Journal of Experimental Social Psychology 48 (2012) 1057–1068

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Social Psychology

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

Manipulating the role of cognitive control while taking the implicit association test

Eric F. Siegel [a], Michael R. Dougherty [a], David E. Huber [b,*]

[a] University of Maryland, College Park, MD, USA
[b] University of California, San Diego, CA, USA

ARTICLE INFO

Article history:
Received 20 January 2011
Revised 13 April 2012
Available online 26 April 2012

Keywords:
Implicit Association Test
Cognitive control
Working memory
Racial attitudes
Political attitudes

ABSTRACT

The Implicit Association Test (IAT) is one of the most widely used methods for measuring attitudes in the behavioral and social sciences. Recent studies have found that individual differences in cognitive control correlate with IAT scores. However, these studies did not collect independent measures of attitude, which makes it difficult to isolate the construct of attitude separate from cognitive control. Furthermore, no study has examined whether the role of cognitive control can be manipulated, which is necessary to establish a causal link between cognitive control and IAT performance. By collecting independent measures of attitude (explicit attitude ratings and the Affect Misattribution Procedure: AMP), Experiment 1 factored out the role of attitude for two different IATs and still found a relationship between IAT scores and cognitive control (Stroop and stop-signal). Experiments 2 and 3 manipulated the role of cognitive control through instructions and feedback regarding the race IAT's measurement goal. These manipulations increased average IAT scores (i.e., stronger preference for whites), increased the relationship with cognitive control (Stroop), and decreased the relationship with attitude (AMP). These results demonstrate that cognitive control influences IAT performance rather than merely correlating with IAT performance.

Introduction

Over the last decade, there has been a rapid increase in the use of implicit measures in psychological research, especially for the measurement of attitudes. The popularity of implicit measures is due in large part to the recognition that explicit measures are prone to a variety of response biases. This is especially true when the attitudes being assessed are socially unacceptable, such as racial prejudice. Responses to explicit measures (e.g., survey response scales) might be disingenuous, but responses to implicit measures are less susceptible to fakery because they measure automatic or reflexive responses elicited by stimuli that are congruent or incongruent with one's attitudes. Specifically, implicit measures assess attitudes by tapping associations between the relevant group or category (e.g., insects) and the affect associated with that group or category (e.g., insects are bad). These associations can be viewed either as implicit attitudes (Greenwald et al., 2002) or as the implicit component of a more general attitude construct (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Olson & Fazio, 2003). Because implicit measures are assumed to reflect automatic responses, they are often interpreted as unbiased assessments of underlying beliefs.

Our research focuses on the Implicit Association Test (IAT: Greenwald, McGhee, & Schwartz, 1998) because it is a popular measure of implicit attitudes used in hundreds of papers across a variety of disciplines (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). The IAT involves classification of four sets of stimuli, two based on their valence information (good/bad), and two based on category information (e.g., insects/flowers). However, the members of the category sets also have valence information (a mosquito is an insect but it is also bad). Just as in the Stroop task (Stroop, 1935), the IAT requires that the test taker classify a stimulus based on one feature (its group identity) while not responding based on another feature (its valence) that might indicate an incorrect response. Critically, while taking the IAT, test takers only use two response keys for these four categories. During one block of trials, the positive response is the same response key as one of the two groups (e.g., flowers, which is a compatible response mapping) while during another block of trials, the positive response is the same response key as the other group (e.g., insects, which is an incompatible response mapping). A comparison of reaction times during the compatible and incompatible blocks of trials (i.e., the IAT effect) yields a relative preference for one group compared to the other. To avoid making errors, the test taker must focus on the judgment that is along the relevant dimension for the current trial and avoid contradictory pre-potent responses to an incongruent piece of information. The test taker's attitudes towards the groups creates response interference; the stronger an association with a specific valence, the greater the response interference, and the stronger the IAT effect.

Since the IAT measures attitudes indirectly through response interference, it is possible that the test taker's natural ability to deal with that response interference influences IAT performance. Prior

* Corresponding author at: University of California, San Diego, Department of Psychology, 9500 Gilman Dr., La Jolla, CA 92093-0109, USA. Fax: +1 858 534 7190.
E-mail address: dhuber@ucsd.edu (D.E. Huber).

0022-1031/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2012.04.011

work has supported this theory. For instance, Blanton, Jaccard, Gonzales, and Christe (2006) used a factor analysis and came to the conclusion that "general processing speed" was a major source of variance underlying IAT scores. Mierke and Klauer (2001) analyzed the pattern of trials in the IAT, finding that responses were faster when the preceding trial was of the same judgment type (e.g., evaluative rather than group membership) rather than the opposite type. This indicated an important role for task switching, which is a key component of cognitive control. McFarland and Crouch (2002) found that different IATs assessing different attitudes tend to have a moderate correlation with each other, indicating that they share a common source of variance. This suggests that there is a shared method variance due to the common procedures used with all IATs (McFarland & Crouch, 2002; Mierke & Klauer, 2003).

To address these issues, Greenwald, Nosek, and Banaji (2003) developed the *D* scoring algorithm of IAT data to compensate for this shared method variance. However, despite their use of the *D* scoring algorithm, Klauer, Schmitz, Teige-Mocigemba, and Voss (2010) found that the method variance of the IAT still loaded onto a task switching component of cognitive control. Klauer et al. collected a variety of cognitive control measures to differentiate between an account of method variance based on task-switching versus an account based on response inhibition, which was predicted to be an important component of IAT performance according to the Quadruple process model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Sherman et al., 2008). Klauer et al. concluded that task switching provided a better explanation than response inhibition. However, from our perspective, the more important conclusion from these studies is that some aspect of cognitive control correlates with IAT performance and this is true even when using the *D* scoring algorithm.

Overview of experiments

The current research is closely related to the findings of Klauer et al. (2010), but addresses two key limitations of their study. By collecting independent measures of cognitive control, they measured the correlation between cognitive control and IAT performance. However, because they did not collect independent measures of attitude, it cannot be determined whether the correlation between cognitive control and IAT performance exists separate from the attitude construct. For instance, it is possible that people who have poor cognitive control actually hold stronger attitudes, which gives the false appearance that cognitive control directly influences IAT performance. Klauer et al. reported correlations between measures of cognitive control and a political attitudes IAT (Experiment 1) and a prejudice IAT (Experiment 2). Based on these correlations, it was assumed that poor cognitive control caused larger IAT scores. However, a recent study by Hodson and Busseri (2012) suggests an intriguing alternative, reporting evidence that individuals who have low cognitive ability early in life become prejudiced later in life and, furthermore, that political conservatism is a mediating factor. In other words, it may be that individuals who have low cognitive ability tend to produce IAT scores reflecting conservatism and prejudice because they are actually conservative and prejudiced. It is important to note that Klauer et al. included additional IAT tests (e.g., flowers versus insects and numbers/equations versus letters/words) , which also correlated with these behaviors. It is not clear whether this alternative account could explain the full set of correlations, but this is an empirical question, requiring additional measures of attitude besides the IAT. To separate the construct of attitude from the construct of cognitive control, we collected independent measures of cognitive control and independent measures of attitude. In doing so, we combined the methodology of Klauer et al. with the methodology of Nosek and Smyth (2007). Nosek and Smyth collected independent measures of attitude besides the IAT and used factor analyses to measure the role of attitude in the IAT separate from cognitive control. But again, as with the Klauer et al. study, it is possible that their results reflect an unfortunate correlation between strength of attitude and cognitive control, thus giving the false appearance that the IAT is a valid measure of attitude because their study only included independent measures of attitude. To properly measure the separate roles of attitude and cognitive control while taking the IAT, it is necessary to collect independent measures of *both* attitude and cognitive control, thus determining the latent correlation between attitude and cognitive control independent of IAT performance. This was the approach taken in Experiment 1, which allowed us to separate IAT performance into a factor based on cognitive control (as determined by independent measures of cognitive control) versus a factor based on attitude (as determined by independent measures of attitude).

