side of Table 5) was likely a consequence of the positive correlations of predictor type with two other significant moderators of ECCs' predictive validity — IEC magnitude and correspondence (see Table 7). Self-report attitude measures had higher correlations with IAT measures and higher levels of rated correspondence to criterion measures than did (collectively) self-report measures of stereotypes, self-concepts, and self-esteem. As in the case of controllability, this may be a situation in which these dimensions are sufficiently entwined in nature to make it impractical to examine their effects separately.

Methodological and Publication Moderators

The analyses of methodological and publication moderators were remarkable for the near-absence of effects (see Tables 8 and 9). The only noteworthy effect was that involving the contrast between measures produced by subject behavior and those coded as a result of experimenters' observations. Predictive validity effect sizes were larger for the approximately two-thirds of studies in which criterion measures were produced by subject behavior. Perhaps subjects were finding ways to inflate consistency with predictors when they responded to criterion measures, or perhaps experimenters' coding introduced greater error into recorded criterion responses. More important than this isolated finding was the absence of effects of procedural variations involving the relative temporal position of predictor and criterion measures. Although there is no reason to oppose standard practices of counterbalancing orders of these measures, it also appears that there is little harm — at least for the purpose of estimating predictive validity — in using fixed orders of measurement. Additionally, there was no indication that having self-report and IAT predictors in the same vs. separate sessions affected predictive validity effect sizes. ICC effect sizes were slightly higher with separate sessions and ECC effect sizes were slightly lower with separate sessions, but neither of these differences was statistically significant (see Tables 8 and 9).

Dual-Representation vs. Single-Representation vs. Dual-Construct Theoretical Interpretations

The hypothesis that IAT measures and self-report measures capture distinct phenomena is supported by two observations: (a) mutual incremental validity for the two types of measures — which indicates that they predicted different aspects of criterion behavior, and (b) the finding that social sensitivity of topic affected predictive validity of self-report measures much more strongly than it affected predictive validity of IAT measures. Some theorists will interpret these findings to indicate that "implicit attitudes" and "explicit attitudes" are distinct entities, as suggested in dual-representation theories such as those of Wilson et al. (2000) and Strack and Deutsch (2004). However, other theorists (e.g., Fazio & Olson, 2003; Kruglanski & Thompson, 1999) point out that these apparent empirical implicit–explicit dissociations can be accounted for by a single-representation form of theory. In the single-representation approach, implicit and explicit attitudes (for example) are conceived not as distinct mental entities, but rather as distinct types of measures that can derive from a single form of underlying representation. Greenwald and Nosek (2008) have advocated a middle course by treating implicit and explicit measures as empirically distinct *constructs*, noting that, at present, the question of single versus dual representations appears empirically unresolvable.

Continuity with Important Previous Reviews

This review continues themes developed in Crosby, Bromley, and Saxe's (1980) review of unobtrusive-measure research on race discrimination and in Kraus's (1995) meta-analysis of attitude–behavior relations. Crosby et al. drew attention to the substantial divergence of results between survey studies of racially discriminatory attitudes and results of the unobtrusive-measure experiments that they reviewed. They found that "discriminatory behavior is more prevalent in the body of unobtrusive studies than we might expect on the basis of survey data" (p. 557). After they noted that self-report measures were often found to be poor predictors of racial discrimination in studies that used unobtrusive measures, Crosby et al. "inferred from [this] literature that whites today [i.e., in 1980] are, in fact, more prejudiced than they are wont to admit."

Crosby et al. (1980, p. 557) identified five studies as showing poor predictive validity of self-report racial attitude measures. These studies did not appear in Kraus's (1995) meta-analysis, either because they did not meet Kraus's inclusion criterion of having attitude measures in a separate session preceding criterion measurement (see Kraus, p. 62)[8] or because some of them, instead of reporting numerical effect sizes, described attitude–behavior correlations only as "non-significant". Kraus may therefore not have had the possibility of identifying interracial behavior as a domain in which attitude–behavior correlations were relatively low.

Conclusion

This review justifies a recommendation to use IAT and self-report measures jointly as predictors of behavior. Even though the relative predictive validities of the two types of measures varied considerably across domains, each type generally provided a gain in predictive validity relative to using the other alone. The review found that, for socially sensitive topics, predictive validity of self-report measures was remarkably low, and incremental validity of IAT measures was relatively high. In the studies examined in this review, high social sensitivity of topics was most characteristic of studies of racial and other intergroup behavior. In those topic domains, predictive validity of IAT measures significantly exceeded predictive validity of self-report measures.

References

* Indicates studies included in the meta-analysis.

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. Psychological Bulletin, 84, 888-918.

*Ames, S. L., Grenard, J. L., Thush, C., Sussman, S., Wiers, R. W., & Stacy, A. W. (2007). Comparison of Indirect Assessments of Association as Predictors of Marijuana Use Among At-Risk Adolescents. Experimental and Clinical Psychopharmacology, 15, 204-218.

*Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. Journal of Personality and Social Psychology, 91, 652-661.

*Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C. & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of the decided and undecided voters. Political Psychology, 29, 369-387.

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?" Psychological Inquiry, 15, 257-279.

*Asendorpf, J.B., Banse, R. & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. Journal of Personality and Social Psychology, 83, 380-393.

*Ashburn-Nardo, L., Knowles, M.L., & Monteith, M.J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. Social Cognition, 21, 61-87.

Ayres, I. (2001). Pervasive prejudice? Unconventional evidence of race and gender discrimination. Chicago: University of Chicago Press.

*Bain, T., Oakes, M., Ottaway, S., & Greenwald, A. G. (2004). Identifying the Unidentified: Predicting voting behavior of Independents using the Implicit Association Test. Unpublished manuscript.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roedeger, III, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), The nature of remembering: Essays in honor of Robert G. Crowder. (pp.117-150). Washington, DC: American Psychological Association.

Banaji, M.R., & Bhaskar, R. (2000). Implicit stereotypes and memory: The bounded rationality of social beliefs. In D.L. Schacter & E. Scarry (Eds.). Memory, brain, and belief. (pp. 139-175). Cambridge, MA: Harvard University Press.

Banaji, M.R., & Dasgupta, N. (1998). The consciousness of social beliefs: A program of research on stereotyping and prejudice. In V.Y. Yzerbyt, G. Lories & B. Dardenne (Eds). Metacognition: Cognitive and social dimensions. (pp. 157-170). Thousand Oaks, CA: Sage Publications.

*Banse, R. & Fischer, I. (2002). Implicit and explicit aggressiveness and the prediction of aggressive behaviour. Poster presented at the 11th conference on personality of the European Association of Personality Psychology, Jena.

*Banse, R. (2007). Implicit attitudes towards romantic partners and ex-partners: A test of the reliability and validity of the IAT. International Journal of Psychology, 42, 149-157.

*Banse, R., Grüne, J., & Kreft, V. (2002). Implicit attitudes towards romantic partners: Convergent and discriminant validity. Poster presented at the 13th General meeting of the European Association of Experimental Social Psychology, San Sebastian.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. Journal of Personality and Social Psychology, 62, 893-912.

Bargh, J.A., & Chartrand, T.L. (1999). The unbearable automaticity of being. American Psychologist, 54, 462-479.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 6, pp. 1-62). New York: Academic Press.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. American Psychologist, 61, 27 - 41.

*Bosson, J.K., Swann, W.B., Jr., & Pennebaker, J.W. (2000). Stalking the perfect measure of implicit self-esteem: The blind man and the elephant revisited? Journal of Personality and Social Psychology, 79, 631-643.

*Brochu, P. M., & Morrison, M. A. (2007). Implicit and explicit prejudice toward overweight and average-weight men and women: Testing their correspondence and relation to behavioral intentions. Journal of Social Psychology, 147, 681-706.

*Brockmyer, B., & Oleson, K. (2004, February). Implicit and explicit measures of age prejudice: Predictions for behavior. Poster Presented at the 6th Annual Meeting of the Society for Personality and Social Psychology, New Orleans, LA.

*Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition? Journal of Consumer Psychology, 14, 385–404.

*Brunstein, J.C. & Schmitt, C.H. (2004). Assessing individual differences in achievement motivation with the Implicit Association Test. Journal of Research in Personality, 38, 536-555.

*Carney, D. R. (2006). The faces of prejudice: On the malleability of the attitude–behavior link. Unpublished manuscript.

*Carney, D. R., Olson, K. R., Banaji, M. R., & Mendes, W. B. (2006). The faces of race-bias: Awareness of racial cues moderates the relation between bias and in-group facial mimicry. Unpublished manuscript, Harvard University, Cambridge, MA.

*Carpenter, S.J. (2000). Implicit gender attitudes. Unpublished doctoral dissertation, Yale University.

Chugh, D. (2004). Societal and managerial implications of implicit social cognition: Why milliseconds matter. Social Justice Research, 17, 203-222.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. (Revised Edition) New York: Academic Press.

*Conner, T., & Barrett, L.F. (2005). Implicit self attitudes predict spontaneous affect in daily life. Emotion, 4, 476-488.

Conrey, F.R., Sherman, J.W., Gawronski, B., Hugenberg, K., & Groom, C.J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. Journal of Personality and Social Psychology, 89, 469-487.

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. Psychological Bulletin, 87, 546-563.

Crowne, D.P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24, 349-354.

*Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Neural components of conscious and unconscious evaluations of Black and White faces. Psychological Science, 15, 806-813.

Cvencek, D., Greenwald, A. G., Brown, A., Gray, N. S., & Snowden, R. J. (2008). Faking of the Implicit Association Test is statistically detectable and partly correctable. Manuscript submitted for publication.

*Czopp, A.M., Monteith, M.J., Zimmerman, R.S., & Lynam, D.R. (2004). Implicit attitudes as potential Applied Social Psychology, 26, 227-237.protection from risky sex: Predicting condom use with the IAT. Basic and Applied Social Psychology, 26, 227-237.

*Dal Cin, S., Gibson B., Zanna, M.P., Shumate, R., & Fong, G.T. (2007) Smoking in movies, implicit associations of smoking with the self, and intentions to smoke. Psychological Science, 18, 559-563.

Dasgupta, N., & Greenwald, A.G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. Journal of Personality and Social Psychology, 81, 800-814.

*Dasgupta, N., & Rivera, L.M. (2006). From automatic antigay prejudice to behavior: The moderating role of conscious beliefs about gender and behavioral control. Journal of Personality and Social Psychology, 91, 268-280.

Dasgupta, N., McGhee, D. E., Greenwald, A.G., & Banaji, M. R. (2000). Automatic preference for White Americans: Ruling out the familiarity explanation. Journal of Experimental Social Psychology, 36, 316-328.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*.

