Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Experimental Social Psychology 42 (2006) 259–272
Journal of Experimental Social Psychology
www.elsevier.com/locate/jesp

The influence of experimentally created extrapersonal associations on the Implicit Association Test ☆

H. Anna Han [a], Michael A. Olson [b], Russell H. Fazio [a,*]

[a] Department of Psychology, 1885 Neil Avenue, Ohio State University, Columbus, OH 43210-1222, USA
[b] Department of Psychology, Austin Peay Building, University of Tennessee, Knoxville, TN 37996, USA

Received 1 July 2004; revised 10 February 2005
Available online 13 June 2005

Abstract

We examined the influence of extrapersonal associations (Olson & Fazio, 2004)—associations that neither form the basis of the attitude nor become activated automatically in response to the object—on the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) by experimentally creating both attitudes and extrapersonal associations. The results revealed that participants who were given extrapersonal information that was inconsistent with their own attitudes were affected by this information when they later performed an IAT. They exhibited significantly reduced IAT scores compared to participants who were provided attitude–consistent extrapersonal information. This attenuation of the IAT effect occurred despite the fact that participants rated the source of the attitude–inconsistent extrapersonal information as irrational and foolish. On the other hand, the extrapersonal associations did not influence a subliminal priming measure in Experiment 1, nor a personalized version of the IAT (Olson & Fazio, 2004) in Experiment 2. These measures proved sensitive to the attitude, regardless of the congruency of the extrapersonal information.

Keywords: Attitude measurement; Implicit measure of attitudes; Implicit Association Test; Priming

Introduction

From mundane daily routines to major life decisions, our behavior is often influenced by relatively automatic and efficient processes (Bargh, 1999), and implicit measures have allowed researchers a glimpse into the cognitive apparatus that presumably produces these behaviors (Banaji, 2001; Fazio & Olson, 2003). The Implicit Association Test (IAT, Greenwald et al., 1998) in particular has become a frequently employed means of assessing attitudes and beliefs implicitly (Fazio & Olson, 2003; Greenwald & Nosek, 2001). Although some of this work suggests that the IAT offers a valid estimate of these associations (e.g., McConnell & Liebold, 2001; Olson & Fazio, 2001), other work indicates that it may be influenced by more than researchers typically consider (e.g., Karpinski & Hilton, 2001; Olson & Fazio, 2004).

In fact, some recent work suggests that IAT can be influenced by attitude–irrelevant associations. These *extrapersonal associations*, as they have been termed, are "associations that are available in memory but are irrelevant to the perceived likelihood of personally experiencing positive or negative outcomes upon interacting with the attitude object" (p. 653, Olson & Fazio, 2004; see Karpinski & Hilton, 2001, for a related view). The research presented here explores the degree to which

☆ This research was supported by Senior Scientist Award MH 01646 and Grant MH 38832 to the third author, and while the first author was supported by Grant T32-MH19728 from the National Institute of Mental Health. The research comprised a portion a master's thesis conducted by H. Anna Han under the supervision of the third author. We thank Suzanne Miller for her assistance with stimulus preparation and programming, Nicole Deland and Sarah Lifland for their assistance with data collection, and Jan De Houwer for his thoughtful feedback on an earlier version of the manuscript.
* Corresponding author. Fax: +1 614 688 5414.
E-mail address: fazio.11@osu.edu (R.H. Fazio).

0022-1031/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2005.04.006

such extrapersonal associations influence individuals' responses on the IAT.

The IAT

The IAT (Greenwald et al., 1998) measures the degree to which two target concepts (e.g., Black/White) are associated in memory with positive and negative evaluations (e.g., Good/Bad). In this measure, the task is to quickly categorize the stimuli by using one of two key assignments (Left/Right). The assumption is that when two highly associated concepts are assigned to the same response key (e.g., White and Good vs. Black and Bad), the participant will be faster to categorize stimulus examples than when two non-associated concepts are assigned to the same key (e.g., White and Bad vs. Black and Good; Greenwald & Nosek, 2001). The difference in response latencies between the categorizations of the two target concepts provides an attitude estimate (e.g., "racial bias"). Unfortunately, even though a myriad of empirical work has been devoted to the IAT, uncertainty remains about its underlying mechanisms and exactly what it measures.

What does the IAT measure?

A number of suggestions have been made regarding the possible mechanisms underlying the IAT: (a) shift-in-response criteria (Brendl, Markman, & Messner, 2001), (b) figure-ground asymmetry (Rothermund & Wentura, 2001, 2004; also see Greenwald, Nosek, Banaji, & Klauer, in press), and (c) task-set-switching (Klauer & Mierke, 2005; Mierke & Klauer, 2001). However, most pertinent to our current view is the idea that the IAT operates at the level of the category. For example, recent work by De Houwer (2001) suggests that the IAT may be measuring evaluative associations with the category response labels and *not* attitudes automatically activated by the category exemplars. De Houwer (2001) created a "British-Foreigner" IAT and presented to participants both positively and negatively evaluated British (e.g., Princess Diana, a mass murderer) and foreign (e.g., Einstein, Hitler) exemplars. The findings indicated that British participants showed a preference for Brits over foreigners regardless of whether the exemplars were positive or negative, corroborating the idea that the IAT effects are substantially determined by associations to the category labels. More specifically, both Einstein and Hitler appear to have been readily categorized as "foreigner," and how easily participants could associate "foreigner" with positive versus negative determined performance. The instruction to categorize the target items appears to encourage their construal only as representatives of their respective categories, resulting in IAT scores based predominantly on associations to category labels. The suggestion that the IAT operates at the level of the category is buttressed further by Olson and Fazio's (2003) demonstration that a racial IAT correlates more strongly with a priming measure when the latter explicitly encourages categorization by race.

This is not to say that the IAT is never affected by the nature of the exemplars used to represent the categories. The argument is simply that such influence occurs, at least in part, through construal of category labels—a point nicely demonstrated by Govan and Williams (2004). These investigators employed the traditional insect vs. flowers IAT, but changed the entire set of exemplar stimulus items from positive flowers (e.g., tulip) to unpleasant flowers (e.g., Venus flytrap) and from unpleasant insects (e.g., cockroach) to more positively viewed insects (e.g., ladybugs). Participants' IAT performance indicated an apparent preference for insects over flowers. Govan and Williams (2004) argued that this reversal of the typical IAT effect was due to the atypical valence of the stimulus items having led participants to redefine, at least temporarily, the category labels of flower and insects. A second experiment replicated the reversal of the IAT effect, and also showed evidence for category re-definition. When participants first performed a plants/animals IAT with nice plants/nasty animals (e.g., lily/pit-bull) or nasty plants/nice animals (e.g., pondweed/puppy), they came to construe the category labels plants and animals in a manner that reflected the valence of the presented exemplars, i.e., nice plants and nasty animals or nasty plants and nice animals. Evidence of this re-definition of the category labels was apparent on a subsequent task; participants responded to the category labels in a manner that was consistent with the previously experienced, albeit atypical, exemplars.