Aside from the need to factor out any correlation between cognitive control and attitude, another limitation of the prior work is that cognitive control was studied as a correlated variable rather than a manipulated variable. Similar to Experiment 1, Experiments 2 and 3 collected independent measures of cognitive control and attitude, but, in addition, these experiments included a manipulation designed to vary the role of cognitive control while taking the IAT. To achieve this, one possibility is a reduction of available cognitive resources with a classic dual-task cognitive load manipulation. However, use of a second task while taking the IAT is likely to affect IAT performance in other ways. For instance, Schmitz, Teige-Mocigemba, Voss, and Klauer (2011) had people perform a random number generator task while taking the IAT. A reduction of cognitive resources should be problematic for incompatible IAT trials, and, as predicted, reaction times during the incompatible block increased under cognitive load. However, reaction time variability greatly increased under cognitive load, resulting in opposite conclusions depending on the method used to score the IAT. In light of this result, Experiments 2 and 3 used more subtle manipulations that do not directly change the IAT procedures.

The ability to accurately categorize a target item (e.g., determination that an individual is Caucasian) without interference from the valence task depends on the relative strengths of two opposing factors: The association strengths between valences and target items versus the ability to use cognitive control to counteract the influence of these associations. If there are no associations, or if the associations are not currently salient, there is no response interference during incompatible trials and cognitive control is not needed to accurately perform the IAT. Conversely, if an individual has a wealth of cognitive resources, even strong valence associations are easily countermanded. We hypothesized that our manipulations might vary the role of cognitive control in one of two ways corresponding to one of these two factors. In Experiment 2, this was achieved by making some test takers aware of the measurement goal of the IAT prior to taking a race IAT (this might increase the salience of the valence associations, and/or impose cognitive load due to anxiety) and in Experiment 3, this was achieved by giving test takers false feedback prior to taking the IAT regarding their racial attitudes as determined by a different measure (again, potentially increasing the salience of associations and/or inducing anxiety that reduces cognitive resources). Because these manipulations occurred prior to taking the IAT rather than during the IAT, they allowed manipulation of the role of cognitive control without altering the IAT procedures. Next, we consider the nature of these manipulations.

First, our manipulations might heighten the saliency of the valence associations attached to the category sets. There is mounting evidence that IAT performance varies when associations are primed or made salient in various ways (Dasgupta & Greenwald, 2001; Han, Czellar, Olson, & Fazio, 2010; Karpinski & Hilton, 2001). Increasing the salience of the associations attached to racial categories should result in greater response conflict, which should increase IAT scores and also increase the cognitive control component of IAT performance at the expense of the attitude component (i.e., the relationship between

IAT scores and independent measures of cognitive control should increase and the relationship between IAT scores and independent measures of attitude should decrease). Second, our manipulations might impose an additional cognitive load on the test taker if he or she is overly concerned about being labeled a racist as measured by the IAT. For instance, correlations between trait anxiety, performance on a response conflict task, and brain activation of prefrontal control regions have established a link between anxiety and cognitive control (Bishop, 2009). On this account, the anxiety caused by our manipulations should diminish one's ability to use cognitive control to ignore the task irrelevant valence associations of the target items, again resulting in greater response conflict. For individuals who naturally have more cognitive resources, this effect should be minimal, but for other individuals, this reduction should pose a serious problem, resulting in greater IAT scores. In summary, regardless of whether our manipulations increase the salience of associations or reduce cognitive resources, we predict that on average IAT scores should increase and IAT performance should more strongly reflect individual differences in cognitive control rather than individual differences in attitude.

Independent measures of cognitive control

Changes in cognitive control have been used to explain the development of cognitive abilities (Diamond & Gilbert, 1989; Ridderinkhof, vanderMolen, Band, & Bashore, 1997) and age-related declines in cognitive abilities (Hasher, Rypma, Stoltzfus, & Zacks, 1989). Moreover, an abundance of research illustrates that individual differences in cognitive control are related to individual differences in working memory span, reading comprehension, problem solving, general cognitive ability, and judgment and decision making (De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; Dempster & Corkill, 1999; Dougherty & Hunter, 2003; Friedman & Miyake, 2004; Gernsbacher, 1993; Kane & Engle, 2002). Cognitive control is not a single mechanism, and likely involves a number of sub-processes, such as 'shifting', 'updating', and 'inhibition' (Miyake et al., 2000). In the attempt to identify these sub-processes, Friedman and Miyake found that inhibition tasks are related to a factor specific to response inhibition as well as general cognitive capacity. We focus on these inhibition tasks because they contain many of the same conceptual and methodological elements as the IAT.

Two of the inhibition tasks identified by Friedman and Miyake (2004) were the color word Stroop task (see Macleod, 1991, for a review) and the Stop signal task (Logan, Cowan, & Davis, 1984). Like the IAT, both of these tasks involve the suppression of a prepotent response in the effort to respond accurately. In addition, both of these tasks involve mixtures of different trial types (i.e., congruent/incongruent or stop/go) where the nature of the next trial is unknown in advance. For incongruent trials of the Stroop task (the word RED in green font), participants must overcome response competition to respond accurately. In the stop signal task, participants are asked to identify shapes, but on a quarter of the trials they hear a tone that indicates they should refrain from responding. In this case, they must control their automatic tendency to make the standard response. Thus, these two measures of cognitive control use tasks that are similar to the IAT and Experiment 1 used them as independent measures of cognitive control.

Independent measures of attitude

To provide an independent measure of attitude, Experiment 1 collected explicit attitude ratings. However, many theories of attitude claim that implicit attitudes exist independent of explicit attitudes (e.g., Wilson, Lindsey, & Schooler, 2000). Therefore, to factor out the separate contribution from implicit attitude while taking the IAT, we needed a secondary measure of implicit attitude that was immune to cognitive control. The Affect Misattribution Procedure (AMP: Payne, Cheng, Govorun, & Stewart, 2005) relies on the test taker misattributing an affective reaction evoked by a prime stimulus (for example, a Caucasian face) to a subsequent neutral stimulus (e.g., a Chinese ideograph). The task of the test taker is to say whether they find the neutral stimulus to be pleasant or unpleasant; differences in these judgments as a function of prime stimulus are used to measure implicit attitudes. Payne et al. found that this measure is unaffected by attempts to correct for the influence of the prime stimulus. In their experiments, some participants were instructed that they should not let the prime influence their responses while others were told that the prime was merely a warning that the target would soon appear—AMP scores were nearly identical regardless of these instructions. Furthermore, AMP scores did not vary as a function of individual's self-reported motivation to control prejudice for a black/white AMP. Considering that even strategic efforts at control do not affect the AMP, these results suggest that the AMP is well suited to provide a secondary measure of implicit attitude that is uncontaminated by cognitive control. We do not imply that the AMP is a better measure of implicit attitude than the IAT; instead, we use the AMP because the AMP does not include task-switching or response competition and therefore does not suffer contamination from individual differences in cognitive control.

Experiment 1

The purpose of Experiment 1 was to determine the separate roles of cognitive control and attitude while taking the IAT by including independent measures of both these factors—previous work only included independent measures of one factor or the other and therefore the results of previous work may have reflected a latent correlation between cognitive control and attitude. Testing whether the role of cognitive control generalizes across different IATs, we examined two attitude domains: racial attitudes and political attitudes. We focused on these two attitude domains because the political IAT has one of the strongest correlations with explicit political measures whereas the race IAT has one of the lowest correlations with explicit measures of racial attitudes. This difference has been attributed to the influence of social desirability on explicit measures of racial prejudice; in contrast, explicit measures of political attitude are assumed to be unaffected by social desirability (Greenwald et al., 2009; Nosek & Smyth, 2007).

The primary hypotheses for Experiment 1 focus on the factor structure underlying the IAT. In particular, we hypothesized a three factor model, consisting of racial attitudes, political attitudes, and cognitive control. We hypothesized that the AMP and explicit measures would load onto common factors defined by their respective construct (political attitude and racial attitude). In contrast, we hypothesized that each IAT would load onto its respective attitude factor as well as a common cognitive control factor. We refer to this dual-factor model as the *cross-loading model*. The cross-loading model can be contrasted with an alternative model. The most obvious alternative model is a three-factor model where the IATs load only onto their target attitudes (henceforth called the *attitude-only model*). Because we hypothesized that the IAT measures both attitudes and cognitive control, we anticipated that the attitude-only model should yield a significantly worse fit to the data as compared to the cross-loading model.

Method

Participants

The participants were 213 University of Maryland undergraduate students. There were 128 females and 85 males. There were 27 self-identified Asians-Americans in the experiment, four participants indicated their race as "other", and the rest self-identified as Caucasian. Because they might have a different set of associations regarding

race, African-Americans were not included in this experiment. Participants received course credit for their participation.

