

RUNNING HEAD: Detecting Attitude Change with the Implicit Association Test

Detecting Attitude Change with the Implicit Association Test

Sandor Czellar

HEC Paris

Russell H. Fazio

Ohio State University

February 8, 2008

## Abstract

The Implicit Association Test and its variants have become pervasive measures of attitudes in a variety of domains and contexts. In two experiments, we provide evidence that a recent variant, the Personalized IAT developed by Olson and Fazio (2004) may more accurately detect changes in personal attitudes than the conventional Traditional IAT devised by Greenwald, McGhee, and Schwartz (1998). Our findings suggest that the Personalized IAT may be more sensitive to detecting attitude changes than the Traditional IAT because it is less affected by extrapersonal associations (i. e. salient associations not contributing to personal evaluations of the object). More generally, this research suggests that for attitude domains characterized by potentially strong extrapersonal associations, using the Personalized and Traditional IATs may provide researchers with complementary insights about knowledge structures.

Keywords: attitudes, attitude change, implicit attitude measures, IAT

## Detecting Attitude Change with the Implicit Association Test

The Implicit Association Test (IAT) devised by Greenwald, McGhee and Schwartz (1998) is a widely used tool for the study of implicit cognitions in a variety of domains. Considerable empirical evidence has accumulated in recent years about the validity of the IAT as an appropriate measure of automatically-activated attitudes about a given object. Extant research has investigated several validity aspects of the IAT including nomological, convergent, discriminant and predictive validity (Lane, Banaji, Nosek, & Greenwald, 2007). In terms of predictive validity, the IAT has been shown to relate to object-related memory biases, explicit attitude measures and preferences, as well as some forms of behavior (Greenwald, Poehlman, Uhlmann, & Banaji, *in press*; Nosek, 2005).

In the present research, we investigate yet another facet of the IAT's validity—its capacity to capture attitude change. Prior research evidences mixed results about the sensitivity of IAT measures to persuasive messages (for a review, see Briñol, Petty, & McCaslin, *in press*). Some studies show that IAT effects, unlike explicit attitude measures, seem quite resistant to various forms of persuasion (e. g., Gawronski & Strack, 2004; Gregg, Seibt, & Banaji, 2006). However, other research demonstrates that sometimes even subtle manipulations can induce changes in attitudes as reflected by IAT effects (e. g., Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Lowery, Hardin, & Sinclair, 2001). Indeed, Lane et al. (2007) recognized a gap in our understanding of why these discrepancies may precisely occur. The purpose of the present research is to bring some elements of answer to this complex issue from a methodological perspective.

We argue that one of the reasons for the differential sensitivity of IAT measures across domains and contexts is the extent to which extrapersonal associations affect the outcome of a

particular measurement technique implemented. Extrapersonal associations constitute object-related information that is accessible in memory yet does not contribute to an individual's evaluation of an object (Olson & Fazio, 2004). On the basis of Olson and Fazio (2004) and Han, Olson, and Fazio (2006), we propose that for some attitude domains, the traditional IAT developed by Greenwald et al. (1998) is more open to the unwanted influence of extrapersonal associations than the personalized IAT. The Han et al. (2006) findings demonstrate that extrapersonal associations can create the appearance of change on a traditional IAT, even when there is good reason to believe that no such change occurred. For example, a clearly invalid assertion uttered by a 10-year old child did not impact the personalized IAT, a measure of automatically-activated attitudes based on a subliminal evaluative priming paradigm, or choice behavior. Yet, the extrapersonal association offered by the child did affect the traditional IAT.

We believe that it also is possible for extrapersonal associations to influence a traditional IAT in such a way as to obscure the presence of real attitude change. The purpose of the present research is to explore this heretofore unexamined possibility. If actual attitude change is induced, extrapersonal associations may mask the attitude change more in the traditional IAT than in the personalized IAT. For this reason, the personalized IAT should perform better in detecting real attitude changes than the traditional IAT. We present two experiments examining this reasoning.

### The Two IAT Variants and Persuasion Attempts

#### *Features of the Two IAT Variants*

The original IAT developed in Greenwald et al.'s seminal paper (1998) compares response latencies in sorting tasks involving compatible and incompatible combinations of attitude objects (e.g. *flowers vs. insects*) and valence attributes (e.g. *pleasant vs. unpleasant*). It is based on the rationale that if individuals have a positive attitude toward flowers but not insects,

they should more easily (i.e. more quickly) associate pleasant words with flower names than with insect names.