*DeSteno, D., Valdesolo, P., & Bartlett, M. Y. (2006). Jealousy and the threatened self: Getting to the heart of the green-eyed monster. Journal of Personality and Social Psychology, 91, 626-641.

Devine, P.G. (1989). Stereotypes and prejudice: Their automatic and controlled components. Journal of Personality and Social Psychology, 56, 5-18.

Devine, P.G., Plant, E.A., Amodio, D.M., Harmon-Jones, E., & Vance, S.L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. Journal of Personality and Social Psychology, 82, 835-848.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. Journal of Experimental Social Psychology, 33, 510-540.

Eagly, A.H., Johannesen-Schmidt, M.C., & van Engen, M.L. (2003). Transformational, transactional, and laissez-faire leadership styles: a meta-analysis comparing women and men. Psychological Bulletin, 129, 569-591.

*Egloff, B. & Schmukle, S.C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. Journal of Personality and Social Psychology, 83, 1441-1455.

*Ellwart, T., Rinck, M., & Becker, E.S. (2006). From fear to love: Individual differences in implicit spider associations. Emotion, 6, 18-27.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. American Psychologist, 49, 709-724.

*Eyssel, F., & Bohner, G. (2007). Measuring sexist behavior in the laboratory: The role of implicit and explicit hostile sexism in predicting the endorsement of sexist humor. Sex Roles, 57, 651–660.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), Advances in experimental social psychology (Vol. 23, pp. 75-109). New York: Academic Press.

Fazio, R., Jackson, J., Dunton, B., & Williams, C. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? Journal of Personality and Social Psychology, 69, 1013-1027.

Fazio, R.H., & Olson, M.A. (2003). Implicit measures in social cognition research: Their meaning and uses. Annual Review of Psychology, 54, 297-327.

Fazio, R.H., & Zanna, M. P. (1981). Direct experience and attitude—behavior consistency. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 14, pp. 161—202). San Diego, CA: Academic Press.

Festinger, L. (1964). Behavioral support for opinion change. Public Opinion Quarterly, 28, 404-417.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. Basic and Applied Social Psychology, 27, 307–316.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). European Review of Social Psychology, 17, 74–147.

*Field, M., Mogg, K., & Bradley, B.P. (2004). Cognitive bias and drug craving in recreational cannabis users. Drug and Alcohol Dependence, 74, 105-111.

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. Psychological Review, 81, 59-74.

*Florack, A., Scarabis, M. & Bless, H. (2001). Der Einfluß wahrgenommener Bedrohung auf die Nutzung automatischer Assoziationen bei der Personen-beurteilung (The impact of perceived threat on the use of automatic associations in person judgments). Zeitschrift für Sozialpsychologie (continued as Social Psychology), 32, 249-259.

*Florack, A., Scarabis, M., & Gosejohann, S. (2004). Choices Between Binary Alternatives: On Automatic Preferences in Consumer Choice. Research Report. University of Basel.

*Friedman, M., Nosek, B.A., Gordon, K., Miller, I., & Banaji, M.R. (2001). Implicit hopelessness and depressive symptoms among married couples. Unpublished manuscript.

*Friese, F., Bluemke, M., & Wänke, M. (2007). Predicting voter behavior with implicit attitude measures: The 2002 German parliamentary election. Experimental Psychology, 54, 247-255.

*Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses takes over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. British Journal of Social Psychology, 47, 397-419.

*Friese, M., Wänke, M., & Plessner, H. (2006). Implicit consumer preferences and their influence on product choice. Psychology and Marketing, 23, 727-740.

*Gabriel, U., Banse, R., & Hug, F. (2007). Predicting private and public helping behavior by implicit prejudice and the motivation to control prejudiced reactions. British Journal of Social Psychology, 46, 365-382.

Gaertner, S.L., & Dovidio, J.F. (1986). The aversive form of racism. In J. F. Dovidio, & S. L. Gaertner (Eds). Prejudice, discrimination, and racism. (pp. 61-89). San Diego, CA: Academic Press.

*Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J., & Klauer, K.C. (2003). It's in the mind of the beholder: Individual differences in associative strength moderate category based and individuating impression formation. Journal of Experimental Social Psychology, 39, 16-30.

*Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. European Journal of Social Psychology, 33, 573-589.

*Gibson, B. (2008). Can evaluative conditioning change attitudes toward mature brands? New evidence from the implicit association test. Journal of Consumer Research, 35, 178-188.

*Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. Journal of Experimental Social Psychology, 44, 164–172.

*Gray, N.S., Brown, A.S., MacCulloch, M.J., Smith, J., & Snowden, R.J. (2005). An implicit test of the associations between children and sex in pedophiles. Journal of Abnormal Psychology, 114, 304-308.

*Green, A.R., Carney, D.R., Pallin, D.J., Ngo, L.H., Raymond, K.L., Iezzoni, L.I., & Banaji, M.R. (2007). The presence of implicit bias in physicians and its prediction of thrombolysis decisions for black and white patients. Journal of General Internal Medicine, 22, 1231-1238.

Greenwald, A. G., & Nosek, B.A. (2008). Attitudinal dissociation: What does it mean? In Petty, R. E., Fazio, R. H., & Briñol, P. (Eds.), Attitudes: Insights from the new implicit measures (Pp. 65–82). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: 1. An improved scoring algorithm. Journal of Personality and Social Psychology, 85, 197-216.

Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 102, 4-27.

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.

Greenwald, A.G., & Nosek, B.A. (2001). Health of the Implicit Association Test at age 3. Zeitschrift für Experimentelle Psychologie, 48, 85-93.

Greenwald, A.G., 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.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. Journal of Personality and Social Psychology, 74, 1464-1480.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. Psychological Review, 108, 814-834.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

*Heider, J. D., & Skowronski, J. J. (2007). Improving the predictive validity of the Implicit Association Test . North American Journal of Psychology, 9, 53-76.

*Hofmann, W., & Friese, M. (2008). Impulses got the better of me: Alcohol moderates the influence of implicit attitudes toward food cues on eating behavior. Journal of Abnormal Psychology, 117, 420-427.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. Personality and Social Psychology Bulletin, 31(10), 1369–1385.

*Hofmann, W., Gschwendner, T., Castelli, L., & Schmitt, M. (2008). Implicit and explicit attitudes and interracial interaction: The moderating role of situationally available control resources . Group Processes & Intergroup Relations, 11, 69-87.

*Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. Journal of Experimental Social Psychology, 43, 497-504.

*Houben, K., & Wiers, R.W. (2006a). Assessing implicit alcohol associations with the Implicit Association Test: Fact or artifact? Addictive Behaviors, 31, 1346-1362.

*Houben, K., & Wiers, R. W. (2006b). A test of the salience asymmetry interpretation of the alcohol-IAT. Experimental Psychology, 53, 292–300.

*Houben, K. & Wiers, R.W., (2007a). Are drinkers implicitly positive about drinking alcohol? Personalizing the alcohol-IAT to reduce negative extrapersonal contamination. Alcohol and Alcoholism. Vol. 42, No. 4, pp. 301–307.

*Houben, K., & Wiers, R. W. (2007b). Personalizing the Alcohol-IAT with individualized stimuli: Relationship with drinking behavior and drinking-related problems. Addictive Behaviors, 32, 2852-2864.

*Houben, K. & Wiers, R.W., (2008). Implicitly positive about alcohol? Implicit positive associations predict drinking behavior, Addictive Behaviors, 33, 979-986.

*Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. Psychological Science, 14, 640-643.

*Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. Psychological Science, 15, 342-345.

*Huijding, J., de Jong, P.J., Wiers, R.W., & Verkooijen, K. (2005). Implicit and explicit attitudes toward smoking in a smoking and a nonsmoking setting. Addictive Behaviors, 30, 949-961.

*Jajodia, A., & Earleywine, M. (2003). Measuring alcohol expectancies with the Implicit Association Test. Psychology of Addictive Behaviors, 17, 126–133.

*Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: Ingroup preferences and overt behaviors among gay and straight men. Personality and Social Psychology Bulletin, 30, 629-642.

Jordan, C.H., Spencer, S.J., Zanna, M.P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. Journal of Personality and Social Psychology, 85, 969-978.

*Kaminska-Feldman, M. (2004). The relation between two different implicit attitude measures: Asymmetry in self-others distance rating and Implicit Association Test. Polish Psychological Bulletin, 35, 65-70.

*Karpinski, A., & Hilton, J.L. (2001). Attitudes and the Implicit Association Test. Journal of Personality & Social Psychology, 81, 774-788.

*Karpinski, A., & Steinman, R.B. (2006). The single category Implicit Association Test as a measure of implicit social cognition. Journal of Personality and Social Psychology, 91, 16-32.

*Karpinski, A., Steinman, R.B., & Hilton, J.L. (2005). Attitude importance as a moderator of the relationship between implicit and explicit attitude measures. Personality and Social Psychology Bulletin, 31, 949-962.

Kelman, H.C. (1974). Attitudes are alive and well and gainfully employed in the sphere of action. American Psychologist, 29, 310-324.

Kim, D-Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). Social Psychology Quarterly, 66, 83–96.

Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. Personality and Social Psychology Bulletin, 21, 58–75.

Krosnick, J.A. (1988). The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. Journal of Personality and Social Psychology, 55, 196-210.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. Psychological Inquiry, 10, 83-109.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) (Pp. 59–102). In B. Wittenbrink & N. S. Schwarz (Eds.). Implicit measures of attitudes: Procedures and controversies. New York: Guilford Press

*Lemm K.M. (2000). Personal and social motivation to respond without prejudice: Implications for implicit and explicit attitudes and behavior. Unpublished doctoral dissertation, Yale University.

*Levesque, C., & Brown, K.W. (2004). Overriding motivational automaticity: Mindfulness as a moderator of the influence of implicit motivation on day-to-day behavior. Unpublished manuscript, Southwest Missouri State University.

Lipsey, M.W. & Wilson, D.B. (2001). Practical meta-analysis. Thousand Oaks, CA: SAGE Publications, Inc.

*Livingston, R. W. (2002). Bias in the absence of malice: The paradox of unintentional deliberative discrimination. Unpublished manuscript.

*Maison, D., Greenwald, A.G., & Bruin, R. (2001). The Implicit Association Test as a measure of consumer attitudes. Polish Psychological Bulletin, 2, 79.

*Maison, D., Greenwald, A.G., & Bruin, R.H. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes and behavior. Journal of Consumer Psychology, 14, 405-415.