Materials

Participants completed a battery of eight tasks measuring three core constructs: (1) Racial attitudes, (2) Political attitudes, and (3) Cognitive control. In addition, participants took Plant and Devine's (1998) motivation to control prejudice scale, although this was included for other reasons and was not analyzed.

Racial attitudes

Race IAT

Implicit attitudes were assessed using a black/white, good/bad IAT. The format and instructions were taken from the sample version of the IAT included with the DirectRT software package. Before each block of trials, participants were given onscreen instructions explaining the new button assignments. Throughout the IAT, they were instructed to respond as quickly and as accurately as possible. The IAT consisted of five blocks:

Block 1: participants classified adjectives according to valence (Bad or Good). There were 20 trials in this block.
Block 2: they classified faces according group membership (black or white). There were 20 trials in this block.
Block 3: was a mixed block with all four types of stimuli. Participants classified all four types of stimuli using only two answer keys. There were 60 trials in this block.
Blocks 4 and 5 were similar to Blocks 2 and 3 except that the answer keys assigned to the black and white categories were reversed.

The faces used in the IAT were those included in the DirectRT software package. There were six white faces and six black faces. The order of the blocks was counterbalanced such that half the participants were first given blocks in which black faces were identified by the left answer key, and the other half were first given blocks in which black faces were identified by the right answer key. The groups featured in each block appeared in the upper corners of the screen, and their left/right location corresponded to the button assignments for that block. Following incorrect responses, a red x appeared on the screen; the next trial was not initiated until mistakes were corrected.

The race AMP

The format and instructions of the AMP were taken from the original version of the AMP used by Payne et al. (2005). Participants were primed with pictures of black faces, white faces, or a gray box (control), and then asked to categorize an unfamiliar Chinese ideograph as "pleasant" or "unpleasant." There were 120 ideographs sampled randomly without replacement. The faces used as primes were the same faces used in the race IAT. The AMP measure was calculated from 72 trials, evenly split between the 3 types of primes. Payne et al. found that this measure is similar regardless of whether the primes are displayed supraliminally or subliminally. In our experiment, the primes were displayed supraliminally (for 200 milliseconds).

Modern Racism Scale

Explicit self-reported attitudes towards race were measured with the Modern Racism Scale (MRS: McConahay, 1986). The MRS is designed to measure racial attitudes through questions on issues that are tied to race relations within the United States. The MRS contains seven items measured on a 7-point Likert scale, such as "Discrimination against blacks is no longer a problem in the United States."

Political attitudes

Political IAT

The political IAT was identical in format to the race IAT. However, instead of showing black and white faces, participants were shown pictures of democrat and republican presidential candidates and were asked to categorize them based on their party affiliation. Pictures of two democratic candidates (Hillary Clinton and Barack Obama) and two republican candidates (John McCain and Rudy Giuliani) were used in the IAT. There were four pictures of each candidate. These pictures were taken from the political IAT featured on the Project Implicit IAT website (implicit.harvard.edu, retrieved October 2008).

Political AMP

The political AMP was identical in format to the race AMP, but used the pictures of the democratic and republican candidates rather than the pictures of black and white faces. The pictures of the candidates used for this measure were the same as the pictures used in the political IAT.

Political attitude measure

The explicit measure of political attitude was a questionnaire designed by the experimenters to test for political attitudes through questions on a number of political issues. Examples of questions include "I support a timetable for the withdrawal of troops from Iraq" and "I support the construction of a fence along the US/Mexico border." There were 14 questions scored on a six point Likert scale. A complete list of the questions is included in the Appendix A.

Cognitive control measures

Stroop task

The Stroop task included 100 trials during which participants identified a series of color words (red, green, blue) that were displayed in a congruent or incongruent colored font. Participants were required to respond based on the words' font color while ignoring the words' meaning. A Stroop task score was calculated by comparing the average time to correctly name congruent color words from the average time to correctly name incongruent color words. The Stroop task was a computer-based version with 80% congruent trials on which the words' font and meaning matched.

Stop signal task

The stop-signal task (Logan et al., 1984) was run using the STOP-IT executable for windows and was analyzed with the ANALYZE-IT program (Verbruggen, Logan, & Stevens, 2008). The end result of this task and analysis is the Stop Signal Reaction Time (SSRT), which is an estimate in milliseconds of how quickly an individual can stop a pre-potent response. Therefore, a higher score indicated worse cognitive control. Details of the procedure and analysis are reported by Verbruggen et al. and are briefly summarized here. On each trial participants saw either a square or a circle and identified the shape as quickly as possible without making errors. However, on a randomly selected 25% of trials (i.e., stop trials) there was an auditory tone indicating that the participant should withhold their response. This tone could occur at any of a number of delays after onset of the square or circle. This delay was systematically varied using a staircase procedure to identify the delay at which the participant was able to withhold their response 50% of the time. The SSRT was then calculated by subtracting the mean stop-signal delay from the mean reaction time for go trials (i.e. trials that did not present a tone). These SSRT values were used for all analyses.

Procedure

The entire experiment was run on Windows computers, using DirectRT software (www.empirisoft.com), which allows for millisecond timing accuracy. Participants responded using a standard keyboard. Participants completed each task and the race and political IATs were completed twice to test the reliability of the IAT. They were offered a 5 minute break half way through the experiment, but none of the participants opted to take a break. The order in which the tasks were presented to participants was randomized for each participant, with the constraint that participants could not take the race IAT twice in a row. The entire experiment took approximately 45 minutes. After completing the experiment, participants were debriefed about its purpose and were excused.

Results and discussion

Data reduction

We set the direction of the IAT scores such that test takers who preferred whites more than blacks on the race IAT had positive IAT scores and test takers who preferred Democrats over Republicans had positive IAT scores on the political IAT. The IAT scores were transformed using the D_1 transformation recommended by Greenwald et al. (2003), which takes the difference in average reaction time between the mixed blocks and divides this difference by the standard deviation pooled across these blocks[1]. Each IAT was taken twice in the experiment and the correlation between each instance of the same IAT was high, indicating good reliability for both the race ($r = .62$) and political ($r = .81$) IATs. Consequently, the D_1 scores for the two race IATs were averaged to create a single score, as were the D_1 scores for the two political IATs. Responses to the explicit scale measures were also set such that higher values indicated a preference for whites or Democrats, which necessitated a reverse coding of the questions in some cases (so that high values were replaced with low values), depending on the content of the question. These values were then averaged for each individual. The Cronbach alphas were $\alpha = .69$ and $\alpha = .77$ for the political attitude scale and the MRS, respectively, indicating that it was acceptable to combine items into a single score for each scale. The AMP scores were calculated by tallying the number of ideographs judged to be pleasant for each condition that presented a different type of prime stimulus prior to the ideograph. Then, the totals from each prime condition were subtracted. As with the other measures, the direction of subtraction was set such that positive numbers indicated a relative preference for whites or Democrats.

Directionality of the measures

Table 1 shows the descriptive statistics of each measure using the calculations described above. Before comparing these measures using correlation and factor analyses, we note that the cognitive control measures are unidirectional (scores are positive, with smaller values indicating better cognitive control and larger values indicating worse cognitive control) whereas the attitude measures are bidirectional (larger versus smaller values indicate a relative preference for one group or the other and intermediate values indicates an equal preference). The instructions for the IAT were to respond quickly and accurately for both the compatible and incompatible trial blocks. However, the labeling of 'compatible' and 'incompatible' is arbitrary, and only serves to determine the orientation of the bidirectional scale. In the current situation we were not interested in how cognitive control relates to the direction of one's preferences, but rather how cognitive control relates to the ability to reduce interference arising from one's attitude towards the target items, regardless of which group is preferred. Therefore, we transformed all of the attitude measures into a unidirectional scale prior to performing the correlation and factor analyses. For the IAT and AMP this was done by taking the absolute magnitude[2]. For the MRS and explicit measure of political attitudes, this was done by subtracting the neutral midpoint of the scale (4 for the MRS and 3.5 for the political attitudes measure) and then taking the absolute value. After this transformation, zero indicated an equal preference for both groups and positive values indicated stronger preferences for one group over the other, regardless of which group was preferred.

Table 1
Experiment 1 descriptive statistics.