Olson and Fazio (2004) argued that the measurement outcome of the traditional IAT could reflect the combined influence of two types of associations: personal and extrapersonal. Personal associations reflect accessible information that contributes to individual attitudes toward the attitude object; whereas extrapersonal associations correspond to accessible information that does not contribute to individual attitudes. For example, these extrapersonal associations may form on the basis of information acquired through social interaction and from the media and can reflect culturally stereotypic knowledge (Karpinski & Hilton, 2001). Extant research shows that the traditional IAT is contaminated by extrapersonal associations that are accessible at the time of categorizing an attitude object (Olson & Fazio, 2004; Han et al., 2006). For this reason, Olson and Fazio (2004) devised a “personalized” version of the IAT that focuses on personal associations to a greater extent than the traditional IAT. The two most important changes with respect to the traditional IAT involve suppressing error feedback and changing the category labels from *pleasant* to *I like* and from *unpleasant* to *I don't like*. These two changes are intended to focus thoughts on personal likes and dislikes. Error feedback, at least with respect to the trials involving *pleasant/unpleasant* categorizations, encourages a normative focus by indicating that a correct answer, independently of the person's own liking or disliking of the target objects, exists. The labels *pleasant/unpleasant* or *good/bad* allow for such an interpretation, enabling extrapersonal associations to intervene in solving the mapping problem of the IAT. In contrast, the use of the labels *I like/I don't like* precludes such construals because it focuses attention on the personal thoughts and feelings regarding the object.

Across several attitude domains, evidence supports the idea that these modifications to the traditional IAT procedure may indeed reduce the influence of extrapersonal associations on the

IAT effect. In Olson and Fazio's (2004) experiments 1-2, the personalized IAT exhibited less racial bias than the traditional IAT, confirming the assumption that negative extrapersonal associations about Black people may reflect negative culturally-shared stereotypes. In experiment 3, Olson and Fazio (2004) found the traditional IAT to reflect greater preference for apples over candy bars than the personalized IAT, suggesting that positive extrapersonal associations about apples may inflate the traditional IAT effect. Han et al. (2006) experimentally created both attitudes and extrapersonal associations about game characters and showed that both attitudes and extrapersonal associations affected the traditional IAT while only attitudes affected the personalized IAT. In all these studies, differences between the two IAT variants were apparent with both the conventional (Greenwald et al., 1998) and D-scoring procedures (Greenwald, Nosek, & Banaji, 2003) for the calculation of the IAT effects.

#### *Predicting Attitude Changes with the Two IAT Versions*

Extant research suggests that persuasive messages can exert widely different effects on attitude measures captured by the IAT. Recently, Briñol et al. (*in press*) provided a comprehensive review of that research and proposed a framework for the possible mechanisms underlying attitude change captured by the IAT. We build on that framework to develop propositions about the processes underlying the differential sensitivity of the two IAT variants to persuasion attempts.

Briñol et al. argue that detection of attitude changes with implicit techniques depends on the level of elaboration involved in the persuasion process (ranging from fully automatic to fully deliberative processes). Existing literature is quite equivocal about the conclusion that low-thought processes can alter, at least temporarily, the types of automatic associations linked to the object. In this respect, the IAT has been shown to be sensitive to persuasion techniques like classical conditioning and mere exposure (e. g., Dasgupta & Greenwald, 2001; Lowery et al.,

2001; Petty, Tormala, Briñol, & Jarvis, 2006). Under high elaboration conditions, the picture is somewhat different. Recently, Brinol et al. proposed a compelling explanation for the mechanism underlying implicit attitude change through deliberative processes. When a message is processed in detail, each positive or negative argument is carefully assessed in terms of its validity with respect to one's own positive or negative attitude toward the object. If the message presents strong convincing arguments, this thoughtful process should lead to attitude rehearsal, whereby the positive or negative aspects of one's attitude become more accessible regarding the object. The effect of this attitude rehearsal should be apparent in the IAT, by facilitating the completion of either its compatible or incompatible task, whichever response mapping more closely corresponds to the individual's attitude. However, if the message presents weak arguments, argument-consistent attitude rehearsal is less likely to occur because individuals would easily dismiss or counterargue the message. Thus, a weak message is expected to produce less attitude change on the IAT than a strong message under high elaboration conditions. Brinol et al. provide several experiments that support this theorizing.