The findings of De Houwer (2001), Olson and Fazio (2003), and Govan and Williams (2004) imply that the IAT does *not* measure the automatic activation of evaluations in response to exemplars, but rather, the association between the dual meanings of any given response option. Thus, IAT performance depends on one's ability to solve the response mapping problem posed by the IAT. To perform well, one needs to remember the dual meaning of any given key. Doing so may be affected by more than attitudes toward the relevant categories. Participants may be influenced by *any* information that can facilitate the task at hand, including information that is potentially inconsistent with their attitudes (Olson & Fazio, 2004). In fact, upon failing to find correspondence between an IAT and an explicit measure regarding preferences for apples vs. candy bars (i.e., individuals preferred candy bars on the explicit measure and apples on the IAT), Karpinski and Hilton (2001) argued that the IAT is influenced by cultural norms, and that the normative information need not necessarily reflect one's own attitudes. They argued, "In our society, there are an abundance of positive associations and virtually no

negative associations with apples. For candy bars, however, the message is much more mixed" (p. 783). Karpinski and Hilton suggested that such normative information may affect participants' approach to the mapping difficulties posed by the IAT.

Olson and Fazio (2004) extended the idea that the IAT is influenced by normative information and argued that the IAT may be influenced by any "extrapersonal association." That is, the IAT may be influenced not only by information that is prevalent in our culture, but also by information that, although not the basis of one's personal evaluations and irrelevant to any privately made approach–avoidance decision, nevertheless is valenced and available in memory. According to Olson and Fazio, an extrapersonal association may emanate from the media, a colleague, a significant other, and, as we will argue later, anyone—even if the person is thought to be obtuse and unintelligent.

Such extrapersonal associations are to be distinguished from one's own attitude (or personal associations) (Fazio, 1995; Fazio, Chen, McDonel, & Sherman, 1982; Fazio, Powell, & Herr, 1983; Katz, 1960). Unlike extrapersonal associations, one's own attitude serves the functional value of guiding approach and avoidance decisions (e.g., Fazio, 1995; Katz, 1960; also see Fazio, 2001; for more details). It must be noted that the distinction between extrapersonal and more personal associations does not imply that one's attitudes are uninfluenced by other individuals or the culture. Instead, Olson and Fazio (2004) call attention to the possibility that an individual's attitude can deviate from the opinion of others or from the cultural norm, and it is ultimately the personal association that will automatically be activated and guide one's behavior, especially when little motivation or opportunity exists to engage in more deliberative processing (see Fazio & Towles-Schwen, 1999).

Olson and Fazio (2004) argued that features of the IAT encourage the use of extrapersonal associations. The category labels "pleasant" and "unpleasant" imply that a "correct" answer exists and, hence, transmit a normative implication. Moreover, feedback is provided when one makes an error. Olson and Fazio proposed a "personalized" IAT designed to be *less open* to the influence of extrapersonal associations. In this version of the IAT, the implied normative responses were eliminated by changing the labels "pleasant" and "unpleasant" to "I like" and "I don't like" and by removing the error feedback.[1]

How do extrapersonal associations influence the traditional IAT and how does the personalized version remove such influences? Let's imagine that you are participant in a study and are asked to complete an IAT regarding apples and candy bars. You have been told that the task is to categorize the items presented on the screen as quickly as possible. In the traditional version of the IAT, you are presented with response options labeled "apples/pleasant" and "candy bars/unpleasant." You are an insatiable candy bar lover. How would you confront this mapping problem? Some relevant information from memory might be activated, albeit not necessarily to the point of conscious awareness. For example, it might occur to you that candy bars are fattening and can cause cavities, or perhaps you may recall the saying, "an apple a day keeps the doctor away." Maybe the perspective of a close other who love apples is activated in memory. Recall that there is nothing in the traditional IAT that prevents such extrapersonal associations from being activated in memory. Your job is to successfully categorize the stimuli in terms of "apples/pleasant" and "candy bars/unpleasant." On the other hand, in the personalized version of the IAT, the response options that you need to use are "I like/apples" and "I don't like/candy bars." The labels "I like" and "I don't like" direct you to focus on your own evaluation, and not on some normative information or extrapersonal associations. Thus, the personalized IAT reduces the impact of extrapersonal information and increases the likelihood that it will measure the evaluations that you personally associate with apples and candy bars.

Olson and Fazio (2004) demonstrated the value of personalizing the IAT in four studies. In Experiments 1 and 2, they explored how the prevailing cultural view of negativity towards Blacks might serve as extrapersonal associations on the traditional IAT. Their results indicated that personalizing the IAT reduced the racial prejudice score among white participants on the IAT. It appears that less extrapersonal information in the form of the cultural view was activated for participants when the labels were "I like" and "I don't like" versus "pleasant" and "unpleasant." Similarly, in Experiment 3, Olson and Fazio found significant correlations between an apples/candy bars IAT and various explicit measures only when the personalized version of the IAT was used and not when the traditional version of the IAT was employed. The personalized version of the IAT was able to predict expressed preferences, past behaviors, and future behavioral intentions. With the personalized version, it seems less likely that the prevailing view of positivity towards apples was activated for participants, thus increasing the likelihood that their IAT performance would reflect their own preferences. Experiment 4 also revealed stronger explicit/implicit correspondence in preferences for Bush or Gore on the personalized IAT than on the traditional version. Unlike the other three experiments, the Bush/Gore IAT did not involve a prevailing cultural view, but participants certainly had available information suggesting that their non-preferred candidate had some positive qualities, or at least was viewed positively by others. This finding suggests that extrapersonal

[1] The personalized and traditional IAT also differ on other dimensions. In some versions of the personalized IAT, normatively valenced pleasant/unpleasant stimuli (e.g., freedom/murder) are replaced by more idiosyncratic stimuli such as coffee and football. However, Olson and Fazio (2004) found that the idiosyncratic version of the personalized IAT is not necessary, and both the normative and idiosyncratic version yielded equivalently same results.

associations do not have to stem from prevailing cultural norms, but can involve any information in memory relevant to the mapping problem posed by the IAT.

What can we conclude from Olson and Fazio's (2004) results? Given that the personalized IAT yields an improvement in correlations with various explicit measures of attitudes and behavior in domains where self-reports tended to be trustworthy (food preferences and voting), the implication is that the traditional IAT is affected by extrapersonal associations. However, as Olson and Fazio admit, the evidence is indirect: "it is important to note that we are drawing inferences about the operation of extrapersonal associations on the basis of the operational modifications we made to the IAT" (p. 664). The benefits of personalizing the IAT have been demonstrated, but whether this stems from the influence of extrapersonal associations on the traditional version of the IAT has not yet been established conclusively. Thus, the purpose of the present research is to examine the issue of extrapersonal associations more directly and experimentally. To test this, we first created attitudes in the lab, with the goal of having attitudes toward one object be clearly more positive than attitudes toward another. Then we experimentally introduced extrapersonal associations regarding the objects. The question was whether the traditional version of the IAT would be affected by the extrapersonal associations.

Because there is a strong indication that what is automatically activated in response to an object is one's personal attitude, in Experiment 1 we contrasted the traditional IAT to a priming measure of attitude. A priming procedure provides a relatively well-understood means of assessing the evaluation automatically activated in response to some prime. Because the prime spontaneously evokes positivity or negativity, people can readily and quickly indicate the connotations of subsequently presented adjectives that are evaluatively congruent. However, if the adjectives are incongruent with the prime, response competition slows the response time (see Fazio, 2001, for a review). Thus, the priming procedure provides an estimate of the overall evaluation automatically activated by the primed object.