Measure	Mean	SD	Skew	Kurtosis
Race				
IAT	170.62	103.00	.36	−.53
AMP	5.58	5.17	.49	.89
MRS	2.80	.81	.41	.3
Political				
IAT	157.62	94.20	.48	.28
AMP	7.34	6.81	−.26	.07
PAS[a]	4.27	.72	.01	.17
Cognitive Control				
Stroop	135.463	97.32	−.03	−.47
Stop Signal	297.94	75.96	.38	.38

[a] PAS refers to the explicit political attitudes scale.

Stroop scores and outliers

For the descriptive statistics shown in Table 1, Stroop scores were calculated in the traditional manner, subtracting the average correct reaction time for congruent trials from the average correct reaction time for incongruent trials—larger Stroop scores indicated worse cognitive control. However, reaction times are positively skewed, which can result in outliers. Therefore, for the correlation analyses and factor analyses reported below, the Stroop scores where transformed in the following manner. First, the inverse reaction time for each correct trial was calculated. Second, the average inverse reaction time was found separately for the congruent and incongruent conditions. Third, the average for the incongruent condition was subtracted from the average for the congruent condition so that the Stroop scores were still on a scale with positive numbers representing worse cognitive control. An outlier analysis was done on these transformed Stroop scores, as well as all the other variables entered into the correlation and factor analyses. Using the H-spread to calculate the 'inner fence' (i.e., 1.5 times the interquartile range above the 75% or lower than the 25% score; aka 'box and whiskers'), there were no outliers for any of the variables.

Correlation analyses

The zero-order correlations are presented in Table 2 using attitude measures that were transformed to be unidirectional. Several results are of particular note. First, the political IAT, political AMP, and the explicit measure of political attitude are all correlated with one another. Second, the race AMP, but not the race IAT, is significantly correlated with the explicit measure of racial attitude. These first two findings are consistent with prior research. The political IAT tends to have a strong correlation to explicit measures of political attitude, whereas

[1] See Greenwald et al. (2003) for a full description of the D_1 transformation algorithm. Our version of the IAT included a built in penalty for error trials, so it was not necessary to transform the error trials.

[2] A potential limitation when using the absolute magnitude of the IAT and AMP scores is that some of the attitude measure variance is lost because the theoretical range of these measures is cut in half. In practice this was not the case because most people exhibit a preference for Caucasians and because most undergraduate students are liberal-minded. For the political IAT and AMP of Experiment 1, 19% (IAT) and 23% (AMP) of the participants had negative scores (more conservative). For the race IAT, the proportions (IAT/AMP) favoring African-Americans were even smaller: Experiment 1, 10%/10%; Experiment 2, 3%/6%; and Experiment 3, 6%/6%.

Table 2
Experiment 1 correlations.

	Political AMP	Political IAT D_1	Explicit political	Race AMP	Race IAT D_1	Explicit race	Stroop effect	Stop signal
Political AMP	1.00	.39**	.23*	.12	.00	−.03	.00	.10
Political IAT D_1	–	1.00	.29*	−.05	.14*	−.04	.26*	.08
Explicit Political	–	–	1.00	.04	.06	−.06	.1	.1
Race AMP	–	–	–	1.00	.25**	.28**	−.014	.056
Race IAT D_1	–	–	–	–	1.00	.18	.45**	.20**
Explicit Race	–	–	–	–	–	1.00	.08	−.01
Stroop Effect	–	–	–	–	–	–	1.00	.30**
Stop Signal	–	–	–	–	–	–	–	1.00

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

the race IAT does not correlate strongly with explicit measures of racial attitudes (Nosek & Smyth, 2007). The political IAT results suggest that the IAT is capturing some aspect of the target attitudes. However, similar to the conclusions drawn by Klauer et al. (2010), our results also suggest that the IAT reflects individual differences in cognitive control; both the political IAT and the race IAT were significantly correlated with performance on the Stroop task and race IAT correlated with the stop-signal task. This pattern of correlations suggests that IAT scores involve at least two factors. We test for this two-factor structure next.

Confirmatory factor analyses

The goal of the study was to determine the separate roles that attitudes and cognitive control play in determining performance on the IAT while controlling for shared variance between the latent constructs of cognitive control and attitude. This question was addressed by constructing two confirmatory factor analysis models that contain three factors: A political attitude factor, a racial attitude factor, and a cognitive control factor. The attitude-only model assumes that the IAT only reflects attitude, with the IATs loaded onto their respective attitude. In the cross-loading model, both IATs are cross-loaded onto both their respective attitudes and onto the cognitive control factor. Unlike prior studies, this experiment included independent measures of attitude and independent measures of cognitive control. This allowed us to estimate the correlations between the three factors at the same time that we determined how each factor loaded onto IAT performance. As shown by the double-headed arrows in Figs. 1 and 2, the correlation between racial attitudes and cognitive control and the correlation between racial attitudes and political attitudes were estimated; the correlation between cognitive control and political attitudes is simply the product of these two values[3].

We used the AMOS program (Arbuckle, 1999) to perform maximum likelihood estimation based on the covariance matrix. The unidirectional versions of the attitude measures were used (i.e., preference strength, regardless of which group was preferred). To examine if one model was significantly better than another, we performed chi-square difference tests on nested models (e.g., Batchelder & Riefer, 1990; Jang, Wixted, & Huber, 2009). These tests entailed subtracting the chi-square for the full model from the chi- square for a nested, restricted model with fewer free parameters (degrees of freedom were calculated with an

[3] An initial exploratory factor analysis failed to find a correlation between cognitive control and political attitudes and confirmatory factor analyses that included this correlation as a free parameter produced results that were nearly identical to the reported analyses that did not include this correlation.

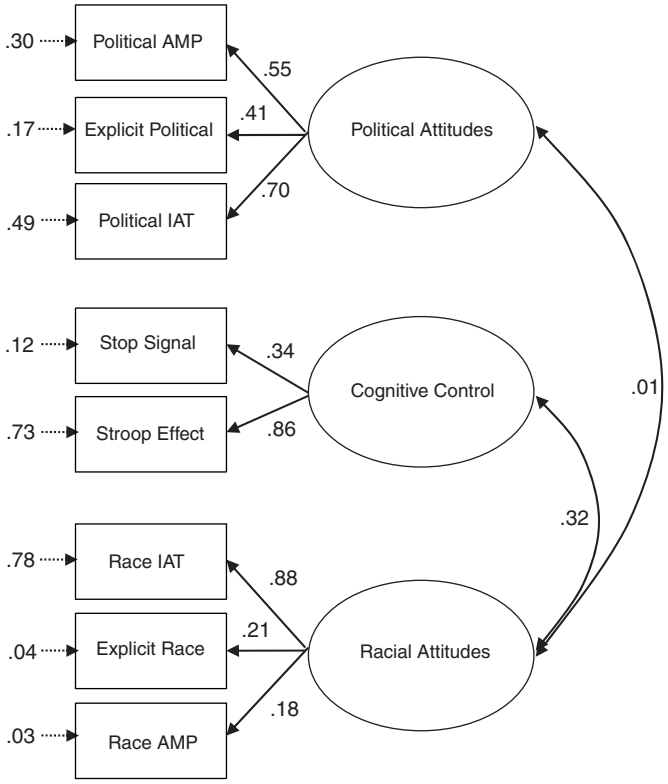

Fig. 1. Path diagram for the attitude only model applied to Experiment 1. Values to the left of the measures at the ends of the dotted straight lines are error variances.