In this research, just like Briñol et al., we focus our attention on persuasion under high elaboration conditions. Because research to date has mostly used the traditional IAT, we extend Briñol et al.'s reasoning by considering the potentially differential sensitivity of IAT type (traditional vs. personalized) on the measurement outcome. We argue that a persuasive message under high elaboration can exert differential effects on the two IAT variants depending on message strength. On the one hand, if the persuasion is sufficiently strong, it is likely to shape personal attitudes through attitude rehearsal, whereby message-consistent personal associations will become more accessible. However, depending on the attitude domain, the traditional IAT may or may not fully reflect these real changes in personal attitudes because extrapersonal associations can potentially interfere with its measurement outcome. For example, in domains

where negative/positive cultural stereotypes are prevalent, such as attitudes toward specific social groups or food preferences, attitude-inconsistent extrapersonal associations may be highly salient whenever one evaluates those groups. In such situations, even if a persuasion attempt successfully manages to change personal attitudes, the presence of extrapersonal associations may still mask the full extent of attitude change on the traditional IAT. Because of its potentially greater sensitivity to personal associations, we expect the measurement outcome to be more sensitive to strong persuasion attempts in the personalized IAT than in the traditional IAT.

To illustrate our reasoning, consider the example of a gourmet person who is fond of chocolate, mainly for hedonic reasons. His positive attitude toward chocolate thus stems from the positive personal associations about the taste of this delicacy. He is of course conscious that chocolate is a high calorie food; however, let's assume that this negative association does not contribute to his overall positive attitude about chocolate and therefore remains an extrapersonal association. Because the traditional IAT is affected by this extrapersonal association, it is likely that it would yield a less positive attitude estimate for this person than the personalized IAT would. If this person were to read a strong positive message about some additional benefits of chocolate (e.g., contains antioxidants), his attitude will probably become even more positive to chocolate than before. However, the negative extrapersonal associations can partially mask this positive attitude change on the traditional, but not on the personalized IAT. We thus expect that a strong persuasive message will affect the measurement outcome more in the personalized IAT than in the traditional IAT.

On the other hand, if the persuasion is weak, no real attitude change occurs. Such messages can probably not affect personal associations but can, at least in the short term, remain salient as an extrapersonal association about the object. To return to our previous example, if our gourmet person hears a friend telling him a rather unbelievable story about chocolate causing



choking hazards, this message may, at best, become a negative extrapersonal association. Due to its sensitivity to extrapersonal associations, the traditional IAT can indicate some change in such situations; the personalized IAT is less likely to do so. Extant research seems supportive of this possibility. For example, Han et al. (2006) showed that traditional (but not personalized) IAT measures were indeed affected by an unjustified, objectively inaccurate statement of preference on the part of two young children – in effect, by a weak counterattitudinal message. We thus expect that a weak persuasive message will affect the measurement outcome more in the traditional than in the personalized IAT.

### Experiment 1

The purpose of our first experiment was to investigate the potentially greater sensitivity of the personalized (vs. traditional) IAT in detecting attitude change after a strong persuasion attempt in a domain where extrapersonal associations are prevalent. Our choice fell on the domain of luxuries because luxury products and brands are characterized by a great variety of potential associations, including quality-related beliefs, hedonic benefits and prestige connotations. It is likely that between-person discrepancies characterize the valence of those associations (e. g., the high price of luxuries may be a positive aspect for one person and a negative aspect for another person). In addition, given the substantial social visibility and social function of those items, the evaluation of luxury items is subject to attitude reports where social norms, as well as socially desirable responding could be prevalent. That is, people may easily associate luxuries with extrapersonal information they do not necessarily endorse personally. For example, a person with a favorable attitude to BMW cars may evoke a variety of thoughts about the BMW brand, including its sporty engines, comfortable interior, high reliability, that is, thoughts that contribute to his or her positive attitude toward the brand (i. e., personal associations). At the same time, this person may acknowledge that some people find BMW to be

overpriced and snobbish (extrapersonal associations). These latter two associations may be salient at the time the attitude measure is taken, yet they may not affect this person's globally positive attitude toward BMW.

### *Method*

*Materials.* The experiments were conducted in the Paris area in France. Experimental materials were developed on the basis of 16 interviews with participants aged between 17 and 70. Participants were asked to remember and describe their recent acquisitions of, or experiences with luxury items. They were then asked to describe their views about luxuries in general and the various product/service categories and brands associated with them. We adapted characteristic sections from those interviews to design various persuasive messages presenting luxuries in either a negative or positive light (see the Appendix for the complete set of selected materials).