*Maner, J. K., Kenrick, D. T., Neuberg, S. L., Becker, D. V., Robertson, T., Hofer, B., Delton, A., Butner, J., & Schaller, M. (2005). Functional projection: How fundamental social motives can bias interpersonal perception. Journal of Personality and Social Psychology, 88, 63-78.

*Marsh, K.L., Johnson, B.T., & Scott-Sheldon, L.A.J. (2001). Heart versus reason in condom use: Implicit vs. explicit attitudinal predictors of sexual behavior. Zeitschrift fur Experimentelle Psychologie, 48, 161-175.

*Martens, A., Jonas, Zanna, M., & Greenberg, J. (2005). Investigating the association between death-denial and cultural worldview defense. Poster presented at the 2005 meeting of the Society for Personality and Social Psychology.

*Mauss, I.B., Evers, C., Wilhelm, F.H., & Gross, J.J. (2006). How to bite your tongue without blowing your top: Implicit evaluation of emotion regulation predicts affective responding to anger provocation. Personality and Social Psychology Bulletin, 32, 589-602.

*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.

*McGraw, K.M. & Mulligan, K. (2008). Implicit political attitudes: Measurement and consequences for political judgment. Unpublished manuscript, Ohio State University.

McGregor, I., & Marigold, D.C. (2003). Defensive zeal and the uncertain self: What makes you so sure? Journal of Personality and Social Psychology, 85, 838-852.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. Zeitschrift für Experimentelle Psychologie, 48, 107-122.

*Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. Neuron, 50, 655-663.

Monteith, M.J., Voils, C.I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. Social Cognition, 19, 395-417.

*Neumann, R., Hülsenbeck, K., & Seibt, B. (2004). Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. Journal of Experimental Social Psychology, 40, 543-550.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

*Nock, M. K., & Banaji, M. R. (2007a). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. Journal of Consulting and Clinical Psychology, 75, 707-715.

*Nock, M.K., & Banaji, M.R. (2007b). Assessments of self-injurious thoughts using a behavioral test. American Journal of Psychiatry, 164, 820-823.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. Journal of Experimental Psychology: General, 134, 565-584.

Nosek, B. A., & Banaji, M. R. (2002). [Polish language] (At least) two factors moderate the relationship between implicit and explicit attitudes. In R.K. Ohme & M. Jarymowicz (Eds.) Natura Automatyzmow (pp. 49-56), Warszawa: WIP PAN & SWPS.

Nosek, B. A., Banaji, M. R., & Greenwald, A.G. (2002b). Harvesting implicit group attitudes and beliefs from a demonstration website. Group Dynamics, 6, 101-115.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review (pp. 265–292). In J. A. Bargh (Ed.), Automatic processes in social thinking and behavior. Psychology Press.

*Nosek, B. A., Hansen, J. (2008). Personalizing the Implicit Association Test increases explicit evaluation of target concepts. European Journal of Personality Assessment, in press.

*Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002a). Math = male, me = female, therefore math ? me. Journal of Personality and Social Psychology, 83, 44-59.

Nosek, B.A., Greenwald, A.G., & Banaji, M.R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. Personality and Social Psychology Bulletin, 31, 166-180.

*Olson, K.R., Parsons, A., Rowatt, W., Nosek, B., Mangu-Ward, K., & Banaji, M. R. (2006). Implicit and explicit attitudes toward gays and lesbians: Evidence from national webbased and university lab-based samples. Unpublished manuscript.

*Olson, M.A., & Fazio, R.H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. Journal of Personality and Social Psychology, 86, 653-667.

Ottaway, S., Hayden, D. C., & Oakes, M. (2001). Implicit attitudes and racism: Effect of word familiarity and frequency on the Implicit Association Test. Social Cognition, 19, 97-144.

*Perugini, M. (2005). Predictive models of implicit and explicit attitudes. British Journal of Social Psychology, 44, 29-45.

Petty, R.E., Tormala, Z.L., Briñol, P., & Jarvis, W.B.G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. Journal of Personality and Social Psychology, 90, 21-41.

*Phelps, E.A., O'Conner, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C., & Banaji, M.R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. Journal of Cognitive Neuroscience, 12, 1-10.

*Plessner, H., Haar, T., Hoffman, K., Stark, R., & Wänke, M. (2006). Directness of attitude measure and speed of information processing as constituents of consumer's attitude behavior correspondence. Unpublished manuscript.

*Powell, A.J., & Williams, K.D. (2000). The role of racial attitudes and experience in crossracial identification among Asian, Caucasian and Eurasian Australians. Paper Presented at the Society for Australian Social Psychology.

Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. Journal of Experimental Social Psychology, 44, 386-396

*Redker, C., & Gibson, B. (2008). Music as an unconditioned stimulus: Positive and negative effects of country music on implicit attitudes, explicit attitudes, and product choice. Journal of Applied Social Psychology, in press.

*Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F, Wyland, C. L., Trawalter, S., & Shelton, J. N. (2003). An fMRI examination of the impact of interracial contact on executive function. Nature Neuroscience, 6, 1323-1328.

*Richeson, J.A., & Shelton, J.N. (2003). When prejudice doesn't pay: Effects of interracial contact on executive function. Psychological Science, 14, 287-290.

*Robinson, M.D., Meier, B.P., Zetocha, K.J., & McCaul, K.D. (2005). Smoking and the Implicit Association Test: When the contrast category determines the theoretical conclusions. Basic and Applied Social Psychology, 27, 201-212.

*Robinson, M.D., Mitchell, KA., Kirkeby, B.S., & Meier, B.P. (2006). The self as a container: Implications for implicit self-esteem and somatic symptoms. Metaphor and Symbol, 21, 147-167.

*Ronay, R., & Kim, D. (2006). Gender differences in explicit and implicit risk attitudes: A socially facilitated phenomenon. British Journal of Social Psychology, 45, 397-419.

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. Journal of Experimental Psychology: General, 133, 139-165.

*Rudman, L. A., & Ashmore, R. D. (2007).Discrimination and the Implicit Association Test. Group Processes & Intergroup Relations, 10, 359-372.

*Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. Journal of Social Issues, 57, 743-762.

*Rudman, L. A., & Lee, M. R. (2002). Implicit and explicit consequences of exposure to violent and misogynous rap music. Group Processes and Intergroup Relations, 5, 133-150.

Rudman, L.A. (2004). Social justice in our minds, homes, and society: The nature, causes, and consequences of implicit bias. Social Justice Research, 17, 129-142.

Rudman, L.A., Greenwald, A.G., Mellott, D.S., & Schwartz, J.L.K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. Social Cognition, 17, 437-465.

*Rudman, L.E., & Heppen, J. (2003). Implicit romantic fantasies and women's interest in personal power: A glass slipper effect? Personality and Social Psychology Bulletin, 29, 1357-1370.

*Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. Journal of Personality and Social Psychology, 91, 995-1008.

*Sargent, M.J., & Theil, A. (2001). When do implicit racial attitudes predict behavior? On the moderating role of attributional ambiguity. Unpublished manuscript.

*Scarabis, M., Florack, A. & Gosejohann, S. (2006). When consumers follow their feelings: The impact of affective or cognitive focus on the basis of consumers' choice. Psychology & Marketing, 23, 1005-1036.

Schmidt, F. L., Pearlman, K., Hunter, J. E., & Hirsch, H. R. (1985). Forty questions about validity generalization and meta-analysis. Personnel Psychology, 38, 697-798.

*Schnabel, K., Banse, R., & Asendorpf, J. B. (2006a). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: The Implicit Association Procedure (IAP). Experimental Psychology, 53, 69-76.

*Schnabel, K., Banse, R., & Asendorpf, J. B. (2006b). Assessment of implicit personality selfconcept using the Implicit Association Test (IAT): Concurrent assessment of anxiousness and angriness. British Journal of Social Psychology, 45, 373-396.

*Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. Journal of Experimental Social Psychology, 39, 75-82.

*Sherman, S.J., Rose, J.S., Koch, K., Presson, C.C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. Journal of Social and Clinical Psychology, 22, 13-39.

*Shoda, Y. & Zayas, V. (1999, October). Individual differences in automatic evaluative associations toward significant persons. A paper presented at the 1999 Annual Meeting of The Society of Experimental Psychology.

*Smoak, N. D., Glasford, D. E., Portnoy, D. B., Marsh, K. L., & Scott-Sheldon, L. A. J. (2006). Impulsive contributors to steady and casual partner condom use in an at-risk sample and implications for HIV prevention. Unpublished ms., University of Connecticut.

*Spicer, C.V., & Monteith, M.J. (2001). Implicit outgroup favoritism among African Americans and vulnerability to stereotype threat. Unpublished manuscript.

Steffens, M.C. (2004). Is the Implicit Association Test immune to faking? Experimental Psychology, 51, 165-179.

*Steffens, M.C., & König, S.S. (2006). Predicting spontaneous Big Five behavior with Implicit Association Tests. European Journal of Psychological Assessment, 22, 13-20.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. Personality and Social Psychology Review, 8, 220-247.

*Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude–behavior consistency for stigmatized behavior. Cognition and Emotion, 15, 207-230.

*Teachman, B. A. (2007). Evaluating implicit spider fear associations using the Go/No-go Association Task. Journal of Behavior Therapy and Experimental Psychiatry, 38, 157-167.

*Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. Journal of Abnormal Psychology, 110, 236-235.

*Teachman, B. A., Smith-Janik, S. B., & Saporito, J. (2007). Information processing biases and panic disorder: Relationships among cognitive and symptom measures. Behaviour Research and Therapy, 45, 1791-1811.

*Teachman, B., & Woody, S. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. Journal of Abnormal Psychology, 112, 100-109.

*Teachman, B.A. (2005). Information processing and anxiety sensitivity: Cognitive vulnerability to panic reflected in interpretation and memory biases. Cognitive Therapy and Research, 29, 479-499.

*Thush, C., & Wiers, R. W. (2007). Explicit and implicit alcohol-related cognitions and the prediction of future drinking in adolescents. Addictive Behaviors, 32, 1367-1383.

*Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2007). Apples and oranges? Comparing indirect measures of alcohol-related cognition predicting alcohol use in at-risk adolescents. Psychology of Addictive Behaviors, 21, 587-591.

*van den Wildenberg, E., Beckers, M., van Lambaart, F., Conrod, P. J., & Wiers, R. W. (2006). Is the strength of implicit alcohol associations correlated with alcohol-induced heart-rate acceleration? Alcoholism: Clinical and Experimental Research, 30, 1336-1348.

*Vanman, E. J., Saltz, J. L., Nathan, L. R., & Warren, J. A. (2004). Racial discrimination by low-prejudiced Whites facial movements as implicit measures of attitudes related to behavior . Psychological Science, 15, 711-714.