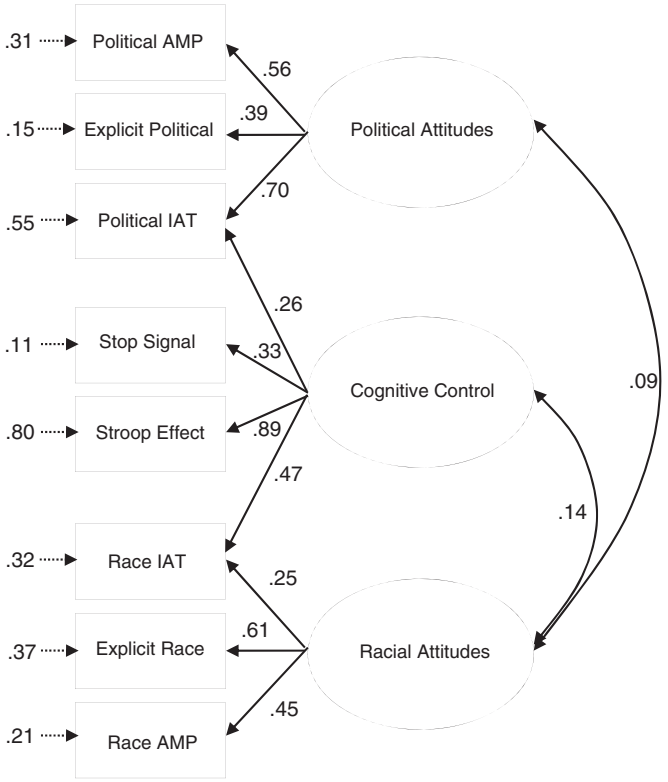

Fig. 2. Path diagram for the cross-loading model applied to Experiment 1. Values to the left of the measures at the ends of the dotted straight lines are error variances.

analogous subtraction). If the resulting chi-square difference is significant, then the less restricted model provides a significantly better fit. Additional model fit indices included the Akaike Information Criterion (AIC: Akaike, 1973), Bentler's comparative fit index (CFI), and the root mean square error of approximation (RMSEA: Hu & Bentler, 1998). Lower values of the AIC and RMSEA, and higher values of CFI indicate better fit. Values of RMSEA less than .05 (Browne & Cudeck, 1993), and values of CFI greater than .95 (Hu & Bentler, 1998) are considered good fits.

Model 1: the attitude-only model

Fig. 1 presents the path diagram representing the attitude-only model. The numbers next to the straight solid line arrows are the standardized factor loadings (interpretable as standardized regression coefficients). The numbers on the left, at the ends of the straight dotted line arrows, are the error variances for each task and represent the variance attributable to idiosyncratic task requirements and measurement error (i.e., unexplained variance). The numbers next to the curved double-headed arrows are the correlations between the latent variables. This model provided a poor fit with a significant chi-square, $\chi^2(18, N=213)=70.00$, $p<.01$; an RMSEA greater than .05 (RMSEA = .11); a CFI value less than .95 (CFI = .65); and an AIC value of 119.96. Assuming that IAT scores are systematic and reliable, and assuming that our attitude measures are fairly accurate and reliable, this result suggests that there must be something more to IAT scores than attitude.

Model 2: the cross-loading model

We constructed a second model in which the IATs load onto their respective attitudes as well as cognitive control, as seen in Fig. 2. The fit of this model was good, with a non-significant chi-square, χ^2 (16, $N=213$) = 16.62, $p=.41$; an RMSEA less than .05 (RMSEA = .014); a CFI value greater than .95 (CFI = .99); and an AIC value of 72.62. The cross-loading model was directly compared to the attitude-only model, revealing that the extra parameters of the cross-loading model were justified by providing a significantly better fit ($\chi^2_{diff}(2)=53.38$, $p<.01$). In addition, all of the eight tasks loaded significantly onto their respective factors. As seen in Fig. 2, the political IAT had a stronger loading onto political attitudes than cognitive control whereas the race IAT loaded more strongly onto cognitive control than racial attitude.

These results support the hypothesis that a significant portion of the variance in the IAT is attributed to attitude-free cognitive control. Previously work assumed that cognitive control is independent of attitude, and thus measured the contribution of cognitive control without including independent measures of attitude. In contrast, the current analyses were able to separately measure attitude and cognitive control to determine whether this assumption was valid. These analyses indicate that the IAT is (at least) a dual factor measure that reflects a combination of attitude and cognitive control. However, there are two caveats to this interpretation of Experiment 1. First, the attitude measures were category rating measures whereas the IAT and the measures of cognitive control were reaction time measures. Thus, the common method variance between the IAT and the cognitive control measures may be some aspect of responding quickly other than cognitive control. Second, these results are correlational rather than experimental. Both of these issues are addressed in Experiments 2 and 3; Experiments 2 and 3 demonstrate that the correlation between cognitive control and IAT performance can be changed by experimental manipulations that are unrelated to a generalized ability to respond quickly.

Experiment 2

The findings of Experiment 1 provide support for the hypothesis that attitude alone is insufficient to explain individual differences in IAT performance; the inclusion of cognitive control provided a significantly improved account of the data. However, the results of Experiment 1 are correlational. To determine whether there is a causal relationship between cognitive control and IAT performance, Experiments 2 and 3 attempted to manipulate the role of cognitive control. Similar to Experiment 1, this was done while collecting an independent measure of attitude (the AMP) and an independent measure of cognitive control (Stroop) so that the correlation between cognitive control and attitude could be measured independent of IAT performance. In all conditions, participants performed the Stroop task followed by the race AMP. At this point, we manipulated participants' beliefs regarding the experiment: half of the participants were told that we were interested in measuring categorization ability (the low knowledge condition) and half were told that we were interested in measuring racial attitudes (the high knowledge condition). Next, all participants completed the race IAT. By making participants aware of the measurement goal of the race IAT, we sought to increase the role of cognitive control. We hypothesized that this manipulation might either increase the salience of the positive or negative associations for different racial groups and/or this manipulation might impose a cognitive load, reducing the availability of cognitive resources necessary for cognitive control. Either of these effects should increase response interference while taking the IAT. If the role of cognitive control increases (i.e., the extent to which IAT performance reflects individual differences in cognitive control), we predict: 1) increased IAT scores; 2) an increased relationship between the IAT and Stroop performance; and 3) a decreased relationship between the IAT and the AMP.

While taking the IAT, cognitive control is needed to switch between valence trials and group membership trials (Mierke & Klauer, 2001) and cognitive control is also needed to inhibit contradictory responses arising from the valence associated with members of a group (Conrey et al., 2005; Sherman et al., 2008). Both these aspects of cognitive control are expected to vary to the extent that target items from the categorization set activate a valence response. If there is no valence associated with group members, then there is no need to inhibit a contradictory response, and, also, task switching is less relevant (i.e., using both response mappings simultaneously will not lead to errors). Thus, the stronger the test taker's valence associations, the stronger the interference on the IAT and the greater the IAT effect. Based on prior evidence that the IAT is malleable and varies with priming and framing effects (Han et al., 2010; Karpinski & Hilton, 2001; Olson & Fazio, 2003), we hypothesized that knowledge of the IAT's purpose might prime the valences associated with racial groups in the high knowledge condition. Alternatively, these associations might become more salient through the ironic process effect (Wegner, Schneider, Carter, & White, 1987) in which the attempt to suppress a particular thought paradoxically results in greater elicitation of that thought. Whether it occurs through priming or the failed attempt to suppress, greater salience of race valence associations should elicit greater response interference during the mixed blocks of the IAT.

Besides increasing the salience of racial valence associations, our knowledge manipulation might give participants performance anxiety, similar to what occurs with stereotype threat (Steele & Aronson, 1995). In the current situation, this anxiety would not be due to a particular stereotype of the participant, but rather the fear held by most individuals that they might be publically labeled a racist. The mechanisms behind stereotype threat are still under debate although one account supposes that performance deficits arise from decreased executive functioning (Schmader, 2010). Similarly, it may be that our knowledge manipulation induces a cognitive load (see Richeson & Shelton, 2003 for evidence that interracial interaction produces a cognitive load), which reduces the availability of cognitive resources while taking the IAT. If cognitive control is reduced in this manner, this should increase response interference during the mixed blocks of the IAT.

Whether through increased salience or decreased cognitive resources, individual with poor cognitive control should experience

greater response interference while taking the IAT in the high knowledge condition. Therefore, on average, IAT scores should increase. Furthermore, this should enhance the role of cognitive control while taking the IAT, producing an increased relationship with Stroop scores. If IAT scores more strongly reflect individual differences in cognitive control, it follows that the role of attitude while taking the IAT should be reduced, producing a decreased relationship with AMP scores.

Method

Participants

Participants were 102 psychology students enrolled in psychology courses at the University of Maryland. There were 61 females and 41 males. Twelve of the participants self-identified as Asian-American, and the rest self-identified as Caucasian. Because they might have a different set of associations regarding race, African-Americans were not included in this experiment. Participants received partial course credit for their participation.

Procedure

The experiment used a between subjects design, with 53 participants randomly assigned to the high knowledge condition and 49 to the low knowledge condition. Participants in both groups were administered the Stroop task followed by the race AMP. The knowledge manipulation was achieved by providing participants with different descriptions regarding what the AMP and IAT measured. In the low knowledge condition, participants were told "The purpose of the first task was to test your ability to make quick and accurate categorizations. You will now complete another measure of categorization ability to see if you perform in a similar manner on that measure." Participants in the high knowledge condition received the following task description:

The first judgment task you completed incorporated black and white faces in the procedure. This is because one purpose of this type of judgment task is to measure racial bias. You will now take another task that is meant to test for unconscious thoughts and feelings, and will specifically test for racial preferences. We are trying to see if people perform on the two tests in a similar way. You should be aware you may not agree with the interpretation of your performance of this test.