*Participants and procedure.* Two-hundred and nineteen undergraduate students participated as part of a course requirement. The experiment was divided into two sessions, with a two-week delay between them. In session 1, participants were asked to write down their "thoughts coming to mind when thinking about luxury brands." Boxes were provided for each of the thoughts and participants were asked to use up as many boxes as necessary. After this, they were asked to code each of the thoughts according to their valence (negative/positive/neutral). Session 2 was a 2 (message arguments: negative, positive)  $\times$  2 (IAT type: traditional, personalized) between-subjects design. Participants were told that the purpose of the study was to investigate "how people form impressions of other people on the basis of interview excerpts." Participants were randomly assigned to the negative and positive arguments conditions, the full details of which are presented in the Appendix. The negative arguments emphasized the snobbery that characterizes users of luxury brands, whereas the positive arguments focused on considerations of exceptional quality. Participants read an introductory sentence explaining that

the excerpts came from an interview with “Eric who is 50 and lives in Paris.” Participants were then presented with the interview excerpts and were asked to read them carefully so as to be able to form a first impression of the source. After reading the interview materials, participants responded to the following items on a 1 (strongly disagree) to 7 (strongly agree) scale: “This person thinks that luxury brands mean snobbery”(reverse-coded), “This person thinks that luxury brands are well regarded in society”, and “Overall, this person thinks that luxuries are a good thing”. These items were averaged to serve as checks for the message valence manipulation. Participants also responded to three items intended to capture their first impressions of the source (unpleasant/pleasant, not sympathetic/sympathetic, negative/positive), all measured on seven-point scales. This latter scale was intended to measure argument strength, with the expectation that participants would form more favorable impressions of people presenting more (vs. less) credible arguments. After this, participants were randomly assigned to the traditional or personalized IAT, implemented on Inquisit 2.0 software.

Both IATs featured the target concepts *luxury brands vs. common brands*. The representative exemplars for the concept of *luxury brands* were selected on the basis of a pretest and featured prestigious international brands in the couture and car industries including *Chanel, Vuitton, Dior, Gucci, Mercedes* and *Ferrari*. For common brands, the most frequently cited in the pretest and, hence, selected for use were *Zara* (clothing brand), *H&M* (clothing brand), *Ariel* (detergent), *Monoprix* (grocery retailer), *Bic* (office furniture) and *Carrefour* (retailer). Valence attributes included *peace, paradise, joy, love, pleasure* and *happiness* for positive attributes and *disaster, grief, accident, pain, bad* and *agony* for negative attributes. In the traditional IAT we employed Greenwald et al.’s (1998) procedure. In the personalized IAT we followed one of the variants developed by Olson and Fazio (2004). The test implemented differed from the traditional IAT in two important respects. First, no error feedback was given to participants. Second, the

category labels were changed from *pleasant* to *I like* and from *unpleasant* to *I don't like*. Thus, the compatible block of the personalized IAT consisted of combined tasks with the labels *luxury brands* and *I like* and *common brands* and *I don't like*; the incompatible block featured the labels *luxury brands* and *I don't like* and *common brands* and *I like*. The purpose of these changes was to limit the potentially strong normative effects on the IAT, thus reducing the relative influence of extrapersonal associations and increasing the relative influence of personal associations on the outcome of the measure.

### *Results*

*Descriptive statistics in Session 1.* On average, 47 % of the thoughts participants wrote were positive, 30.9% were negative and 21.1% were neutral, reflecting a globally rather positive view of luxury brands in the sample.