*Vantomme, D., Geuens, M., De Houwer, J., & De Pelsmacker, P. (2005). Implicit attitudes toward green consumer behavior. Psychologica Belgica, 45, 217-239.

*Vargas, P. T., von Hippel, W., & Petty, R.E. (2004). Using partially structured attitude measures to enhance the attitude–behavior relationship. Personality and Social Psychology Bulletin, 30, 197-211.

*Verplanken, B., Friborg, O., Wang, C.E., Trafimow, D.,& Woolf, K. (2007). Mental habits: Metacognitive reflection on negative self-thinking. Journal of Personality and Social Psychology, 92, 526-541.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D.T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), Handbook of social psychology (4th ed, Vol. 1, pp. 446-496). New York: McGraw-Hill.

Wicker, A. W., & Pomazal, R J. (1971). The relationship between attitudes and behavior as a function of specificity of attitude object and presence of a significant person during assessment conditions. Representative Research in Social Psychology, 2, 26-31.

Wicker, A.W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. Journal of Social Issues, 25, 41-78.

*Wiers, R. W., Houben, K., & de Kraker, J. (2007). Implicit cocaine associations in active cocaine users and controls. Addictive Behaviors, 32, 1284-1289.

*Wiers, R. W., van de Luitgaarden, J., van den Wildenberg, E., & Smulders, F. T. Y. (2005). Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. Addiction, 100, 806-819.

*Wiers, R.W., van Woerden, N.V., Smulders F. T. Y., & de Jong, P.T. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. Journal of Abnormal Psychology, 111, 648-658.

*Williams, K., Wheeler, L., Edwardson, M., & Govan, C. (2001). The Implicit Association Test as a measure of consumer involvement. Unpublished manuscript.

Wilson, T., Lindsey, S., & Schooler, T.Y. (2000). A model of dual attitudes. Psychological Review, 107, 101-126.

*Yabar, Y., Johnson, L., Miles, L., & Peace, V., (2006). Implicit behavioral mimicry: Investigating the impact of group membership. Journal of Nonverbal Behavior, 30, 97-113.

*Zayas, V., & Shoda, Y. (2005). Do automatic reactions elicited by thoughts of romantic partner, mother, and self relate to adult romantic attachment? Personality and Social Psychology Bulletin, 31, 1011-1025.

*Ziegert, J, C., & Hanges, P. J. (2005). Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. Journal of Applied Psychology, 90, 554-562.

Author Notes

Anthony G. Greenwald, Department of Psychology, University of Washington; T. Andrew Poehlman, Cox School of Management, Southern Methodist University; Eric Luis Uhlmann, Kellogg School of Management, Northwestern University; Mahzarin R. Banaji, Department of Psychology, Harvard University. This research was supported by National Science Foundation Grants SBR-9422242, SBR-9710172, SBR-9422241, and SBR-9709924 and by National Institute of Mental Health Grants MH-01533, MH-41328, and MH-57672. This work was also supported by a grant from the Third Millennium Foundation and a fellowship from the Rockefeller Foundation to M. R. Banaji. The authors are grateful to the many authors of papers reviewed in the meta-analysis who helped by providing additional useful data from their studies. Katie Lancaster provided important assistance to the authors in coding and organizing a major portion of the data set. The authors additionally thank Scott Akalis, John Bargh, Max Bazerman, Victoria Brescoll, Huajian Cai, Dolly Chugh, Geoffrey Cohen, Dario Cvencek, Alice Eagly, Jeff Ebert, Richard Eibach, Larisa Heiphetz, Kristin Lane, Brian Nosek, Elizabeth Levy Paluck, Laurie Rudman, Konrad Schnabel, and Greg Walton for their help and advice in response to material in earlier drafts. Please address correspondence to A. G. Greenwald, Department of Psychology, University of Washington, Seattle, WA 98195-1525, agg@u.washington.edu.

Appendix
Characteristics of the 184 Independent Samples Included in the Meta-Analysis

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
Ames, Grenard, Thush, Sussman, Wiers, & Stacy (2007)	1	1	121	1	drugs/tobacco	.137	3	attitude	.418	3	multiple	.058
Amodio & Devine (2006)	2	1	32	2	race (Bl/Wh)	.156	2	att/belief	.333	1	attitude	.171
	3	2	21	3	race (Bl/Wh)	.325	2	multiple				
Arcuri, Castelli, Galdi, Zogmaister, & Amadori (2008)	1	1	52	1	politics	.642	1	attitude				
	2	2	37	1	politics	.414	1	attitude				
Asendorpf, Banse, & Mücke (2002)	1	1	138	3	personality	.274	1	self	.260	3	self	.440
Ashburn-Nardo, Knowles, & Monteith (2003)	1	1	77	1	race (Bl/Wh)	.230	1	attitude	.152	7	belief	.400
*Bain, Oakes, Ottoway, & Greenwald (2004)	1	1	138	2	politics	.462	1	attitude	.672	2	belief	.660
*Banse & Fischer (2002)	1	1	94	1	personality	.219	2	self	-.160	1	self	.150
Banse (2007)	1	1	132	1	personality	.370	1	attitude	.480	1	attitude	.490
*Banse, Grüne, & Kreft (2002)	1	1	96	2	relationships	.133	1	attitude	.166	1	attitude	.230
Bosson, Swann, & Pennebaker (2000)	1	1	83	6	personality	.161	1	self	.396	4	self	.220
Brochu & Morrison (2007)	1	1	37	3	other intergroup	.199	1	attitude	.459	1	attitude	.210
*Brockmeyer & Olson (2004)	1	1	58	1	other intergroup	.071	1	attitude	.185	2	attitude	-.003
Brunel, Tietje, & Greenwald (2004)	1	1	50	2	consumer	.533	2	belief	.637	1	attitude	.504
Brunstein & Schmitt (2004)	1	1	44	1	personality	.498	1	self	-.170	1	self	-.108
	1	2	44	1	personality	-.009	1	self	-.079	1	self	-.028
*Carney (2006)	1	1	29	19	race (Bl/Wh)	.256	1	attitude	.007	1	attitude	.180
	1	2	33	19	race (Bl/Wh)	.058	1	attitude	-.004	1	attitude	.160
*Carney, Olson, Banaji, & Mendes (2006)	1	1	21	2	race (Bl/Wh)	.266	1	attitude	-.202	1	attitude	-.269
*Carpenter (2000)	1	1	125	1	gender/sex	.241	2	att/belief	.690	1	attitude	.459
Conner & Barrett (2005)	1	1	124	29	personality	.102	1	self	.255	1	self	.230
	2	2	84	13	personality	.086	1	self	.320	1	self	.001
Cunningham et al. (2004)	1	1	13	1	race (Bl/Wh)	.790	1	attitude	.508	1	attitude	.004
Czopp, Monteith, Zimmerman, & Lynam (2004)	1	1	132	3	relationships	.161	1	attitude	.218	1	attitude	.250

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
Dal Cin, Gibson, Zanna, Shumate, & Fong (2007)	1	1	52	1	drugs/tobacco	.461	1	self				
Dasgupta & Rivera (2006)	1	1	82	1	gender/sex	-.060	1	attitude				
Dasgupta & Rivera (2006)	2	1	67	1	gender/sex	.040	1	attitude				
DeSteno, Valdesolo, & Bartlett (2006)	1	1	46	1	relationships	.550	1	self	.170	1	self	-.070
Egloff & Schmukle (2002)	3	1	62	6	clinical	.133	1	self	.093	1	self	-.060
	4	2	33	7	clinical	.201	1	self	.129	1	self	.250
Ellwart, Rinck, & Becker (2006)	1	1	48	3	clinical	.205	1	attitude	.677	2	attitude	.220
	2	2	18	1	clinical	.635	1	attitude	.648	2	attitude	.270
Eyssel & Bohner (2007)	1	1	130	1	gender/sex	-.005	2	belief	.329	3	belief	.042
Field, Mogg, & Bradley (2004)	1	1	33	1	drugs/tobacco	.382	1	attitude				
Florack, Scarabis, & Bless (2001)	1	1	20	1	other intergroup	.580	1	attitude	.200	1	attitude	.210
	1	2	26	2	other intergroup	-.210	1	attitude	-.100	1	attitude	-.190
	1	3	21	3	other intergroup	.170	1	attitude	.630	1	attitude	.110
*Florack, Scarabis, & Gosejohann (2004)	1	1	105	1	consumer	.460	2	att/belief	.710	1	attitude	.470
	1	2	108	1	consumer	.480	2	att/belief	.750	1	attitude	.410
*Friedman, Nosek, Miller, Gordon, & Banaji (2001)	1	1	122	1	clinical	.380	1	self				
	1	2	122	1	clinical	.100	1	attitude				
Friese, Bluemke, & Wänke (2007)	1	1	1,386	10	politics	.302	5	attitude	.560	5	attitude	.409
Friese, Hofmann, & Wänke (2008)	1	1	42	1	consumer	.120	1	attitude	.600	1	attitude	.200
	1	2	43	1	consumer	.450	1	attitude	.240	1	attitude	.370
	2	3	33	1	consumer	-.050	1	attitude	.350	1	attitude	.410
	2	4	33	1	consumer	.290	1	attitude	-.080	1	attitude	.010
	3	5	21	1	drugs/tobacco	-.240	1	attitude	.380	1	attitude	.180
	3	6	25	1	drugs/tobacco	.500	1	attitude	.110	1	attitude	.370
Friese, Wänke, & Plessner (2006)	1	1	25	1	consumer	.080	1	attitude	.360	1	attitude	.390
	1	2	27	2	consumer	.470	1	attitude	.740	1	attitude	.260
Gabriel, Banse, & Hug (2007)	1	1	69	1	gender/sex	-.010	1	attitude	.174	2	attitude	.320
Gawronski, Ehrenberg, Banse, Zukova, & Klauer (2003)	1	1	119	1	gender/sex	.190	1	belief				
Gawronski, Geschke, & Banse (2003)	1	1	35	2	other intergroup	.181	1	attitude	.010	1	belief	.216
	1	2	34	2	other intergroup	.321	1	attitude	.110	1	belief	.161