After receiving this description, participants took the race version of the IAT.

Results and discussion

Data reduction was performed in the same manner as Experiment 1. As with Experiment 1, the descriptive statistics used the standard bidirectional form of the attitude measures and the Stroop scores were calculated in the traditional manner. These are reported in Table 3, collapsed across both knowledge conditions. The regression and correlation analyses used the inverse transformed Stroop scores as well as the absolute magnitude of the IAT D_1 and AMP scores. Outlier analyses failed to find any outliers for any of the variables. As predicted, absolute IAT D_1 scores were higher in the high knowledge condition ($M = .67$, $SD = .33$) than the low knowledge condition ($M = .49$, $SD = .29$), $t(100) = -2.88$, $p < .01$.

Bivariate correlations are reported in Tables 4 and 5 for the low and high knowledge conditions, respectively, and scatterplots of these correlations are shown in Figs. 3 and 4. Because this experiment only included one independent measure of racial attitude (the AMP) and one independent measure of cognitive control (Stroop), we did not perform factor analyses. Instead, the key hypotheses were tested with multiple regressions. First, we ran a regression model with Stroop, AMP, a condition by Stroop interaction term, and a condition by AMP interaction term, as predictors of the IAT D_1 scores. We used the interaction terms to determine whether the Stroop or AMP regression coefficients significantly changed as a function condition (Aiken & West, 1991). This test was followed by separate multiple regressions for each condition with Stroop and AMP scores as predictors of the IAT scores. All regressions used mean centered predictor variables to reduce errors arising from multicollinearity. The logic behind these tests was similar to the rationale behind Experiment 1—by using both Stroop and AMP scores as predictors, we hoped to isolate the separate roles that attitude and cognitive control play while taking the IAT, aside from any correlation between cognitive control and attitude (i.e., aside from any correlation between Stroop and AMP scores).

Table 3
Experiment 2 descriptive statistics.

Measure	Mean	SD	Skew	Kurtosis
IAT	183.19	142.66	0.35	−0.34
IAT D_1	0.58	0.32	−0.31	−0.61
AMP	6.20	7.43	0.05	0.72
Stroop	113.43	118.82	0.69	0.55

Table 4
Experiment 2 correlations for the low knowledge condition.

	IAT D_1	AMP	Stroop
IAT D_1	1	.42**	−.14
AMP		1	−.31*
Stroop			1

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

Table 5
Experiment 2 correlations for the high knowledge condition.

	IAT D_1	AMP	Stroop
IAT D_1	1	.39*	.48**
AMP		1	.43**
Stroop			1

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

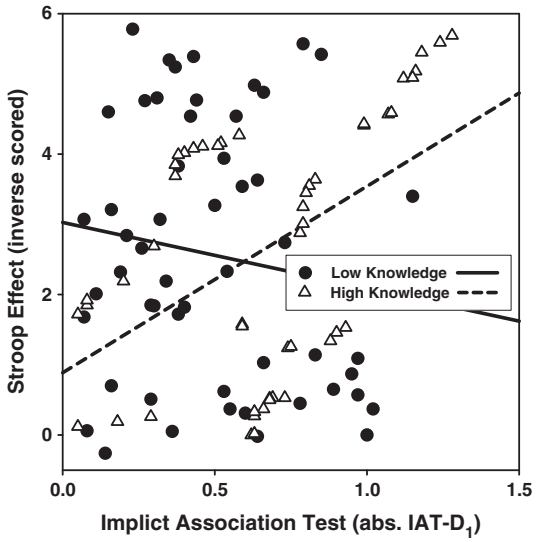

Fig. 3. Scatter plot of the correlations between the race Implicit Association Test (IAT) and the Stroop effect for Experiment 2. Representing implicit attitude strength regardless of preference, the absolute magnitude of IAT scores are shown, with the IAT calculated using the D_1 scoring algorithm. To reduce outliers, the Stroop effect vales were calculated by subtracting the average inverse reaction time of correct incongruent trials from the average inverse reaction time of correct congruent trials. For the high and low knowledge conditions, separate correlations were calculated as indicated by the regression lines (see Tables 4 and 5 for the corresponding correlation coefficients).

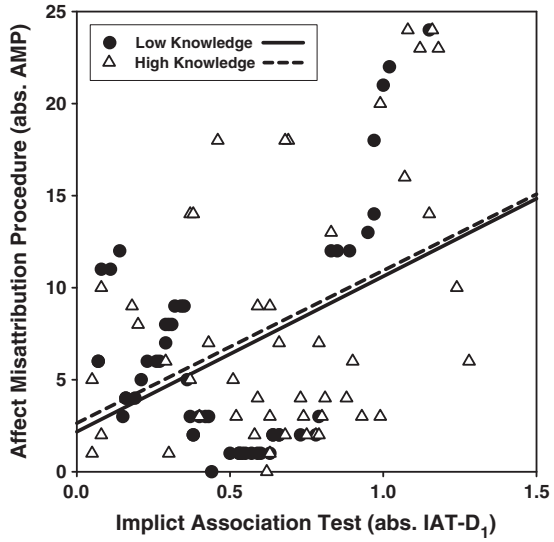

Fig. 4. Scatter plot of the correlations between the race Implicit Association Test (IAT) and the Affect Misattribution Procedure (AMP) scores for Experiment 2. Representing implicit attitude strength regardless of preference, the absolute magnitude of AMP and IAT scores are shown, with the IAT calculated using the D_1 scoring algorithm. For the high and low knowledge conditions, separate correlations were calculated as indicated by the regression lines (see Tables 4 and 5 for the corresponding correlation coefficients).

Because there was a significant interaction between Stroop and condition, $\beta = .28$, $t(96) = 2.20$, $p < .05$, we next checked whether the Stroop scores were a good predictor separately for each condition. This was done using a separate regression for each condition with Stroop and AMP scores as predictors of the IAT: Stroop scores were a good predictor of IAT scores for participants in the high knowledge condition, $\beta = .38$, $t(46) = 2.74$, $p < .001$, but not in the low knowledge condition, $\beta = -.02$, $t(50) = -.125$, $p = .90$. Thus, as predicted, the role of cognitive control was greater in the high knowledge condition. If the role of cognitive control is greater, it follows that the role of attitude should be diminished, which implies that there should be a stronger relationship between IAT and AMP scores in the low knowledge condition compared to the high knowledge condition. While the AMP by condition interaction was a non-significant predictor of IAT performance, $\beta = .13$, $t(96) = 1.03$, $p = .31$, inspection of the regression parameters in each condition was consistent with this prediction: the AMP predicted IAT scores in the low knowledge condition, $\beta = .41$, $t(50) = 3.04$, $p < .01$, but not the high knowledge condition, $\beta = .23$, $t(46) = 1.61$, $p = .11$. Interpreting these regressions cautiously (due to the non-significant interaction), this suggests that knowledge may somewhat compromise the validity of the IAT as a measure of attitude.

In summary, Experiment 2 successfully manipulated the role of cognitive control while taking the IAT through increased awareness for the IAT's measurement goal—through increased salience of racial attitudes or through increased cognitive load, we hypothesized that the role of cognitive control should increase in the high knowledge cognition. It followed that: 1) average IAT scores should increase because individuals with poor cognitive control should experience greater response interference; 2) the relationship between the IAT and measures of cognitive control should increase; and 3) the relationship between the IAT and measures of attitude should decrease. The first two predictions were confirmed, and the results were in the expected direction for the third prediction.

Experiment 3

Experiment 2 manipulated awareness of the IAT's measurement goal, finding that the role of cognitive control varied in the expected manner. However, the effect sizes in Experiment 2 were modest and one of the three key predictions failed to produce a significant interaction—the prediction regarding the correlation between the AMP and the IAT was only confirmed with separate tests that did not correct for multiple comparisons. In Experiment 3 we used a slightly different manipulation in an attempt to manipulate the role of cognitive control more strongly in hope of producing a more powerful test of the prediction that the attitude component of the IAT should reduce as the cognitive control component of the IAT increased.