In fact, a host of studies have indicated that the priming measure activates one's own evaluation of the object, even when that attitude is not shared by the culture or popular opinion (e.g., Bessenoff & Sherman, 2000; Fazio, 1993; Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Sherman, Presson, Chassin, Rose, & Koch, 2002). In what may be the most direct comparative test, Fazio (1993) showed that what is activated in response to a prime is one's idiosyncratic reaction to the attitude object and not the normative, consensual response. In a multiple regression involving simultaneous entry of participants' personal like or dislike of each attitude object, on the one hand, and categorization of each object as either positive or negative on the basis of the mean rating in a large survey sample, on the other hand, the former idiosyncratic measure of liking proved predictive of the priming effect, whereas the latter normative estimate of liking did not. Given that the priming measure assesses one's own automatically activated evaluation of the primed object, we had every reason to believe that it would prove sensitive to the attitudes created experimentally in lab and that it would not be affected by the experimentally manipulated extrapersonal associations.

In Experiment 2, we contrast the traditional IAT to the personalized IAT. If Olson and Fazio (2004) are correct about why the personalized IAT leads to improvements in attitude measurement, then, just as with the priming measure, it should not be affected by extrapersonal associations.

Experiment 1

The current experiment examined the influence of extrapersonal associations on the two implicit measures, the IAT and priming. We predicted that the IAT would be influenced by the introduction of extrapersonal associations while the priming measure would not.

Method

Participants

One hundred and thirty one introductory psychology students participated in this experiment in partial fulfillment of their course requirements. Participants were randomly assigned to one of two experimental conditions. The data from two participants were excluded because they indicated some awareness of the purpose of the experiment, and from another six for their failure to respond correctly on a critical manipulation check item. The final sample consisted of 123 participants (66 males, 57 females).

Procedure

The current experiment was conducted in groups of 2–4 students, who participated simultaneously, albeit in individual cubicles. Participants were told that the experiment was about comparing adults' and children's decision-making skills, and how adults evaluate children's decision-making. Participants were also told that the topic of the experiment would be Pokémon because it was important to engage the children who were participating in the experiment.

Attitude formation. Participants were given a packet with a set of instructions explaining what Pokémon were and how to play the card game. Participants read that Pokémon is a game that has become popular with elementary school children, and the purpose of the game was to

accumulate every possible type of Pokémon. It was explained that there are approximately 250 different types of Pokémon, each with different names and distinctive skills/powers that can be accumulated by using one's own Pokémon to fight and capture others' Pokémon. Hence, it was made clear to the participants that it would be advantageous to start the collection with a Pokémon that was an all around solid character so it could fight repeatedly against an opponent's Pokémon without being seized.

Participants were then told that they would be given a choice of two Pokémon to start the game, and were handed two Pokémon cards with attribute information regarding each of two obscure Pokémon characters (Metapod and Shelder). They were asked to study the cards carefully, and make an informed decision as adults. In fact, the cards were designed such that they clearly led participants to prefer one specific Pokémon (Shelder) over the other (Metapod). Each card showed a picture of the Pokémon, its name, and a brief description and summary of the character's powers. The cards also displayed various attribute information such as attack and defense abilities, special attack power, special defense, hit points, speed, and an overall score, all clearly indexed by strong/weak red bars.

All participants then completed three 7-point semantic differential scales anchored by *bad–good*, *unfavorable–favorable*, and *weak–strong*, for each Pokémon character. Participants were also asked to write up to five reasons why they preferred their chosen Pokémon. All of this was done in the interest of ensuring that participants consolidated the available information about the two Pokémon characters and formed well-rehearsed attitudes that would prove capable of automatic activation (Fazio et al., 1986).

Extrapersonal associations. Next, under the guise of evaluating children's decision-making skills, all participants were introduced to a controlled extrapersonal association via a short video clip. Depending on the condition to which they had been assigned, participants saw a video providing an extrapersonal association that was either consistent (prefer Shelder) or inconsistent (prefer Metapod) with their preference. In the clip, participants saw two 10-year-old boys responding to an interviewer's questions about the same Pokémon characters to which the participants had been exposed. Like the participants, the boys were asked to decide between the Pokémon. The videos showed the boys articulating their preference for a given Pokémon without justifying their preference to any meaningful extent. For example, in the extrapersonal inconsistent information video, one of the boys stated, "He (Shelder) is not very bright… I would not choose him as my first choice at all," and "he (Metapod) resembles a cocoon, kind of, with two eyes… I like Metapod." Given that the participants had been induced to prefer Shelder over Metapod, the content of the videos was either consistent or inconsistent with the preference they had developed earlier.

After exposure to the video clip, all participants completed an open-ended question regarding the boys' choice of Pokémon, simply as a manipulation check. They then were asked to compare the boys to other 10 year olds on a series of 5-point semantic differential scales. They rated the boys' decision-making skills (*very poor–excellent*), rationality (*not rational at all–very rational*), sensibility (*not sensible at all–very sensible*), and the extent to which they had been influenced by extraneous factors, such as appearance of the Pokémon, that should not have any influence on their decisions (*not influenced at all–very influenced*).

Finally, participants were asked to move to another room so the experimenter could test their "attentional abilities." All participants then completed the both subliminal priming measure and the IAT.[2]

Subliminal priming. The experimenter described the priming procedure as a test of attentional abilities and, more specifically, a word connotation task that required both speed and accuracy. Participants were told that they would see a string of alphanumeric characters on the screen (mask) followed by a word. They were instructed to do nothing with respect to the alphanumeric characters except to keep their eyes focused on them. It was explained that they served as a warning signal that a target word would appear. Participants were then instructed to respond to the target word by indicating its meaning as good or bad using a response box. It was emphasized that speed and accuracy were imperative for this task.

The subliminal priming used in this experiment employed the same parameters as in Olson and Fazio's (2002) procedure. First, the monitors and the video cards were set at a screen refreshment cycle of 14 ms. Each trial started with an alphanumeric string pre-mask of 56 ms, followed by a prime (the word "Shelder" or "Metapod") for 28 ms, and then an alphanumeric string post-mask for 42 ms. Ninety-eight milliseconds later, a target word appeared for a maximum of 1.75 s or until the participant responded. The parameters resulted in an SOA of 168 ms.

Participants first completed a 16-trial practice block with no primes. They then completed two critical blocks with 32 trials each. Each critical block contained two primes (the names, "Shelder" or "Metapod") presented 16 times each, followed half of the time by a positive adjective (e.g. "EXCELLENT", "MAGNIFICENT") and the other half of the time by a negative adjective (e.g., "HORRIBLE", "INFERIOR"). Across the two blocks, each

[2] The order was not counter-balanced; the IAT always came second. Given that the primes are presented subliminally, the subliminal priming measure does not prompt any additional conscious consideration of the Pokemon. In contrast, the IAT obviously forces participants to consciously consider the Pokemon. Hence, administering it first had the potential to influence the priming measure.

prime was followed by each of 16 positive and 16 negative target adjectives. The computer program randomly ordered the prime–target pairs for each participant.

The Implicit Association Test. The experimenter described the IAT as a different test of attentional abilities—a categorization task that again required both speed and accuracy. Participants were told that they would be asked to categorize some of the same Pokémon characters they saw earlier in the experiment, along with words that generally have positive and negative connotations. Participants were then instructed to categorize the items on the screen using the keyboard.