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
Gibson (2008)	2	1	30	1	consumer	.487	1	attitude	.421	1	attitude	.349
	2	2	30	2	consumer	.139	1	attitude	.517	1	attitude	-.011
Glaser & Knowles (2008)	1	1	48	1	race (Bl/Wh)	.290	2	att/belief	.248	2	attitude	.298
Gray, Brown, MacCulloch, Smith, & Snowden (2005)	1	1	77	1	clinical	.355	1	belief				
Green, Carney, Pallin, Ngo, Raymond, Iezzoni, & Banaji (2007)	1	1	207	1	race (Bl/Wh)	.138	2	att/belief	.021	2	att/belief	.017
Heider & Skowronski (2007)	1	1	140	2	race (Bl/Wh)	.119	1	attitude	.042	2	attitude	.003
	2	2	55	4	race (Bl/Wh)	.272	1	attitude	.038	2	attitude	.284
Hofmann & Friese (2008)	1	1	29	1	consumer	-.190	1	attitude	.470	1	self	.020
	1	2	29	1	consumer	.400	1	attitude	.250	1	self	.260
Hofmann, Gschwendner, Castelli, & Schmitt (2008)	1	1	85	2	race (Bl/Wh)	.118	1	attitude	.154	1	attitude	.341
	2	2	76	2	race (Bl/Wh)	.192	1	attitude	.002	1	attitude	.001
Hofmann, Rauch, & Gawronski (2007)	1	1	26	1	consumer	.340	1	attitude	-.290	1	self	-.290
	1	2	24	2	consumer	-.090	1	attitude	.480	1	self	.480
Houben & Wiers (2006a)	1	1	96	2	drugs/tobacco	.153	3	multiple	.301	7	multiple	.107
Houben & Wiers (2006b)	1	1	46	2	drugs/tobacco	.386	1	attitude	.110	2	attitude	.346
Houben & Wiers (2007a)	1	1	42	2	drugs/tobacco	.410	1	attitude	.126	4	attitude	.330
Houben & Wiers (2007b)	1	1	46	2	drugs/tobacco	.289	2	attitude	.348	3	attitude	.145
Houben & Wiers (2008)	1	1	62	1	drugs/tobacco	.175	4	attitude	.341	2	attitude	.086
Hugenberg & Bodenhausen (2003)	1	1	24	1	race (Bl/Wh)	.460	1	attitude	.054	1	attitude	.360
	2	2	24	1	race (Bl/Wh)	.424	1	attitude	.187	1	attitude	-.129
Hugenberg & Bodenhausen (2004)	1	1	20	1	race (Bl/Wh)	.345	1	attitude	.032	1	attitude	-.068
	2	2	57	1	race (Bl/Wh)	.163	1	attitude	-.030	1	attitude	-.230
Huijding, de Jong, Wiers, & Verkooijen (2005)	1	1	48	1	drugs/tobacco	.450	1	attitude	.692	1	attitude	.221
Jajodia & Earleywine (2003)	1	1	103	3	drugs/tobacco	.307	1	belief	.387	1	belief	.132
Jellison, McConnell, & Gabriel (2004)	1	1	39	4	gender/sex	.258	1	attitude	.271	1	attitude	.510
Kaminska-Feldman (2004)	1	1	47	1	other intergroup	.227	1	attitude				
Karpinski & Hilton (2001)	2	1	81	1	consumer	.030	1	attitude	.360	3	multiple	.160
Karpinski & Steinman (2006)	1	1	53	1	consumer	.277	3	attitude	.540	2	attitude	.101

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
Karpinski, Steinman, & Hilton (2005)	1	1	155	1	politics	.420	1	attitude	.721	2	attitude	.460
Karpinski, Steinman, & Hilton (2005)	2	2	109	1	consumer	.353	1	attitude	.698	1	attitude	.290
	3	3	72	1	consumer	.550	1	attitude	.960	2	attitude	.520
*Lemm (2000)	1	1	33	1	gender/sex	.380	1	attitude	-.160	1	belief	.130
*Levesque & Brown (2004)	2	1	69	1	personality	.140	1	self	.440	1	self	.190
	3	2	78	2	personality	.060	1	self	.270	1	self	-.020
*Livingston (2002)	1	1	34	1	other intergroup	.370	1	attitude	.025	2	belief	
	2	2	34	1	other intergroup	.040	1	attitude	-.265	2	belief	
	2	3	31	2	race (Bl/Wh)	.430	2	attitude	.262	2	belief	.290
Maison, Greenwald, & Bruin (2001)	1	1	70	1	consumer	.200	1	attitude	.465	3	multiple	.380
	2	2	50	1	consumer	.340	1	attitude			attitude	
Maison, Greenwald, & Bruin (2004)	1	1	32	1	consumer	.535	1	attitude	.697	1	attitude	.474
	2	2	39	1	consumer	.352	1	attitude	.592	1	attitude	.431
	3	3	102	2	consumer	.572	1	attitude	.638	1	attitude	.404
Maner et al. (2005)	2	1	51	2	other intergroup	-.130	1	attitude				
Marsh, Johnson, & Scott-Sheldon (2001)	1	1	36	1	relationships	.107	2	belief	.130	4	multiple	.055
	1	2	71	1	relationships	-.085	2	belief	.461	3	multiple	-.032
	1	3	80	1	relationships	-.050	2	belief	.317	4	multiple	-.013
*Martens, Jonas, Zanna, & Greenberg (2004)	1	1	35	1	personality	.460	1	self	.120	1	attitude	-.120
Mauss, Evers, Wilhelm, & Gross (2006)	2	1	36	8	clinical	.345	1	self				
McConnell & Leibold (2001)	1	1	41	15	race (Bl/Wh)	.229	1	attitude	.085	1	attitude	.420
*McGraw & Mulligan (2003)	1	1	93	2	politics	.418	2	attitude	.549	2	self	.480
Mitchell, McCrae, & Banaji (2006)	1	1	15	2	politics	.434	2	belief	-.021	1	self	.173
Neumann, Hülsenbeck, & Seibt (2004)	1	1	37	2	gender/sex	.160	1	attitude	.238	1	attitude	.190
Nock & Banaji (2007a)	1	1	73	3	clinical	.376	1	self				
Nock & Banaji (2007b)	1	1	89	4	clinical	.493	2	belief				
Nosek & Hansen (2008)	2	1	926	1	politics	.647	1	attitude	.866	2	attitude	.637
	4	2	1,028	2	consumer	.304	1	attitude	.554	2	attitude	.370

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
	5	3	82	1	politics	.552	2	attitude	.837	2	attitude	.576
	7	4	203	2	consumer	.459	1	attitude	.780	2	att/belief	.499
Nosek, Banaji, & Greenwald (2002)	1	1	227	1	personality	.380	1	attitude	.490	1	attitude	.420
Olson & Fazio (2004)	3	1	26	2	consumer	.185	1	attitude	.870	1	attitude	.150
	3	2	33	2	consumer	.526	1	attitude	.880	1	attitude	.670
	4	3	12	1	politics	.310	1	attitude	.830	1	attitude	.560
	4	4	9	1	politics	.610	1	attitude	.850	1	attitude	.730
*Olson et al. (2006)	1	1	76	1	gender/sex	.337	2	attitude	.512	2	attitude	.342
Perugini (2005)	1	1	48	1	drugs/tobacco	.640	1	attitude	.480	1	attitude	.480
	2	2	109	2	consumer	.190	1	attitude	.278	1	attitude	.090
Phelps et al. (2000)	1	1	12	1	race (Bl/Wh)	.576	1	attitude	-.047	1	belief	
*Plessner, Haar, Hoffman, Stark, & Wänke (2006)	1	1	40	1	consumer	.572	1	attitude	.662	1	attitude	.443
	2	2	109	2	consumer	.093	2	attitude	.301	3	attitude	.138
*Powell & Williams (2000)	1	1	55	1	other intergroup	.313	1	attitude				
*Redker & Gibson (2008)	1	1	68	1	consumer	.262	1	attitude	.357	1	attitude	.210
Richeson & Shelton (2003)	1	1	21	2	race (Bl/Wh)	.474	1	attitude	.389	1	attitude	.251
Richeson et al. (2003)	1	1	15	3	race (Bl/Wh)	.554	1	attitude				
	2	2	15	2	race (Bl/Wh)	.610	1	attitude				
Robinson, Meier, Zetocha, & McCaul (2005)	1	1	48	1	drugs/tobacco	.520	1	attitude	.491	1	attitude	.480
	2	2	52	1	drugs/tobacco	.260	2	attitude	.456	1	attitude	.400
Robinson, Mitchell, Kirkeby, & Meier (2006)	1	1	96	1	clinical	.210	1	self				
	2	2	61	4	clinical	.309	1	self				
Ronay & Kim (2006)	1	1	126	2	personality	.204	2	belief	.134	3	belief	.090
Rudman & Ashmore (2007)	1	1	64	3	race (Bl/Wh)	.265	2	att/belief	.395	2	att/belief	.268
	2	2	89	1	other intergroup	.380	1	belief	.239	3	multiple	.530
	2	3	89	2	other intergroup	.275	2	att/belief	.181	3	multiple	.192
	2	4	126	3	race (Bl/Wh)	.205	2	att/belief	.033	3	multiple	.225
Rudman & Glick (2001)	1	1	27	2	gender/sex	.367	1	belief	.095	1	belief	.040
	1	2	19	2	gender/sex	.408	1	belief	.297	1	belief	.040
Rudman & Heppen (2003)	1	1	77	3	gender/sex	.318	1	belief	.057	1	belief	.170
	2	2	121	3	gender/sex	.167	1	belief	-.033	1	belief	-.090
	3	3	73	4	gender/sex	.269	2	att/belief	-.104	2	att/belief	-.242