One weakness of the manipulation used in Experiment 2 is that the IAT is widely known among undergraduate psychology majors. Thus, it is possible that participants in the low knowledge group suspected the true intent of the IAT despite reading a description to the contrary. To avoid this concern, we used a salience manipulation based on feedback rather than knowledge. This allowed us to manipulate the role of cognitive control even though participants in both groups were fully informed regarding the IAT's measurement goal. Before taking the IAT, but after taking the AMP, some participants were told that they had been identified either as scoring high (indicating a tendency toward racial prejudice) or low (indicating no such tendency) based on the AMP. For the high feedback group, we anticipated that this would evoke concern regarding evaluation of racial attitudes. Participants with a greater concern about being prejudiced might experience: 1) greater reflection on racial attitudes, producing more salient associations; 2) an increased attempt to suppress thoughts related to race, producing exactly the opposite through the ironic process effect; and/or 3) greater anxiety, producing a cognitive load that reduced cognitive resources. Any or all of these should increase the role of individual differences in cognitive control while taking the IAT.

Method

Participants

The participants were 98 students enrolled in psychology courses at the University of Maryland. There were 58 females and 40 males. Eight of the participants self-identified as Asian-American and the rest self-identified as Caucasian. Because they might have a different set of associations regarding race, African-Americans were not included in this experiment. Participants received credit towards their course for completing the experiment.

Procedure

As in Experiment 2, participants completed the Stroop task and the AMP prior to the feedback manipulation. Following these two tasks, participants were provided with false-feedback regarding their performance on the AMP. Participants in the high feedback group ($N = 46$) were told their scores on the initial test of racism suggested that they had high levels of racial prejudice. Participants in the low feedback group ($N = 52$) were told that the initial test suggested that they had low levels of racial prejudice. Participants were then told that they would receive a second measure of racial bias as a means of validating the first measure.

Results and discussion

The results of Experiment 3 were analyzed in the same manner as Experiment 2. Outlier analyses failed to find any outliers for any of the variables. Table 6 displays the descriptive statistics for all measures

Table 6
Experiment 3 descriptive statistics.

Measure	Mean	SD	Skew	Kurtosis
IAT	166.61	106.66	.66	.54
IAT D_1	.48	.27	.11	.14
AMP	5.75	5.89	.28	.43
Stroop	94.35	60.24	.50	.51

Table 7
Experiment 3 correlations for the low feedback condition.

	IAT D_1	AMP	Stroop
IAT D_1	1	.29*	−.15
AMP		1	.00
Stroop			1

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

collapsed across feedback condition. As expected, IAT scores were higher in the high feedback condition (M = .61, SD = .24) than the low feedback condition (M = .38, SD = .24), $t(96) = 4.87$, $p<.001$. Next we consider whether this produced the expected changes in the relationship between IAT scores and Stroop or AMP scores.

Bivariate correlations are reported in Tables 7 and 8 for the low and high feedback conditions, respectively, and scatterplots of these correlations are shown in Figs. 5 and 6. Using the same regression tests as Experiment 2, the Stroop by condition interaction was a significant predictor of IAT scores, $\beta = .28$, $t(92) = 2.39$, $p<.05$. Breaking this down for each condition, participants in the high feedback condition showed a significantly positive relationship between Stroop scores and IAT scores, $\beta = .31$, $t(43) = 2.18$, $p<.05$, whereas the was no reliable relationship in the low feedback condition, $\beta = -.15$, $t(49) = -1.13$, $p = .26$. Similarly, the AMP by condition interaction was a significant predictor of IAT scores, $\beta = -0.27$, $t(92) = 2.41$, $p<.05$. Separate analyses of each feedback condition revealed a significant relationship between the IAT and the AMP for the low feedback group, $\beta = 0.29$, $t(49) = 2.15$, $p<.05$, but not the high feedback group, $\beta = -0.19$, $t(43) = -1.37$, $p = .18$.

In summary, providing false feedback regarding an individual's racial prejudice prior to taking the IAT proved to be a robust method for manipulating the role of cognitive control while taking the IAT. This experiment replicated all three of the effects reported in Experiment 2, but with larger effect sizes and with significant condition interaction terms for both the AMP and Stroop scores. Thus, the role of cognitive control increased in the high feedback condition at the same time that the role of attitude decreased. Under conditions in which participants believe that they scored low on the AMP, there was a relationship between the subsequent IAT measure and the AMP and little or no relationship with Stroop. In contrast, under conditions in which participants believed that they scored high on the AMP, there was little or no relationship between the subsequent IAT measure and the AMP, but there was a relationship with Stroop. In summary, the role of cognitive control while taking the race IAT can be reduced (and the role of attitude increased) if the test taker is told in advance of taking the IAT that other indicators suggest that they are not prejudiced.

General discussion

Summary of results

The purpose of our research was to measure and manipulate the separate contributions of attitude and cognitive control while taking the IAT. Previous research measured different forms of cognitive control (e.g., Klauer et al., 2010), or different measures of attitude (e.g., Nosek & Smyth, 2007), of individuals who took the IAT. However, because a correlation might exist between cognitive control and attitudes, conclusions drawn from this work are limited. In contrast, Experiment 1 collected independent measures of both cognitive control (Stroop and stop-signal) and attitude (explicit ratings and the AMP) to properly isolate how each affected performance on two different IATs (race and politics). This allowed us to run factor analyses that included both attitude and cognitive control as separate factors (previous work only included one or the other factor). An attitude-only model failed to capture the data whereas a cross-loading model in which IAT scores reflected both cognitive control and

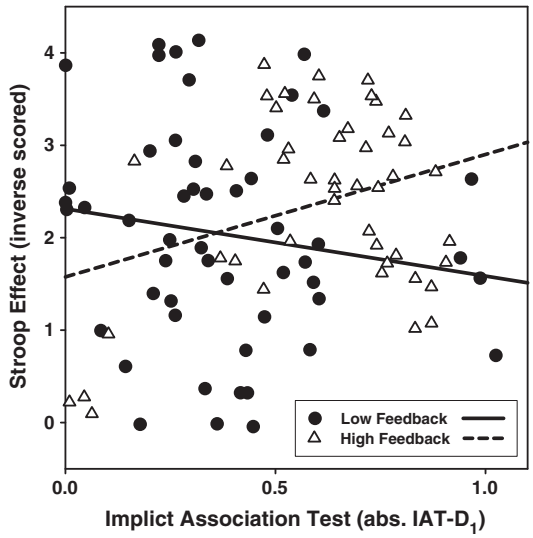

Fig. 5. Scatter plot of the correlations between the race Implicit Association Test (IAT) and the Stroop effect for Experiment 3. Representing implicit attitude strength regardless of preference, the absolute magnitude of IAT scores are shown, with the IAT calculated using the D_1 scoring algorithm. To reduce outliers, the Stroop effect vales were calculated by subtracting the average inverse reaction time of correct incongruent trials from the average inverse reaction time of correct congruent trials. For the high and low feedback conditions, separate correlations were calculated as indicated by the regression lines (see Tables 7 and 8 for the corresponding correlation coefficients).

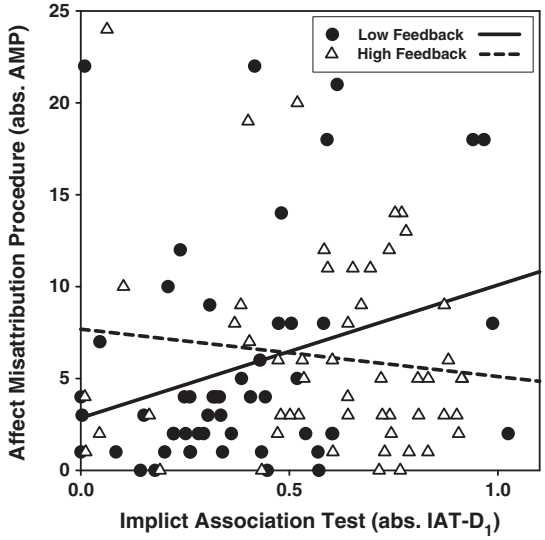

Fig. 6. Scatter plot of the correlations between the race Implicit Association Test (IAT) and the Affect Misattribution Procedure (AMP) scores for Experiment 3. Representing implicit attitude strength regardless of preference, the absolute magnitude of AMP and IAT scores are shown, with the IAT calculated using the D_1 scoring algorithm. For the high and low feedback conditions, separate correlations were calculated as indicated by the regression lines (see Tables 7 and 8 for the corresponding correlation coefficients).