The IAT used in this experiment was designed following a procedure established by Greenwald et al. (1998) and is nearly identical to the one used by Olson and Fazio (2001). A total of 12 blocks were presented with 50 trials each. The first 4 blocks were practice. Blocks 1–2 required categorizing the Pokémon stimuli. Pictures and the names Shelder and Metapod, sometimes written in upper case and sometimes in lower case letters, served as the stimuli. Blocks 3–4 required the categorization of valenced words as pleasant/unpleasant (e.g., "love," "murder"). Three critical combined blocks (blocks 5–7) were then presented with one Pokémon character being paired with the pleasant category and the other with the unpleasant category. The next two practice blocks (8–9) involved categorization of the Pokémon names with the keys reversed relative to blocks 1 and 2. Three more critical combined blocks (10–12) were presented but in reverse categorization from blocks 5–7. Instructions on the meaning of the keys and type of items to categorize were presented at the beginning of each block. The order in which the participants performed the critical combined blocks was counterbalanced.

After completing both the subliminal priming measure and the IAT, participants were probed for suspicion, especially regarding the perceived purpose of the experiment, and awareness of the presence of the primes in the priming phase. In addition, upon completion of the experiment and as a part of the debriefing, participants were asked which Pokémon card they would like to take home as a gift. This served as a final measure of their Pokémon preference. Participants then were debriefed, thanked, and dismissed.

Results

Manipulation check

Upon studying the two Pokémon cards, all participants chose the Pokémon that was objectively superior (Shelder) and accordingly, rated it as more favorable ($M = 2.36$, $SD = .69$) than the objectively inferior Pokémon, Metapod ($M = -1.90$, $SD = .91$), $t(1,128) = .36.198$, $p < .001$. Because it was important to ensure that participants paid attention to the video and were able to remember the extrapersonal association, we analyzed participants' response to the query about the boys' opinion. As mentioned above, six participants recalled the wrong preference, and consequently, their data were excluded from subsequent analyses.

Priming

Because the response latencies displayed a normal distribution, all analyses were conducted using raw latencies. The latencies were submitted to a 2 (Experimental Condition: Extrapersonal Consistent vs. Inconsistent) × 2 (Prime: Shelder vs. Metapod) × 2 (Target Valence: Positive vs. Negative) ANOVA with repeated measures on the latter two factors. The analyses revealed a main effect of target valence, $F(1,121) = 50.98$, $p < .001$, whereby positive adjectives were responded to more quickly than negative adjectives. However, as expected, a Prime × Target interaction emerged, $F(1,121) = 4.58$, $p = .034$, indicating that response latencies to positive versus negative targets varied as a function of which prime had been presented. Participants were relatively fast to respond to negative targets when they were preceded by the objectively more negative Pokémon prime (Metapod) ($M = 681$ ms, $SD = 109$) than when they were preceded by the objectively more positive Pokémon prime (Shelder) ($M = 695$ ms, $SD = 116$). Similarly, participants were relatively faster to respond to positive targets when they were preceded by the objectively more positive Pokémon prime ($M = 653$ ms, $SD = 95$) than when they were preceded by the objectively more negative Pokémon prime ($M = 656$ ms, $SD = 101$). Thus, the priming measure proved sensitive to participants' Pokémon preferences. Importantly, the three-way interaction (Prime × Target × Condition) failed to reveal any moderating effect of extrapersonal consistent vs. inconsistent conditions, $F < 1$. Thus, although the priming measure was sensitive to participant's own preference, it was not influenced by extrapersonal associations[3] (see Fig. 1).

[3] The priming measure was able to distinguish the participants' own preferences towards the Pokemon characters in a pilot study in which half of the participants were led to like Shelder and dislike Metapod and the other half to dislike Shelder and like Metapod. The latencies from the priming measure were submitted to a 2 (Experimental Condition: Like Shelder/Dislike Metapod vs. Dislike Shelder/Like Metapod) × 2 (Prime: Liked vs. Disliked Pokémon) × 2 (Target Valence: Positive vs. Negative) ANOVA with repeated measures on the latter two factors. The expected Prime × Target valence interaction was observed, $F(1,100) = 5.33$, $p = .023$ such that participants were relatively faster to respond to positive targets when they were preceded by the liked Pokémon prime ($M = 623$ ms, $SD = 93$) than when they were preceded by the disliked Pokémon prime ($M = 633$ ms, $SD = 98$), and relatively faster to respond to negative targets valence when they were preceded by the disliked Pokémon prime ($M = 662$ ms, $SD = 105$) than by the liked Pokémon prime ($M = 671$ ms, $SD = 93$). In short, the priming measure was sensitive to participants' own preference and was able to distinguish between the liked and disliked Pokemon, just as it did in Experiment 1.

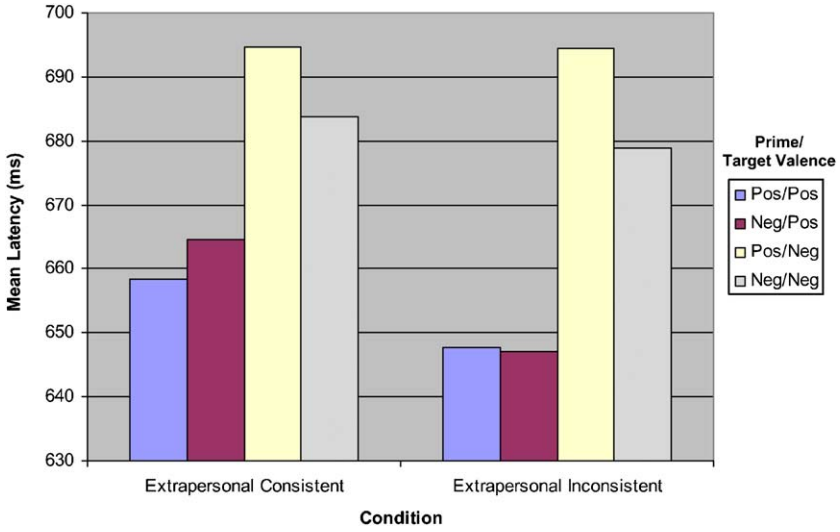

Fig. 1. Mean latency to positive and negative target adjectives as a function of prime valence and extrapersonal association conditions.

IAT

Following the standard procedure established originally by Greenwald et al. (1998), the response latencies for the first two trials of each block were dropped and latencies under 300 and over 3000 recoded to 300 and 3000, respectively. Then the latencies were log transformed. All analyses were done using the log transformation, but raw latencies will be presented for ease of interpretation. Means of each critical block type were then computed (Blocks 5–7 and Blocks 10–12). Block compatibility was defined in reference to the originally developed attitude. Thus, for all participants, compatible blocks were identified as Shelder/+ vs. Metapod/− and incompatible blocks as Metapod/+ vs. Shelder/−.

The log transformed block means were then entered into a 2 (Block Type: incompatible vs. compatible) × 2 (Experimental Condition: Consistent vs. Inconsistent Extrapersonal association) × 2 (IAT Order: Compatible or Incompatible First) ANOVA with repeated measures on the first factor. As expected, the results revealed a typical block type main effect, $F(1,119) = 96.52$, $p < .001$ where participants were quicker to respond in the compatible blocks (Shelder/Pleasant) ($M = 650.38$, $SD = 91.10$) than in the incompatible blocks (Shelder/Unpleasant) ($M = 701.85$, $SD = 114.90$). Furthermore, a Block Type × IAT order interaction emerged, $F(1,119) = 96.52$, $p < .001$, indicating that participants' response latencies were faster in the incompatible blocks if they completed the IAT order that was inconsistent to their attitude first ($M = 675.92$, $SD = 104.85$) than if they completed the IAT order that was consistent to their attitude first ($M = 729.96$, $SD = 119.50$), $t(121) = 2.72$, $p = .008$. Such differences in response latencies did not subsist in the compatible blocks (consistent to attitude first $M = 650.89$, $SD = 85.64$; inconsistent with attitude first $M = 649.83$, $SD = 97.44$), $F < 1$.