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
Rudman & Lee (2002)	2	1	38	3	race (Bl/Wh)	.232	1	belief	.147	2	belief	.190
Rydell & McConnell (2006)	4	1	29	2	relationships	.229	1	attitude	.433	1	attitude	-.030
*Sargent & Theil (2001)	1	1	38	1	race (Bl/Wh)	.320	1	attitude	.070	1	belief	.010
Scarabis, Florack, & Gosejohann (2006)	1	1	25	1	consumer	.595	2	belief	.373	2	attitude	.481
	1	2	24	2	consumer	.287	2	belief	.282	2	attitude	.397
	1	3	25	3	consumer	.165	2	belief	.301	2	attitude	.440
	1	4	24	4	consumer	.066	2	belief	.320	2	attitude	-.039
Schnabel, Banse, & Asendorpf (2006a)	1	1	58	1	clinical	.170	1	self	.360	1	self	.150
Schnabel, Banse, & Asendorpf (2006b)	1	1	100	4	clinical	.021	2	self	.193	7	self	.181
Sekaquaptewa, Espinoza, Thompson, Vargas, & von Hippel (2003)	1	1	79	1	race (Bl/Wh)	.030	1	attitude	.010	1	belief	.160
Sherman, Rose, Koch, Presson, & Chassin (2003)	1	1	54	1	drugs/tobacco	.120	1	attitude				
*Shoda & Zayas (1999)	2	1	84	6	relationships	.060	3	belief	.225	3	belief	.217
	3	2	40	8	relationships	.369	1	belief	.097	1	belief	-.054
*Smoak, Glasford, Portnoy, Marsh, & Scott-Sheldon (2006)	1	1	19	3	relationships	.341	1	belief	.205	3	belief	-.004
*Spicer & Monteith (2001)	2	1	78	7	race (Bl/Wh)	.146	1	attitude				
Steffens & König (2006)	1	1	89	7	personality	.192	5	self	.110	5	self	.091
Swanson, Rudman, & Greenwald (2001)	2	1	101	2	consumer	.250	2	att/belief	.403	2	att/belief	.473
	2	2	98	1	drugs/tobacco	.221	2	att/belief	.547	2	att/belief	.166
	3	3	70	1	drugs/tobacco	.357	2	att/belief	.486	2	att/belief	.276
Teachman & Woody (2003)	1	1	59	3	clinical	.243	4	multiple	.649	1	belief	.340
Teachman (2005)	1	1	103	3	clinical	.277	1	self				
Teachman (2007)	1	1	32	3	clinical	.562	1	attitude	.665	2	attitude	.340
Teachman, Gregg, & Woody (2001)	1	1	67	1	clinical	.543	4	attitude	.856	1	attitude	
Teachman, Smith-Janik, & Saporito (2007)	1	1	81	3	clinical	.271	1	self	.662	3	multiple	.260
Thush & Wiers (2007)	1	1	100	2	drugs/tobacco	.198	3	multiple	.341	3	multiple	.074
Thush, Wiers, Ames, Grenard, Sussman, & Stacy (2007)	1	1	81	1	drugs/tobacco	.075	3	belief	.378	3	belief	.023

citation	expt	sam-ple	N	N crit	topic	ICC	N iat	IAT type	ECC	N expl	expl type	IEC
Van den Wildenberg, Beckers, van Lambaart, Conrod, & Wiers (2006)	1	1	48	2	drugs/tobacco	.183	2	att/belief	.132	2	belief	-.033
Vanman, Saltz, Nathan, & Warren (2004)	1	1	59	1	race (Bl/Wh)	.170	1	attitude	.251	1	attitude	.042
	1	2	21	3	race (Bl/Wh)	.024	1	attitude	.193	1	attitude	.106
Vantomme, Geuens, De Houwer, & De Pelsmacker (2005)	1	1	60	2	consumer	.224	1	attitude	.386	1	attitude	.190
	2	2	67	2	consumer	.298	1	attitude	.493	1	attitude	.330
Vargas, Von Hippel & Petty (2004)	4	1	226	1	personality	.170	1	self	.587	3	multiple	.120
Verplanken, Friborg, Wang, Trafimow, & Woolf (2007)	5	1	125	2	personality	.201	1	self	.441	1	self	.065
Wiers, Houben, & de Kraker (2007)	1	1	32	2	drugs/tobacco	.227	3	multiple	.213	3	multiple	.109
Wiers, van de Luitgaarden, van den Wildenberg, & Smulders (2005)	1	1	92	5	drugs/tobacco	.119	2	att/belief	.035	2	att/belief	.200
Wiers, van Woerden, Smulders, & de Jong (2002)	1	1	48	1	drugs/tobacco	.335	2	att/belief	.320	5	multiple	.183
*Williams, Wheeler, Edwardson, & Govan (2001)	1	1	74	2	consumer	.334	1	self				
Yabar, Johnston, Miles, & Peace (2006)	2	1	48	1	other intergroup	.099	1	attitude	-.282	1	attitude	.550
Zayas & Shoda (2005)	1	1	58	1	relationships	.280	1	attitude				
	2	2	85	3	relationships	.207	3	multiple				
Ziegert & Hanges (2005)	1	1	99	1	race (Bl/Wh)	.259	1	attitude	.056	2	attitude	.116

* = unpublished report; expt = experiment number in report; sample = ordinal independent sample in report; N = number of subjects in independent sample; N crit = number of distinct criterion measures in independent sample; topic = classification into one of the 9 criterion domains used in several analyses ("race (Bl/Wh)" = Black–White race); ICC = IAT–criterion average effect size (*r*) for sample; N iat = number of distinct IAT measures in independent sample; IAT type = classification of IAT measures as attitude, belief, mixture of attitude and belief ("att/belief"), self-concept or self-esteem ("self"), or other mixtures of types ("multiple"); ECC = self-report–criterion average effect size (*r*) for sample; N expl = number of distinct self-report measures in independent sample; expl type = same codes as for IAT type, applied to self-report measures; IEC = IAT–self-report average effect size (*r*) for sample.

Table 1. Description of Conceptual Moderator Variables

	Moderators in analyses of IAT–criterion correlations (ICCs)					Moderators in analyses of explicit–criterion correlations (ECCs)				
Moderator Definition	*k*	Min	Max	Mean	SD	*k*	Min	Max	Mean	SD
IAT–explicit correlation (IEC; Fisher *Z*-transformed)	152	-0.30	0.93	0.23	0.23	152	-0.30	0.93	0.23	0.23
predictor type: attitude = 1; other = 0	184	0.0	1.0	0.69	0.42	156	0.0	1.0	0.64	0.44
social sensitivity of response to the predictor[a] (range = 1–7)	184	1.0	7.0	3.93	2.17	154	1.0	7.0	3.73	2.16
controllability of response to the criterion measure (range = 0–10)	184	0.0	10.0	6.15	2.61	156	0.0	10.0	6.31	2.59
correspondence between IAT or self-report measure and criterion measure (range = 1–7)	184	1.0	5.0	3.20	1.13	155	1.0	5.0	3.26	1.11
complementarity of alternative concepts used in IAT measures[b] (range = 1–9)	177	1.0	8.5	2.29	1.79	149	1.0	8.0	2.23	1.76

Note. Min = minimum observed value of moderator; Max = maximum observed value. Mean = average value of the moderator across the *k* independent samples; SD = standard deviation. Numbers of independent samples (*k*) are sometimes less than their maxima of *k* = 184 for ICCs and *k* = 156 for ECCs because reports did not always contain sufficient information to code the moderator.

[a] For IAT measures, this was rated social sensitivity of responding to a *self-report* measure of the measured attitude, belief, or self-concept predictor.

[b] For self-report measures, the (average) rated complementarity for the study's IAT measure(s) was used as the moderator (see text).

Table 2. Description of Methodological and Publication Moderator Variables

	Predictors of IAT–criterion correlations (ICCs)					Predictors of explicit–criterion correlations (ECCs)				
Moderator Definition	*k*	Min	Max	Mean	SD	*k*	Min	Max	Mean	SD
Number of effect sizes available in the independent sample	184	1	29	2.86	3.49	156	1	29	3.57	4.53
Mean sample size, averaged over effect sizes in the independent sample	184	9	1386	81.0	141.5	156	9	1386	83.8	152.9
Number of IAT measures obtained from each subject	184	1	6	1.51	0.97	—	—	—	—	—
Criterion data collection method: subject response = 0; experimenter observation = 1	184	0	1	0.33	0.46	156	0	1	0.35	0.46
IAT scoring method: *D* algorithm = 1; other = 0	145	0	1	0.48	0.50	—	—	—	—	—
Predictor–criterion ordinal position relation: predictor first = 1; counterbalanced = 2; predictor last = 3	156	1	3	1.81	0.95	126	1	3	1.82	0.90
Predictor–criterion session relation: predictor and criterion in same session = 0; separate sessions = 1	171	0	1	0.18	0.38	141	0	1	0.18	0.39
Publication year	184	1999	2008	2004.6	2.37	156	1999	2008	2004.7	2.38
Publication status: unpublished = 0; published = 1	184	0	1	0.83	0.38	156	0	1	0.83	0.37

Note. Min = minimum observed value of moderator; Max = maximum observed value. SD = standard deviation. Numbers of independent samples (*k*) are sometimes less than their maxima of *k* = 184 for ICCs and *k* = 156 for ECCs because reports did not always contain sufficient information to code the status of the moderator. The third and fifth moderators applied only to IAT measures. The *D* algorithm of the fifth moderator is the scoring procedure introduced by Greenwald et al. (2003).

Table 3. Weighted Mean Effect Sizes and Homogeneity Tests for ICCs, ECCs, and IECS in all Independent Samples and within Nine Criterion Measure Domains

	IAT–criterion correlations (ICCs)				explicit–criterion correlations (ECCs)				implicit–explicit correlations (IECs)			
Criterion domain	*r* (95% CI)	*k*	*N*	*SD*	*r* (95% CI)	*k*	*N*	*SD*	*r* (95% CI)	*k*	*N*	*SD*
All independent samples	.274 (±.029)	184	14,900	.215	.361 (±.056)	156	13,068	.391	.214 (±.039)	155	13,121	.258
race (White vs. Black)	.236 (±.062)†	32	1,699	.186	.118 (±.108)†	28	1,568	.295	.117 (±.074)†	27	1,589	.198
other intergroup behavior	.201 (±.093)†	15	678	.189	.120 (±.165)	12	525	.297	.148 (±.115)	12	544	.207
gender and sexual orientation	.181 (±.081)†	15	1,094	.164	.224 (±.151)	12	828	.279	.172 (±.101)	12	876	.182
consumer preferences	.323 (±.049)	40	3,257	.171	.546 (±.065)	38	3,126	.258	.319 (±.056)	38	2,994	.190
political preferences	.483 (±.071)	11	2,903	.145	.709 (±.094)	9	2,810	.231	.537 (±.082)	9	2,858	.158
personality traits	.277 (±.064)	24	1,456	.169	.353 (±.105)	21	1,317	.270	.166 (±.078)†	21	1,326	.186
alcohol and drug use	.221 (±.069)	16	1,718	.147	.269 (±.121)	16	1,712	.262	.159 (±.080)	16	1,736	.166
clinical (e.g., phobia, anxiety)	.296 (±.068)	19	1,318	.161	.537 (±.127)	10	547	.257	.248 (±.113)	10	558	.190
close relationships	.171 (±.094)	12	777	.169	.247 (±.164)†	10	635	.279	.091 (±.116)†	10	640	.189

Note. Note. Aggregate effect sizes were computed for Fisher's Z-transformed *r* values. For "all independent samples" weighted mean effect sizes (*r*), their 95% confidence intervals (CIs), and their weighted standard deviations (*SD*s), transformed back to the *r* metric, were obtained from a random effects test. For the nine categories, these results were from a mixed model analysis of variance of differences among the categories. *k* = number of samples associated with each weighted mean effect size; *N* = summed numbers of subjects in the *k* samples.