Table 8
Experiment 3 correlations for the high feedback condition.

	IAT D_1	AMP	Stroop
IAT D_1	1	−.12	.32*
AMP		1	−.04
Stroop			1

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

attitude provided a significantly better explanation. Klauer et al. also concluded that cognitive control plays a role in IAT performance, but our study ruled out the possibility that this conclusion was an artifact of a correlation between cognitive control and attitudes.

Similar to previous studies, Experiment 1 was purely correlational and it was not clear whether individual differences in cognitive control caused differences in IAT performance. Therefore, Experiments 2 and 3 manipulated the role of cognitive control while taking the race IAT. In Experiment 2 this was done by informing participants regarding the measurement goal of the IAT for the high knowledge condition and in Experiment 3 this was done by giving participants positive or negative false feedback regarding their racial prejudice prior to taking the IAT. These manipulations were chosen because they might make the associations of racial prejudice more salient and/or they might induce anxiety, resulting in reduced cognitive resources. Regardless of the mechanism, it was predicted that these manipulations should increase response interferences, producing larger IAT effects on average, an increased correlation with cognitive control (Stroop) and a decreased correlation with attitude (AMP). All three of these predictions were confirmed.

Implications of results

The IAT has become increasingly popular for the advantages it offers over traditional explicit measures. Not only does the IAT allow measurement of socially sensitive attitudes (Greenwald et al., 2009) but it can also be used to measure unconscious attitudes such as self-esteem (Greenwald & Farnham, 2000; Greenwald et al., 2002). Although the IAT is widely used, its validity remains a concern. Some studies considered whether the IAT reflects attitudes or merely salient associations (e.g., Han et al., 2010; although see Siegel, Sigall, & Huber, 2012) whereas other work considered whether individual differences on the IAT reflect strength of attitude or merely cognitive control (e.g., Klauer et al., 2010). Our result validated some of these concerns, demonstrating that 1) an attitude-only account of the IAT is insufficient as compared to an account that also includes cognitive control; and 2) making the IAT's measurement goal salient or worrisome increased IAT scores and increased the role of cognitive control.

In light of these results, research using the IAT should proceed cautiously in the absence of measuring cognitive control. For instance, correlations between IAT scores and other behaviors may be due to latent (i.e., third variable) correlations between cognitive control and the correlated behaviors. However, most researchers use the IAT as an aggregate measure to compare implicit attitudes for different groups of participants or participants in different conditions. When using the IAT in this manner, any differences or lack of differences may be confounded with cognitive control differences between the groups that occur naturally or due to cognitive load differences imposed by different conditions. For example, because cognitive control increases (Davidson, Amso, Anderson, & Diamond, 2006) and then decreases (Hasher et al., 1989) across the lifespan, comparisons of IAT scores from younger versus older participants (e.g., Baron & Banaji, 2006; Nosek & Smyth, 2007) may reflect differences in cognitive control rather than different attitude strengths. Aside from differences in cognitive control, the manipulations of Experiments 2 and 3 suggest a second issue when comparing IAT scores from different groups. If different groups of participants, or participants in different conditions, are predisposed to different levels of concern or awareness regarding the IAT's measurement goal, this factor may give the false impression that the groups have different implicit attitudes.

Acknowledgments

This research was supported by National Science Foundation Grants BCS-0843773 to David Huber and SES-0620062 to Michael Dougherty.

Appendix A. Political Attitude Scale (PAS)

1. I support setting a withdrawal timetable for US troops to leave Iraq.
2. I support unrestricted free trade.
3. Human pollution is a significant cause of global warming.
4. There should be a way for illegal immigrants to apply for citizenship.
5. The government should provide access to health care regardless of peoples' ability to pay.
6. The country should have a temporary guest worker program for immigrants.
7. I think the wealthiest Americans should pay more taxes.
8. Gay marriage should be legal.
9. I support tax cuts for middle-class families.
10. The US should have a long-term presence in Iraq.[1]
11. The government should construct a better fence along the US and Mexico border. [1]
12. The United States should be able to the use military force on countries it sees as a threat without the approval of the United Nations. [1]
13. Cities that don't enforce immigration laws should not receive funding from the federal government. [1]
14. Private, market based health insurance is the right way to provide health care to Americans. [1]

[1] Reverse coding was used for these questions.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, California: Sage Publications.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, & F. Caski (Eds.), *Proceedings of the second international symposium on information theory* (pp. 267–281). Budapest, Hungary: Akademiai Kiado.

Arbuckle, J. L. (1999). *Amos 4.0.* Chicago, IL: SmallWaters.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes—Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, *17*(1), 53–58.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*(4), 548–564.

Bishop, S. J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nature Neuroscience*, *12*(1), 92–98.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christe, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, *42*(2), 192–212.

Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen, & J. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Prk, CA: Sage.

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*(4), 469–487.

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*(5), 800–814.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078.

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology*, *51*(2), 305–320.

Dempster, F. N., & Corkill, A. J. (1999). Interference and inhibition in cognition and behavior: Unifying themes for educational psychology. *Educational Psychology Review*, *11*(1), 1–88.

Diamond, A., & Gilbert, J. (1989). Development as progressive inhibitory control of action—Retrieval of a contiguous object. *Cognitive Development*, *4*(3), 223–249.

Dougherty, M. R. P., & Hunter, J. E. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychologica*, *113*(3), 263–282.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology. General*, *133*(1), 101–135.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science*, *4*(5), 294–298.

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*(1), 3–25.

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*(6), 1022–1038.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41.

Han, H. A., Czellar, S., Olson, M. A., & Fazio, R. H. (2010). Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology*, *46*(2), 286–298.

Hasher, L., Rypma, B., Stoltzfus, E. R., & Zacks, R. T. (1989). Age deficits in inhibitory mechanisms—Data and theory. *Bulletin of the Psychonomic Society*, *27*(6), 496.

Hodson, G., & Busseri, M. A. (2012). Bright minds and dark attitudes: lower cognitive ability predicts greater prejudice through right-wing ideology and low intergroup contact. *Psychological Science*, *23*(2), 187–195.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453.

Jang, Y., Wixted, J. T., & Huber, D. E. (2009). Testing Signal-Detection Models of Yes/No and Two-Alternative Forced-Choice Recognition Memory. *Journal of Experimental Psychology. General*, *138*(2), 291–306.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, *81*(5), 774–788.

Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the Implicit Association Test: Why flexible people have small IAT effects. *Quarterly Journal of Experimental Psychology*, *63*(3), 595–619.

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the Ability to Inhibit Simple and Choice Reaction-Time Responses - a Model and a Method. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(2), 276–291.

Macleod, C. M. (1991). Half a Century of Research on the Stroop Effect - an Integrative Review. *Psychological Bulletin*, *109*(2), 163–203.

McConahay, J. B. (1986). Modern racism, ambivelance, and the modern racism scale. In J. F. Dovidio, & S. L. Gaertner (Eds.), *Prejudice, discrimination and racism* (pp. 91–126). New York: New York Academic.

McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, *20*(6), 483–510.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift Fur Experimentelle Psychologie*, *48*(2), 107–122.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology*, *85*(6), 1180–1192.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the implicit association test - Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology*, *54*(1), 14–29.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science*, *14*(6), 636–639.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277–293.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811–832.

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*(3), 287–290.

Ridderinkhof, K. R., vanderMolen, M. W., Band, G. P. H., & Bashore, T. R. (1997). Sources of interference from irrelevant information: A developmental study. *Journal of Experimental Child Psychology*, *65*(3), 315–341.

Schmader, T. (2010). Stereotype threat deconstructed. *Current Directions in Psychological Science*, *19*(1), 14–18.

Schmitz, F., Teige-Mocigemba, S., Voss, A., & Klauer, K. C. (2011). When scoring algorithms matter: Effects of working memory load on different IAT scores. *The British Journal of Social Psychology*, *19*, http://dx.doi.org/10.1111/j.2044-8309.2011.02057.x.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*(2), 314–335.

Siegel, E., Sigall, H., & Huber, D. E. (2012). The IAT is sensitive to the perceived accuracy of newly learned associations. *European Journal of Social Psychology*, *42*, 189–199.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test-performance of African-Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods*, *40*(2), 479–483.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*(1), 5–13.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*(1), 101–126.