The predicted Block Type × Condition interaction also emerged, $F(1,119) = 4.63$, $p = .033$. In the extrapersonal inconsistent condition, the mean for the compatible blocks was 652.80 ($SD = 97.67$), in contrast to 693.51 ($SD = 124.38$) on the incompatible blocks. In the extrapersonal consistent condition, the mean latencies on compatible blocks was 648.16 ($SD = 85.34$), in contrast to 709.53 ($SD = 105.81$) on the incompatible blocks. Therefore, the extrapersonal inconsistent condition was characterized by a significant reduction on the IAT effect ($M = 40.71$, $SD = 65.68$) relative to the extrapersonal consistent condition ($M = 61.37$, $SD = 66.05$). That is, participants in the inconsistent extrapersonal information exhibited a significantly reduced preference for the obviously superior Pokémon compared to the participants who were exposed to the consistent extrapersonal information (see Fig. 2).

The new IAT scoring algorithm. In late 2003, after most of the present experimental work was completed, Greenwald, Nosek, and Banaji suggested a new scoring algorithm for the IAT based on their large web-based samples. A set of analyses was conducted closely following their 12 step procedure, with some minor modifications as delineated below. Because the IAT employed in the experiment used a 12 block format as opposed to the 7 block IAT Greenwald, Nosek, and Banaji (2003) considered when developing the new algorithm, critical blocks B5, B6, B7, B10, B11, and B12 were used in the current analysis. According to their recommendation, trials greater than 10,000 ms (0 cases) and participants for whom more than 10% of trials have latencies less than 300 ms (0 cases) are to be eliminated. All error latencies were replaced by the block mean plus a 600 ms "penalty." Then, the means for each block and three pooled standard deviations for all trials in B5 and B10,

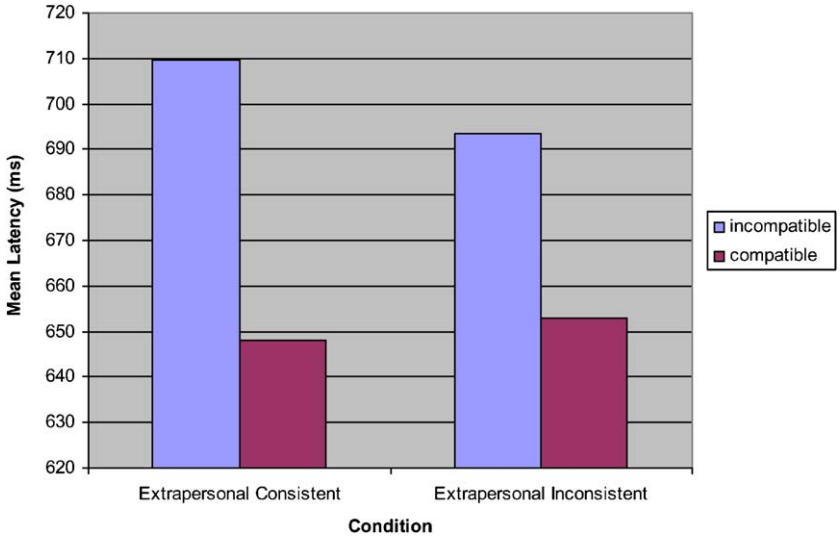

Fig. 2. Mean latency on attitudinally compatible and incompatible blocks of the IAT as a function of extrapersonal association conditions.

B6 and B11, and B7 and B12 were computed. Next, the difference scores of mean latencies of B10 from B5, B11 from B6, and B12 from B7 were computed, and depending on the IAT order, the difference score was reversed (i.e., B5–B10). Each difference score was then divided by its corresponding pooled standard deviation calculated earlier, and finally, these three quotients were averaged.

The new scoring algorithm mirrored the previous results. The single IAT "*D*" score the new algorithm produces proved to be significantly different from 0 ($M = .16$, $SD = .25$), $t(122) = 6.92$, $p < .001$, indicating that, on the average, participants preferred the objectively more positive Pokémon. However, further analyses showed that the extent to which this was true varied by condition, just as predicted. We submitted the IAT scores to a 2 (Experimental Condition: Consistent vs. Inconsistent Extrapersonal association) × 2 (IAT Order: Compatible or Incompatible First) ANOVA. The analysis revealed the IAT order main effect,($M = .05$, $SD = .24$ vs. $M = .28$, $SD = .20$, incompatible vs. compatible block first, respectively) $F(1,119) = 34.53$, $p < .001$. More importantly, the experimental condition main effect $F(1,119) = 8.20$, $p = .005$ indicated that the extrapersonal inconsistent condition displayed a reduced preference toward the objectively more positive Pokémon ($M = .10$, $SD = .24$) compared to the extrapersonal consistent condition ($M = .21$, $SD = .25$).[4] The two-way interaction of Condition and IAT order was not significant, $F < 1$.

Ancillary comparative analysis. The results presented thus far show the IAT to have been affected by whether the extrapersonal associations were or were not attitudinally consistent, whereas the priming measure was not affected. To explore the idea that the IAT was more influenced by the extrapersonal associations than the priming measure was, we conducted an ancillary analysis comparing the two measures. We first computed a single score for each participant from the latency data in the priming task (the difference between the mean latencies for positive and negative targets when preceded by Shelder versus Metapod primes), such that higher scores indicated greater preference for Shelder. These scores, as well as the IAT "*D*" scores were standardized so as to place them on a comparable metric. These *z*-scores were submitted to a 2 (Experimental Condition: Extrapersonal Consistent vs. Inconsistent) × 2 (Measure: IAT vs. priming) ANOVA with repeated measures on the last factor. The results revealed a Measure × Condition interaction that approached a conventional level of statistical significance, $F(1,121) = 3.55$, $p = .073$. The nature of the extrapersonal associations tended to have a greater impact on the IAT scores than on the priming scores.

Ratings of the source of extrapersonal associations. We analyzed participants' ratings of the source of the extrapersonal associations (i.e., the boys). Recall that participants rated the boys' decision-making skills, rationality, sensibility, and the extent to which they had been influenced by extraneous factors compared to other 10-year-olds. Because participants' ratings on the various scales were highly correlated, these ratings were averaged to form an overall index ($\alpha = .86$). Participants in the extrapersonal inconsistent condition ($M = -.35$, $SD = .87$) viewed the boys less positively than did participants in the extrapersonal consistent condition ($M = 1.34$, $SD = .38$), $t(121) = 14.18$, $p < .001$. Both means significantly differed from 0, $t(63) = 23.43$, $p < .001$ and

[4] On average, participants responded with 96 % accuracy. This accuracy rate was unaffected by either the compatible vs. incompatible nature of the response mapping or the manipulation of extrapersonal information, all F's < 1.

$t(58) = 3.11$, $p < .003$, for consistent and inconsistent conditions, respectively. Since "0" was labeled as average and the questions asked for comparison with other 10-year-olds, these findings indicate that the boys were viewed as "above average" in the consistent condition and "below average" in the inconsistent condition.