† $p > .05$ for homogeneity test (i.e., homogeneous effect sizes), from fixed-effect analysis of the nine categories. All category aggregate effect sizes not marked with"†" were significantly heterogeneous (i.e., $p \leq .05$ for homogeneity test).

Table 4. Tests of Weighted Regression Models for Conceptual Moderators of IAT–Criterion Correlations (ICCs)

	Weighted Regression Analyses of ICCs								
	Univariate effects					Simultaneous effects (*k* = 145)			
Moderators[a]	*B*	*β*	*k*	*z*	*p*	*B*	*β*	*z*	*p*
IEC	.471	.547	152	7.92	10^{-15}	.384	.451	5.49	10^{-8}
predictor type	.059	.126	184	1.57	.12	.024	.051	0.68	.50
social sensitivity	-.017	-.184	184	-2.31	.02	-.020	-.220	-2.18	.03
controllability	.006	.079	184	0.97	.33	-.001	-.018	-0.25	.81
correspondence	.044	.254	184	3.27	.001	-.026	-.151	-1.46	.15
complementarity	.032	.313	177	4.24	10^{-5}	.025	.259	3.24	.001

Note: Analyses were conducted using Fisher's *Z*-transformed *r* values and mixed effects models (fixed slopes, random intercepts). Summary statistics for the simultaneous regression analysis: R^2 = .401, 2-tailed $p = 10^{-18}$; random effects variance component = .0058; mean effect size (*r*) = .280. *k* = number of samples in each analysis; *B* = unstandardized regression coefficient; β = standardized regression coefficient; *z* = critical ratio test for the regression coefficient; *p* = 2-tailed probability of *z*. IEC = Fisher's *Z*-transformed implicit–explicit correlation.

[a] See Table 1 for descriptions of the conceptual moderator variables.

Table 5. Tests of Weighted Regression Models for Conceptual Moderators of Explicit–Criterion Correlations (ECCs)

	Weighted Regression Analyses of ECCs								
	Univariate effects					Simultaneous effects (k = 144)			
Moderators[a]	B	β	k	z	p	B	β	z	p
IEC	.936	.586	152	9.17	10^{-20}	.471	.292	4.28	10^{-5}
predictor type	.262	.313	156	3.74	.0002	.012	.015	0.23	.82
social sensitivity	-.084	-.494	154	-6.26	10^{-10}	-.037	-.217	-2.52	.01
controllability	.043	.296	156	3.32	.0009	.011	.075	1.20	.23
correspondence	.193	.594	155	8.62	10^{-17}	.090	.278	3.13	.002
complementarity	.064	.308	149	3.85	.0001	.037	.187	2.87	.004

Note: Analyses were conducted using Fisher's *Z*-transformed *r* values and mixed effects models (fixed slopes, random intercepts). Summary statistics for the simultaneous regression analysis: R^2 = .554; 2-tailed $p = 10^{-35}$; random effects variance component = .0368; mean effect size (*r*) = .393. *k* = number of samples in each analysis; *B* = unstandardized regression coefficient; β = standardized regression coefficient; *z* = critical ratio test for the weighted coefficient; *p* = 2-tailed probability of *z*. IEC = Fisher's *Z*-transformed implicit–explicit correlation.

[a] See Table 1 for descriptions of the conceptual moderator variables.

Table 6. Tests of Weighted Regression Models for Conceptual Moderators of Implicit–Explicit Correlations (IECs)

	Weighted Regression Analyses of IECs								
	Univariate effects					Simultaneous effects (k = 145)			
Moderators[a]	*B*	*β*	*k*	*z*	*p*	*B*	*β*	*z*	*p*
IAT predictor type	.167	.310	152	3.63	.0003	.056	.105	0.88	.38
self-report predictor type	.157	.310	152	3.64	.0003	.048	.094	0.82	.41
social sensitivity	-.033	-.313	152	-3.58	.0003	-.029	-.268	-2.42	.02
controllability	.028	.312	152	3.32	.0009	.009	.100	1.27	.21
correspondence	.081	.398	152	4.93	10^{-6}	.012	.060	0.52	.61
complementarity	.046	.378	145	4.49	10^{-5}	.038	.322	3.91	.0001

Note: Analyses were conducted using Fisher's *Z*-transformed *r* values and mixed effects models (fixed slopes, random intercepts). Summary statistics for the simultaneous regression analysis: R^2 = .341, 2-tailed $p = 10^{-12}$; random effects variance component = .0180; mean effect size (*r*) = .244. *k* = number of samples in each analysis; *B* = unstandardized regression coefficient; *β* = standardized regression coefficient; *z* = critical ratio test for the weighted coefficient; *p* = 2-tailed probability of *z*. The social sensitivity and correspondence predictors were averages of separate ratings for the IAT and self-report predictors.

[a] See Table 1 for descriptions of the conceptual moderator variables.

Table 7. Unweighted Correlations Among Variables in the Regression Analyses of Tables 4, 5, and 6[a]

	Variables as numbered in the left column						
Variables	1	2	3	4	5	6	7
1. ICC or ECC		.574	.293	-.502	.389	.595	.302
2. IEC	.429		.285	-.351	.352	.412	.333
3. predictor type	.173	.298		-.133	.186	.380	.117
4. social sensitivity	-.141	-.351	-.037		-.385	-.695	.095
5. controllability	.055	.340	.120	-.347		.310	.129
6. correspondence	.186	.406	.331	-.682	.274		.162
7. complementarity	.312	.329	.279	.107	.132	.170	

Note. Correlations below the diagonal used the 145 independent samples included in the simultaneous regression of predictors of IAT–criterion correlations (ICCs) in the right side of Table 2. Those above the diagonal used the 144 samples in the simultaneous regression of explicit–criterion correlations (ECCs) in the right side of Table 3. The variables representing ICCs, ECCs, and implicit–explicit correlations (IECs) were Fisher Z-transformed values of aggregated correlations of each type within independent samples.

[a] For sample size = 144, the minimum correlation associated with a p value of .005 is r = .233.

Table 8. Tests of Weighted Regression Models for Methodological and Publication Moderators of IAT–Criterion Correlations (ICCs)

	Weighted Regression Analyses of ICCs								
	Univariate effects					Simultaneous effects (*k* = 129)			
Moderators[a]	*B*	*β*	*k*	*z*	*p*	*B*	*β*	*z*	*p*
Number of effect sizes	-.010	-.189	184	-2.32	.02	-.012	-.171	-1.75	.08
Mean sample size	.000	.101	184	1.24	.22	.000	.022	0.22	.83
Number of IATs	-.016	-.083	184	-1.00	.32	-.018	-.098	-0.94	.35
Criterion data collection method: subject response vs. observation	-.058	-.128	184	-1.58	.11	-.089	-.205	-2.20	.03
IAT scoring method	.018	.044	145	0.48	.63	.057	.146	1.23	.22
Predictor-criterion ordinal position	.020	.101	156	1.28	.20	.023	.108	1.16	.25
Predictor-criterion session relation	.024	.046	171	0.53	.60	.083	.159	1.70	.09
Publication year	.005	.065	184	0.80	.42	-.004	-.047	-0.40	.69
Publication status	-.007	-.013	184	-0.16	.87	-.021	-.037	-0.41	.68

Note: Analyses were conducted using Fisher's *Z*-transformed *r* values and mixed effects models (fixed predictor slopes and random intercepts). Summary statistics for the simultaneous regression analysis: R^2 = .104, 2-tailed *p* = .12; random effects variance component = .0165; mean effect size (*r*) = .269. *k* = number of samples in each analysis; *B* = unstandardized regression coefficient; *β* = standardized regression coefficient; *z* = critical ratio test for the regression coefficient; *p* = 2-tailed probability of *z*.

[a] See Table 2 for descriptions of the methodological and publication moderator variables.

Table 9. Tests of Weighted Regression Models for Methodological and Publication Moderators of Explicit–Criterion Correlations (ECCs)

	Weighted Regression Analyses of ECCs								
	Univariate effects					Simultaneous effects (k = 123)			
Moderators[a]	*B*	*β*	*k*	*z*	*p*	*B*	*β*	*z*	*p*
Number of effect sizes	-.011	-.141	156	-1.57	.12	-.013	-.173	-2.01	.04
Mean sample size	.001	.210	156	2.47	.01	.000	.147	1.67	.10
Criterion data collection method: subject response vs. observation	-.209	-.261	156	-3.05	.002	-.157	-.233	-2.64	.008
Predictor-criterion ordinal position	-.053	-.134	126	-1.56	.12	-.046	-.129	-1.35	.18
Predictor-criterion session relation	-.074	-.084	141	-0.83	.40	-.072	-.090	-1.00	.32
Publication year	.015	.095	156	1.09	.27	.006	.044	0.46	.65
Publication status	.087	.089	156	0.99	.32	-.060	-.075	-0.86	.39

Note: Analyses were conducted using Fisher's *Z*-transformed *r* values and mixed effects models (fixed predictor slopes and random intercepts). Summary statistics for the simultaneous regression analysis: R^2 = .110; 2-tailed p = .03; random effects variance component = .0624; mean effect size (r) = .341. k = number of samples in each analysis; B = unstandardized regression coefficient; β = standardized regression coefficient; z = critical ratio test for the regression coefficient; p = 2-tailed probability of z.

[a] See Table 2 for descriptions of the methodological and publication moderator variables.

Figure Captions

[*Note: Figures and captions appear on the following two pages, but the figures (without captions) are also provided separately in .tiff files*]

1. Weighted average IAT–criterion (ICC) and explicit–criterion (ECC) correlations for nine domains of criterion measures (see Table 3). Significance tests (*p* values) are from paired-sample, fixed-effects tests for difference in magnitudes of the two types of effect sizes. Numbers of samples (*k*) for significance tests are shown in Figure 2 (samples for which both effect sizes were available). However, this figure's plotted average effect sizes and 95% confidence interval error bars are based on all available samples, for which the numbers of samples are given in the axis labels (for ICCs first, ECCs second).

2. Weighted average partial IAT–criterion (IC.E) and explicit–criterion (EC.I) correlations (see text for further description) for nine domains of criterion measures (see Table 3). Significance tests (*p* values) are from paired-sample, fixed-effects tests for difference in magnitudes of the two types of effect sizes. Numbers of samples (*k*) are those for which both types of effect sizes were available. Error bars are 95% confidence intervals.