We also conducted a secondary analysis to consider the possibility that participants who rated the boys more positively might show greater influence of extrapersonal associations on the IAT. Such an effect, especially within the extrapersonal inconsistent condition, might be expected if the attenuated IAT scores we observed are interpreted as arising from the participants' having valued the boys' opinion. However, no correlation was apparent within either the extrapersonal consistent $r = -.08$, or the extrapersonal inconsistent condition, $r = -.14$. Hence, it appears that the IAT was influenced by the extrapersonal associations regardless of what participants thought of the boys. Thus it appears unlikely that the reduction on the IAT scores in the inconsistent condition was due to participants accepting the boys' opinions as a source of good information. Instead, it seems that the extrapersonal information available to these participants eased the incompatible mapping problem on the IAT.

Recall that the participants also had been asked which Pokemon card they would like to have as a token of appreciation for their participation. 92% chose the objectively superior Shelder, and the percentages did not differ by condition (94 vs. 90% in the extrapersonal inconsistent and consistent conditions, respectively). Hence, it seems unlikely that participants in the extrapersonal inconsistent condition experienced any attitude change as a function of their exposure to the boys' judgments. Instead, they belittled the source of this objectively inaccurate information. Nevertheless, this unaccepted extrapersonal association affected their IAT performance.

Discussion

In sum, the current experiment demonstrated an effect of extrapersonal associations on the IAT. Exposure to attitudinally inconsistent extrapersonal information resulted in a significantly reduced preference on the IAT relative to what was observed among participants who were given consistent extrapersonal information. Apparently, when participants in the extrapersonal inconsistent condition were faced with a critical IAT block that was inconsistent with their attitude (Shelder/Unpleasant and Metapod/Pleasant), the extrapersonal information eased the mapping problem posed by the IAT. Although we do not mean to imply that the process involves a conscious strategy, these participants appear to have acted as if they were temporarily adopting the boys' perspective. The reduction in the IAT effect, evident among these participants, occurred despite the fact that they rated the source of the attitudinally inconsistent extrapersonal information (i.e., the boys) as relatively foolish and irrational. Furthermore, it was shown that participants in the extrapersonal inconsistent condition did not change their personal preference in the direction of the source of extrapersonal associations; at the end of the experiment, they displayed a behavioral preference for the objectively superior Pokémon, not for the one the boys preferred. This indicates that even though participants personally did not believe the extrapersonal associations, the attitude-irrelevant information affected their IAT performance. On the other hand, the priming measure proved sensitive to the initial preference and remained unaffected by the presence of the extrapersonal information.

Experiment 2

The previous experiment contrasted the IAT to a priming measure of attitudes. In Experiment 2, we compare two different versions of the IAT—the traditional version and the "personalized" version developed by Olson and Fazio (2004). As noted earlier, these researchers found that modifying the category labels of the IAT from "Pleasant/Unpleasant" to "I like/I don't like" had a marked effect on IAT scores and improved correspondence with various explicit measures of attitudes, past behavior, and behavioral intentions. Olson and Fazio (2004) argued that the traditional IAT is more open to the influence of extrapersonal associations than is the personalized version, although their research provides no direct evidence in support of their interpretation. Experiment 2 tested this hypothesis directly.

We expected that extrapersonal associations would influence the traditional IAT, whereas the personalized IAT would behave more as the priming measure did in Experiment 1. That is, personalizing the IAT should reduce the influence of the extrapersonal associations.

Participants

One hundred and ninety three introductory psychology students participated in this experiment in partial fulfillment of their course requirements. Participants were randomly assigned to one of the four experimental conditions. Data from three participants were excluded for committing a large number of errors (>20%), and two more for failing to provide a correct response on a critical manipulation check. The final sample consisted of 188 participants (130 females, 58 males).

Procedure

The materials and procedure for Experiment 2 were the same as those of Experiment 1 with a few crucial differences. The priming measure was not included. All

participants completed either a traditional or personalized version of the IAT. The resulting design was a 2 (Condition: Extrapersonal Consistent or Inconsistent) × 2 (IAT Type: Traditional or Personalized IAT), with both factors between subjects.

The personalized version of the IAT was identical to the traditional IAT (see the parameters in Experiment 1) with the exception of the category labels, which were changed from "pleasant/unpleasant" to "I like/I don't like" in the personalized IAT. The actual stimuli presented remained the same in both versions. Moreover, in a departure from Olson and Fazio (2004), respondents to both IATs received feedback when they made an error. Thus, the present modification to the IAT is even more focused than in Olson and Fazio (2004); the only difference between the two versions was the category labels.

As in Experiment 1, after having formed attitudes towards the Pokémon and being introduced to an extrapersonal association, participants were escorted to a different room. There they completed either a personalized or traditional version of the IAT. They were then probed for suspicion, debriefed, thanked, and dismissed.

Results and discussion

Manipulation check

All but one participant chose the Pokémon they were expected to choose (Shelder) and accordingly, rated it as more favorable ($M = 2.33$, $SD = .69$) than the objectively inferior Pokémon, Metapod ($M = -1.97$, $SD = .86$), $t(189) = 44.36$, $p < .001$. Participants' recall on the Pokémon preference involved in the extrapersonal association was analyzed. As mentioned above, two participants recalled the preference stated by the boys incorrectly; hence, their data were excluded.

IAT

Mean response latencies from the Shelder/pleasant (compatible) and Shelder/unpleasant (incompatible) blocks were calculated as in Experiment 1. The block means were then entered into a 2 (Block Type: incompatible and compatible) × 2 (Extrapersonal Associations: Consistent vs. inconsistent) × 2 (IAT Type: personalized vs. traditional) ANOVA with repeated measure on the first factor.[5] The analysis revealed a block type main effect, $F(1, 184) = 78.71$, $p < .001$, such that participants were faster to respond during the compatible blocks ($M = 650.88$, $SD = 90.08$) than during the incompatible blocks ($M = 715.72$, $SD = 124.64$). It also showed a marginal IAT type main effect, $F(1, 184) = 3.51$, $p < .07$, such that participants responded in the personalized IAT more slowly ($M = 696.52$, $SD = 99.82$) than in the traditional IAT ($M = 669.78$, $SD = 94.17$).[6] More importantly, the main effects were qualified by the predicted three-way interaction involving Block Type, Condition, and IAT Type, $F(1, 184) = 4.84$, $p = .029$. Collapsing across the IAT block type by computing IAT difference scores revealed that the personalized IAT was unaffected by the extrapersonal associations; the mean IAT score in the extrapersonal inconsistent condition was 80.82 ($SD = 105.34$), which did not differ statistically from the mean of 64.21 ($SD = 106.37$) in the extrapersonal consistent condition, $t < 1$. On the other hand, the traditional IAT was influenced by the extrapersonal associations. The mean IAT score in the extrapersonal consistent condition was 81.49 ($SD = 76.61$), and this effect was reduced by more than half in the extrapersonal inconsistent condition, in which the mean score was 32.53 ($SD = 83.67$), $t(91) = 2.71$, $p = .008$.[7]

The new scoring algorithm also mirrored this effect, $F(1, 182) = 4.61$, ($p = .033$).[8] On the personalized IAT, there was no difference between the extrapersonal inconsistent and consistent conditions ($M = .22$, $SD = .30$ and $M = .20$, $SD = .26$, respectively), $t(92) < 1$, whereas the extrapersonal inconsistent and consistent conditions differed significantly for the traditional IAT ($M = .10$, $SD = .32$ and $M = .27$, $SD = .26$, respectively), $t(90) = 2.76$, $p = .007$.

Ratings of the source of extrapersonal associations

Consistent with the previous experiment, participants' ratings of the source of extrapersonal information

[5] An initial analysis including IAT block order as an additional factor failed to reveal any significant interaction between the variable and the other between subject variable, all F's < 1. As before, inclusion of the IAT order did reveal a Block Type × IAT order effect, $F(1, 180) = 14.92$, $p < .001$. Latencies for the incompatible block were slower when the participants completed the critical pairing that matched their preference first (incompatible block, $M = 734.89$, $SD = 136.47$; compatible block, $M = 640.11$, $SD = 92.4$) compared to when they completed critical pairing that mismatched to their preference first (incompatible block, $M = 697.73$, $SD = 110.13$; compatible block, $M = 640.12$, $SD = 92.40$).

[6] Such slower responding has not always been observed in the previous studies (see Olson & Fazio, 2004), and hence, does not appear necessary for observing benefits to personalizing the IAT.

[7] As in Experiment 1, we explored the possibility of differential accuracy rates on the critical blocks of the IAT as a function of the experimental manipulations. Accuracy rates on the incompatible and compatible blocks were submitted to a 2 (Block type: compatible vs. incompatible) × 2 (Extrapersonals condition: consistent vs. inconsistent) × 2 (IAT type: personalized vs. traditional) ANOVA with repeated measures on the first factor. A main effect of block type emerged, $F(1, 189) = 5.73$, $p = .018$, such that participants displayed a higher accuracy rate (i.e., made fewer errors) in the compatible block ($M = .95$, $SD = .05$) than in the incompatible block ($M = .94$, $SD = .06$). No other main effects or interactions approached significance, all F's < 1. Thus, accuracy rates were unaffected by the type of IAT participants performed or the nature of the extrapersonal information to which they were exposed.

[8] Following Greenwald et al.'s (2003) recommendation for the new scoring algorithm, two participants who were excessively fast in responding (responses on 10% of trials were less than 300 ms) were not included in the analysis.

were combined to form a single index ($\alpha = .86$). As expected, participants in the extrapersonal inconsistent condition rated the boys as less rational, less sensible, and below average in decision making skills ($M = -.25$, $SD = .76$) compared to participants in the extrapersonal consistent condition ($M = 1.32$, $SD = .41$), $t(186) = 17.69$, $p < .001$. Both means significantly differed from 0, $t(96) = 31.84$, $p < .001$ and $t(92) = 3.12$, $p < .002$, for consistent and inconsistent conditions, respectively. As in experiment 1, the boys were viewed as "above average" in the consistent condition and "below average" in the inconsistent condition.

Moreover, the correlation between IAT scores and participants' ratings of the boys mirrored Experiment 1. No correlation was apparent in either conditions (r's of $-.10$ and $-.12$ in the extrapersonal consistent and inconsistent conditions, respectively). Hence, the findings suggest that the participants did *not* view the boys as a credible source of information when the boys expressed a preference for the objectively inferior Pokemon. Moreover, performance on the traditional version of the IAT was affected regardless of what they thought of the boys.

Just as in Experiment 1, then, the traditional IAT was affected by the experimental manipulation of extrapersonal associations. The same was not true of the personalized IAT. It remained sensitive to participant's established attitudes and, in line with Olson and Fazio's (2004) argument, was not influenced by the extrapersonal associations.

General discussion

The two experiments reported here suggest that the traditional IAT is influenced both by one's personal associations and by extrapersonal associations—ones that are attitude-irrelevant but that are valenced and available in memory (Olson & Fazio, 2004). In both experiments, the traditional IAT showed a reduced effect when participants were introduced to an extrapersonal association that was inconsistent with their own attitudes. Importantly, this happened in a context in which individuals were presented with carefully controlled information that clearly dictated the development of a more positive attitude toward an objectively superior object than an inferior alternative. Participants' self-reports demonstrated that they did develop the appropriate attitudes, and one certainly would expect any measure to discriminate between such attitudes. Yet, as demonstrated in the experiments reported here, the capability of the IAT to do so was lessened when an inconsistent extrapersonal association was made available.

It would be difficult to argue that the extrapersonal information (the video of the two boys acclaiming a preference for the objectively inferior alternate) should be considered a legitimate source of attitudinal information. Participants' motives to maintain accurate attitudes presumably led them to disregard the boys as a source of relevant information, especially when compared to the objective information about the attitude objects that the experimenter provided. The boys were viewed by participants as relatively irrational, foolish, and unintelligent—in short, as inappropriate sources of knowledge. In fact, given a behavioral choice between the objectively superior and inferior Pokemon, participants overwhelmingly chose the superior Pokemon. Clearly, participants were successful in disregarding the boys' assertions when it came to their own personal attitudes and choice behavior. However, this extrapersonal information proved relevant and influential when participants were confronted with the mapping problem posed by the traditional IAT, despite the brief exposure to the extrapersonal information (about 35 s of video). Because of the influence of the extrapersonal associations, the attitude estimates of participants completing the traditional IAT represented something of a mix of the attitudes developed via the experimenter-provided information and the opinion expressed by the boys in the video.

On the other hand, the priming measure proved sensitive to the initial preference, but remained unaffected by the presence of the extrapersonal information. Priming measures of attitudes do not force participants to associate a response category representing the attitude object with a given valence as the IAT does. Instead, the attitude object is merely presented as a prime presumably irrelevant to the participant's major task. Indeed, in the current experimental instantiation, the prime was presented subliminally. As a result, it is less likely that information that is available in memory but not attitudinally relevant would be activated. What differentially affects participants' responses to the subsequently presented target words is whatever evaluative response the prime activates automatically. Indeed, this is what makes attitudes and their automatic activation from memory so highly functional for the individual (Fazio, 2000; Smith, Bruner, & White, 1956). Attitudes serve as remarkable tools for efficient object appraisal. As such, they guide approach and avoidance tendencies in a direction that maximizes positive outcomes. Moreover, as noted earlier, priming measures are sensitive to this very functionality. What is automatically activated upon exposure to an object, and can be estimated from the priming data, is the individual's own personal attitudes, not the consensual or cultural view (e.g., Bessenoff & Sherman, 2000; Fazio, 1993; Fazio et al., 1995).

In Experiment 2, it was revealed that the personalized IAT was similar to the priming measure in that it too was not influenced by extrapersonal associations. Experiment 2 went beyond Olson and Fazio's (2004) previous work by experimentally distinguishing personal and extrapersonal associations, i.e., by experimentally creating attitudes in the lab and introducing a controlled

extrapersonal association. Thus, we were able to obtain clear evidence of the causal impact of extrapersonal associations on the traditional IAT. Moreover, by leaving the error feedback intact in the personalized IAT, we were able to show that it is the category labels "Pleasant/Unpleasant" that make the IAT vulnerable to the influence of extrapersonal associations. When participants have to solve an incompatible mapping problem posed by the IAT, any attitudinally inconsistent extrapersonal associations that are available can be of assistance. The extrapersonal knowledge can help them keep in mind the dual meaning of the response keys, and hence, increase the speed with which they can categorize the stimuli in the incompatible block.

A plausible alternative?

There is, however, an alternative way to view the present findings—a possibility that merits consideration. Let us assume both the existence of differential implicit versus explicit representations and the validity of the IAT as an instrument for revealing the nature of these implicit attitudes. With these as inviolate starting points, a proponent would then argue that the boys' comments produced change in the implicit attitudes—change that the traditional IAT, but neither priming nor the personalized IAT, was able to detect. This alternative possibility cannot be denied. However, we would question its plausibility, as well as the scientific utility of such a view of implicit attitudes.