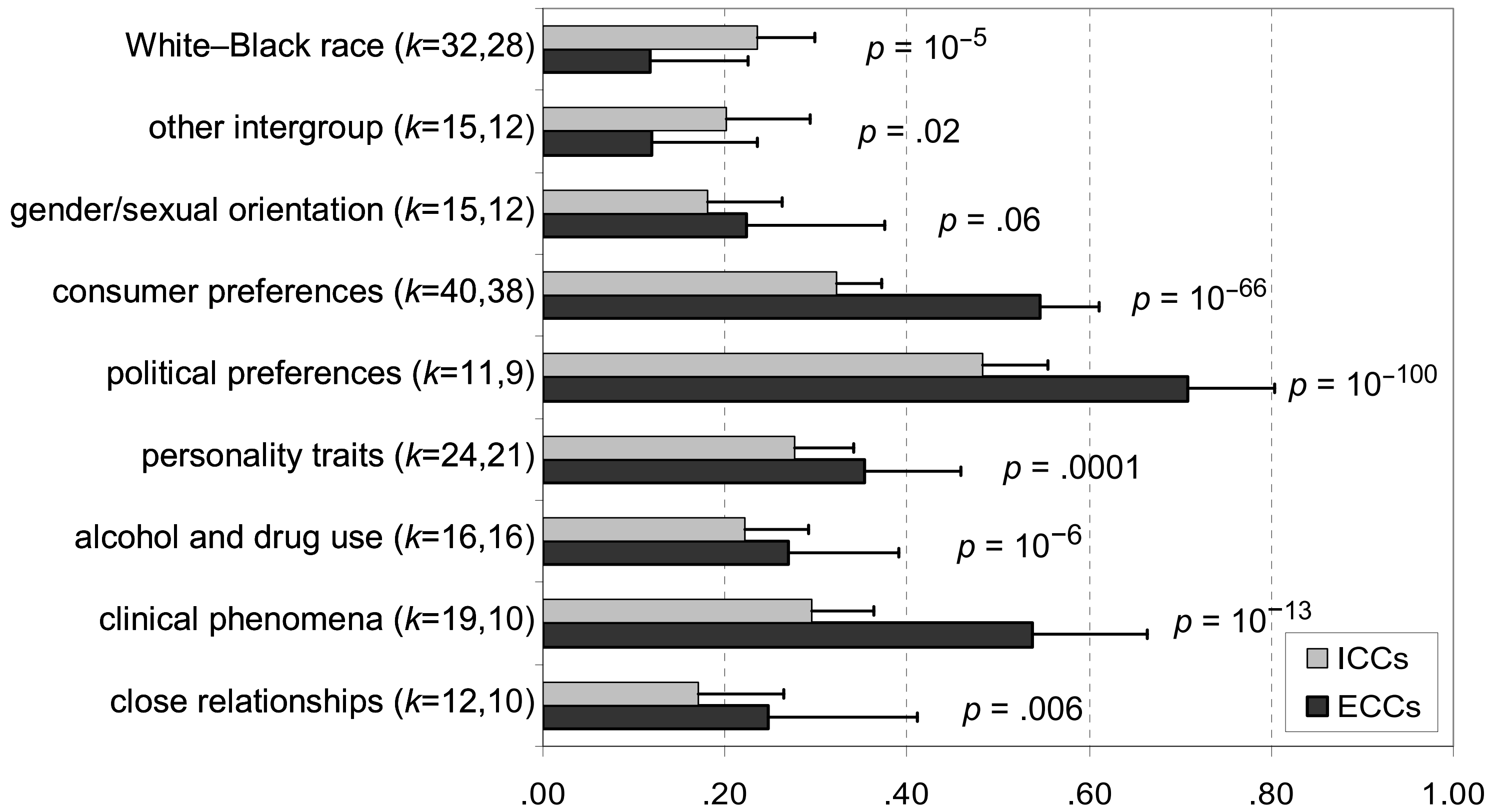

Figure 1. Weighted average IAT–criterion (ICC) and explicit–criterion (ECC) correlations for nine domains of criterion measures (see Table 3). Significance tests (*p* values) are from paired-sample, fixed-effects tests for difference in magnitudes of the two types of effect sizes. Numbers of samples (*k*) for significance tests are shown in Figure 2 (samples for which both effect sizes were available). However, this figure's plotted average effect sizes and 95% confidence interval error bars are based on all available samples, for which the numbers of samples (*k*) are given in the axis labels (for ICCs first, ECCs second).

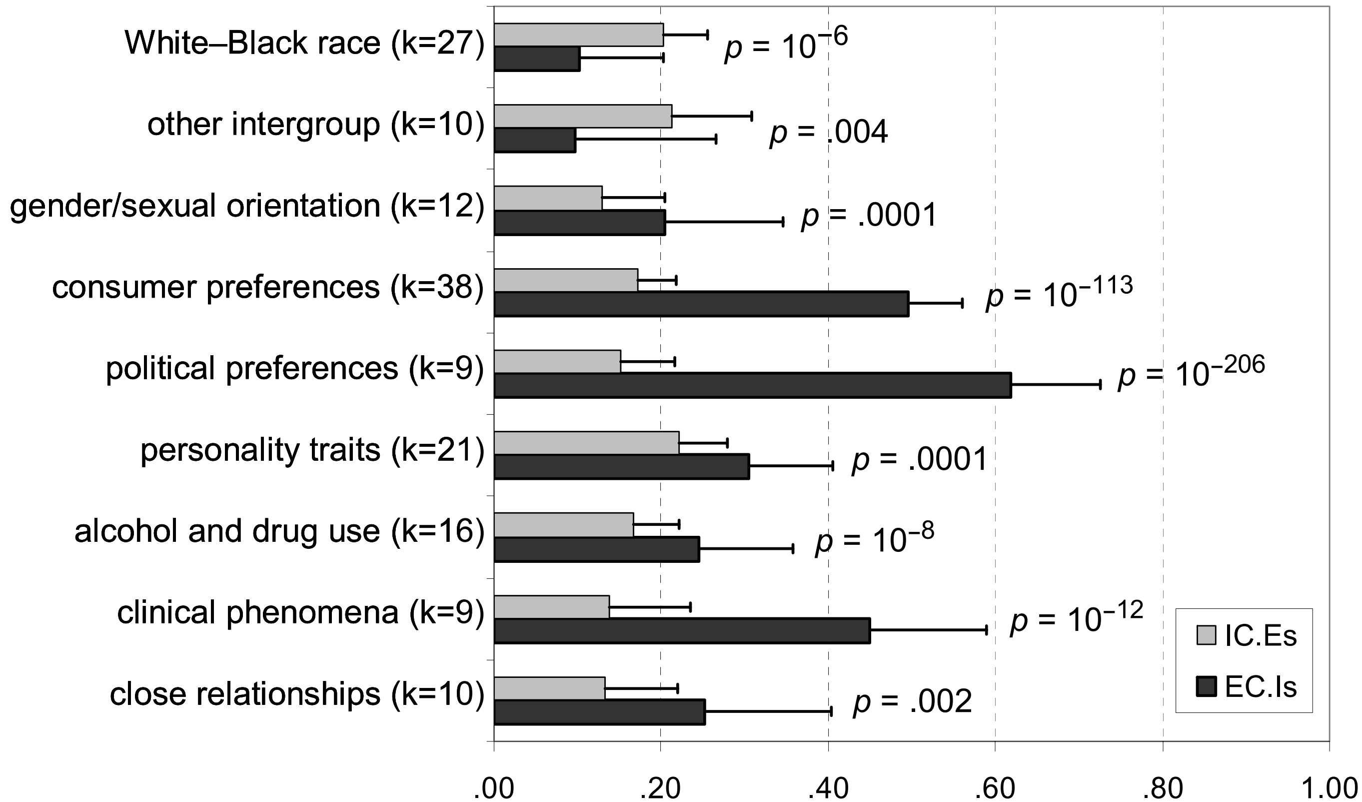

Figure 2. Weighted average partial IAT–criterion (IC.E) and explicit–criterion (EC.I) correlations (see text for further description) for nine domains of criterion measures (see Table 3). Significance tests (*p* values) are from paired-sample, fixed-effects tests for difference in magnitudes of the two types of effect sizes. Numbers of samples (*k*) are those for which both types of effect sizes were available. Error bars are 95% confidence intervals.

Footnotes

[1] The combined-task classifications present random selections of one of the concept pairs (e.g., the two sets of faces) on odd-numbered trials and random selections of the other pair (e.g., the pleasant and unpleasant words) on even-numbered trials. This alternation, or task-switching, has been found to produce measures of association strength (cf. Mierke & Klauer, 2001) that are superior to ones obtained with full randomization of the trial sequence.

[2] Averaging effect sizes within independent samples is statistically desirable, but can be conservative in estimating predictive validity. Consider Study 3 of Amodio and Devine (2006), which included (a) a race attitude IAT that was expected to predict voluntary selected seating distance from an African American and (b) a race stereotype IAT that was expected not to predict estimates of this seating distance measure. In the independent samples analysis, the predictive validity correlations of both IAT measures with the seating distance measure were averaged into the independent-sample ICC.

[3] The record of identification of effect sizes and coding of moderators is available in an archive that includes electronic copies of all of the studies included in the meta-analysis, as well as copies of correspondence with authors that led to obtaining the many effect sizes that were unavailable in original reports. Also in the archive are records of all data analyses reported in this article. This archive can be accessed for downloading at http://xxxxxxxxxxxxxxxxxx.

[4] In his meta-analysis of attitude–behavior relations, Kraus (1995) required (as an inclusion condition) that predictor and criterion measures be obtained in separate sessions. Such separate session designs were quite infrequent in the reports included in this meta-analysis. This observation may indicate a shift in research practices toward single-session studies in recent

years, but it may also indicate that researchers who work with IAT measures have been relatively unconcerned about within-session contamination between IAT measures and criterion measures.

[5] These analyses used SPSS macros described by Lipsey and Wilson (2001).

[6] For this regression equation, predicting Z_{ICC} from social sensitivity ratings, the unstandardized parameter estimates were 0.684 for intercept and −.084 for slope.

[7] Interestingly, Hofmann et al. (2005) reported no “evidence that correlations [of IAT with self-report measures] were influenced by the degree of social desirability . . . associated with the topic” (p. 1380). The disparity between their conclusion and the present one may be explained by the difference between their operational definition of social desirability and the present definition of social sensitivity. For their meta-analysis, Hofmann et al. assessed social desirability with a rating of “How much are people in general concerned about whether their attitudes or personality characteristics are socially acceptable” p. 1373; cf. Nosek, pp. 570–571). This may differ enough from the present operation (a rating of concern about the impression that the self-report response would make on others) to explain the difference in results.

[8] Kraus’s (1995) article has not recently been locatable in electronic library resources. A copy may be requested by email from the first author of this article.