How reasonable is it to assume that the remarks dropped by two boys had no impact on the explicit measure of attitude, choice behavior, the personalized IAT, and the evaluations that were automatically activated when the attitude objects were presented subliminally as primes, yet nevertheless, affected some other representational layer of attitudes that in turn influenced performance on the traditional version of the IAT? What is the nature of this other representation, and by what mechanism is the traditional IAT, but not the personalized IAT or the priming measure, able to unlock its secrets? How ever they might be conceived, are implicit attitudes subject to change in every interaction setting, every time someone voices an opinion to the contrary, even if that opinion advocates an objectively inferior position and is viewed incredulously? If so, then the very idea of implicit attitude becomes a vapid construct, impossible to falsify.

Additional implications

In addition to demonstrating the influence of extrapersonal associations on the traditional version of the IAT, the current experiments also shed some light on the low correspondence that is sometimes observed between the IAT and priming. For instance, it has been repeatedly documented that the IAT shows a greater prejudice towards Blacks by White participants than is apparent with a priming measure (about 80 vs. 50%; e.g., Fazio et al., 1995; Nosek, Banaji, & Greenwald, 2002a). Blacks also do not show an in-group preference on the IAT as they do on the priming measure (e.g., Fazio et al., 1995; Nosek, Banaji, & Greenwald, 2002b). The disparity between the two measures has been found with respect to self-esteem (Bosson, Swann, & Pennebaker, 2000), attitudes towards condom use (Marsh, Johnson, & Scott-Sheldon, 2001), attitudes towards smoking (Sherman et al., 2002), as well as racial attitudes (Olson & Fazio, 2003). However, the lack of correlation between the two measures is not as disturbing if we consider the differential influence of extrapersonal associations on priming and the IAT. Because the IAT is more heavily influenced by extrapersonal associations than the priming measure is, the two need not cohere. Whenever extrapersonal associations are available, these two measures would not be expected to correspond well.

The results of the current research also raise some questions about the conclusions that have been reached concerning the presumed malleability of attitudes, even when measured implicitly (see Blair, 2001; for an in-depth discussion; Fazio & Olson, 2003). The IAT has often been employed as a tool in such research. For example, it has been shown that exposure to counter-stereotypical or counter-attitudinal exemplars (Dasgupta & Greenwald, 2001) or the presence of a Black experimenter reduces the prejudice exhibited towards Blacks on the IAT (Lowery, Hardin, & Sinclair, 2001). Similarly, Blair, Ma, and Lenton (2001) found weaker gender stereotypes after having participants create mental images of counter-stereotypic (strong) women. However, these manipulations may simply have served to enhance the salience of extrapersonal associations. Thus, it is possible that these effects demonstrate, not malleability of attitudes or stereotypes as they are represented in memory per se, but a consequence of the IAT's sensitivity to extrapersonal associations.

Before we conclude, some clarifications are in order. First, it must be emphasized that the personalized IAT employed in the current work and the Olson and Fazio (2004) research does not address other task features that have been viewed as pivotal to IAT performance (e.g., salience asymmetries, Rothermund & Wentura, 2001, 2004; task-set switching, Klauer & Mierke, 2005; representativeness of the exemplars selected for use as stimuli, Govan & Williams, 2004; Mitchell, Nosek, & Banaji, 2003; Steffens et al., 2004). Rather, it appears that the personalized IAT removes the influence of extrapersonal associations, and thus, provides a cleaner measure of one's personal associations than the traditional IAT. As research on the mechanisms that underlie IAT performance accumulates, additional means by which the measure can be improved are likely to come to light.

Second, we are not arguing that extrapersonal associations are irrelevant and do not have the potential to impact behavior. When one's motivation is high or when the situational cues dictate that one behaves in a certain way, extrapersonal associations may predict behavior, perhaps even better than one's personal associations. In fact, many attitude theories, such as the theory of reasoned action (Fishbein & Ajzen, 1975), have long argued that subjective norms can dictate one's volitional behavior. Likewise, the MODE model (Fazio & Towles-Schwen, 1999) maintains that such motivational concerns can counter the influence of the personal evaluation that is automatically activated upon encountering an attitude object. However, the purpose of the current experiments was to show that the traditional version of the IAT is influenced by these extrapersonal associations and hence, does not measure personal associations as purely as might be desired.

References

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

Bargh, J. A. (1999). *The cognitive monster: The case against the controllability of automatic stereotype effects. Dual-process theories in social psychology*. New York: Guilford Press.

Bessenoff, G. R., & Sherman, J. W. (2000). Automatic and controlled components of prejudice toward fat people: evaluation versus stereotype activation. *Social Cognition, 18*, 329–353.

Blair, I. (2001). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242–261.

Blair, I., Ma, J., & Lenton, A. (2001). Imagining stereotypes away: The moderation of automatic stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828–841.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 760–773.

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

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443–451.

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.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Hillsdale, NJ: Erlbaum.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115–141.

Fazio, R. H. (2000). Accessible attitudes as tools for object appraisal: Their costs and benefits. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Function of attitudes* (pp. 1–36). Mahwah, NJ: Earlbaum.

Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: Data reanalysis and commentary on Bargh, Chaiken, Govender, and Pratto (1992). *Journal of Personality and Social Psychology, 64*, 753–758.

Fazio, R. H., Powell, M. C., & Herr, P. M. (1983). Toward a process model of the attitude-behavior relation: Accessing one's attitude upon mere observation of the attitude object. *Journal of Personality and Social Psychology, 44*, 723–735.

Fazio, R. H., Chen, J., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency and the strength of the object-evaluation association. *Journal of Experimental Social Psychology, 18*, 339–357.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (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: Their meaning and use. *Annual Review of Psychology, 54*, 297–327.

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*, 229–238.

Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude–behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 97–116). New York: Guilford.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Govan, C. L., & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal of Experimental Social Psychology, 40*, 357–365.

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.

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., 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*, 197–215.

Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (in press). Validity of the salience asymmetry interpretation of the IAT: Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General*..

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 774–778.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly, 24*, 163–204.

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for task-set switching account for the Implicit Association Test effects. *Personality and Social Psychology Bulletin, 31*, 208–217.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*, 842–855.

McConnell, A. R., & Liebold, J. M. (2001). Relations between the Implicit Association Test, explicit racial attitudes, and discriminatory behavior. *Journal of Experimental Social Psychology, 37*, 435–442.

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

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

Mitchell, J. A., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455–469.

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

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*, 413–417.

Olson, M. A., & Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition, 20*, 89–104.

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

Olson, M. A., & Fazio, R. H. (2004). Reducing the Influence of Extrapersonal Associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653–667.

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 94–106.

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

Sherman, S. J., Presson, C. C., Chassin, L., Rose, J. S., & Koch, K. (2002). Implicit and explicit attitudes toward cigarette smoking: the effects of context and motivation. *Journal of Social and Clinical Psychology, 22*, 13–39.

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and personality*. New York: Wiley.

Steffens, M. C., Lichau, J., Still, Y., Jelenec, P., Anheuser, J., Goergens, N. K., et al. (2004). Individual or group, exemplar or category? A two-factor model for explaining reaction-time differences in the Implicit Association Test (IAT). *Zeitschrift für Psychologie, 212*, 57–65.