*Manipulation checks in Session 2.* The message valence manipulation was successful because compared to the negative arguments condition, participants in the positive arguments condition found the source's thoughts about luxuries to be less snobbish ( $M = 5.61$  vs.  $M = 1.57$ ,  $F(1, 217) = 435.42$ ,  $p < 10^{-6}$ ), better regarded in society ( $M = 5.34$  vs.  $M = 1.42$ ,  $F(1, 217) = 788.80$ ,  $p < 10^{-6}$ ) and globally good things ( $M = 5.80$  vs.  $M = 1.63$ ,  $F(1, 217) = 945.49$ ,  $p < 10^{-6}$ ). Compared to the negative arguments condition, participants in the positive arguments condition also formed better impressions of the source in terms of pleasantness ( $M = 4.28$  vs.  $M = 3.43$ ,  $F(1, 217) = 20.36$ ,  $p = 10^{-5}$ ), sympathy ( $M = 4.30$  vs.  $M = 3.45$ ,  $F(1, 217) = 21.64$ ,  $p < 10^{-5}$ ) and positiveness ( $M = 4.26$  vs.  $M = 3.44$ ,  $F(1, 217) = 20.46$ ,  $p < 10^{-5}$ ). This latter result, combined with the fact that our respondent held rather positive views of luxuries, suggests that IAT scores should be more favorable toward luxury brands in the positive (vs. negative) arguments condition. We are interested in knowing which of the IAT versions, traditional or personalized, prove more sensitive to this exposure to the differently valenced arguments.

*Descriptive IAT statistics.* D-measures for the IAT effect were calculated following the scoring algorithm developed by Greenwald et al. (2003). Mean error rates in both IATs were low (3.9% in the traditional IAT and 5.4% in the personalized IAT). Corroborating session 1 findings, participants across experimental conditions revealed a strong preference for *luxury brands* (vs. *common brands*) on both IATs (traditional:  $M = .38$ ,  $t(112) = 8.96$ ,  $p < 10^{-6}$ ; personalized:  $M = .48$ ,  $t(106) = 11.18$ ,  $p < 10^{-6}$ ).

*Evidence of attitude change with the two IAT versions.* A  $2 \times 2$  ANOVA was conducted with IAT effect as a dependent variable and message arguments and IAT type as factors. A marginally significant effect emerged for IAT type,  $F(1, 215) = 2.82$ ,  $p = 0.09$ , qualified by a significant two-way interaction between message arguments and IAT type,  $F(1, 215) = 3.77$ ,  $p = 0.05$ . The IAT effect as a function of message arguments and IAT type are represented in Figure 1. Simple contrasts showed that in the personalized IAT, the IAT effect was more positive after the positive arguments ( $M = .59$ ) than after the negative arguments ( $M = .38$ ),  $F(1, 215) = 5.76$ ,  $p = .017$ ; however, such a difference was not observed in the traditional IAT ( $M = .40$  vs.  $M = .37$ ,  $F(1, 215) = 5.76$ ,  $p = .746$ ). These results suggest that if attitude change occurred as a result of the positive arguments, this change seems to have been detected with the personalized but not the traditional IAT.

We also investigated how IAT scores were affected by impressions about the source across the experimental conditions. If attitude change indeed occurred and if it was only captured by the personalized IAT, then IAT scores should be related to impression about the source in the personalized but not in the traditional IAT. By conducting multiple linear regressions using Aiken and West's methods (1991), the following findings emerged. In the personalized IAT, IAT scores were regressed on impression about source (average of three items), message arguments (0 = negative, 1 = positive), and the impression about source  $\times$  message arguments cross-product.

This regression resulted in a negative coefficient for impression about interviewee ( $t(106) = -2.62, p = .01$ ), a positive coefficient for experimental condition ( $t(106) = 3.27, p = .001$ ) and a positive coefficient for the interaction term ( $t(106) = 2.49, p = .01$ ),  $R^2 = .14, F(3, 104) = 5.74, p < .01$ . To probe the significant interaction, the slopes of regression lines for the negative and positive experimental conditions were tested. In the negative arguments condition, as expected, a negative relation between impression about the source and IAT scores was observed,  $t = -2.61, p = .01$ ; the more favorably the interviewee was viewed, the less positive the IAT score. In the positive arguments condition, again as expected, a positive relation between impression about the source and IAT scores was observed,  $t = 9.49, p < .01$ ; more favorable views of the interviewee were associated with more positive IAT scores. A similar regression analysis conducted on the traditional IAT scores yielded non-significant estimates,  $R^2 = .002, F(3, 106) = .09, p = .97$ .

To provide further evidence of attitude change being captured more by the personalized IAT than the traditional IAT, we investigated the relationships between thought measures in session 1 with IAT effects in session 2. To do so, we first created a thought listing index reflecting pre-exposure attitudes by subtracting, for each participant, the percentage of positive thoughts from the percentage of negative thoughts. If the traditional IAT is less affected by attitude change than the personalized IAT, then traditional IAT effects should be more correlated with the pre-exposure thought index than is true for the personalized IAT scores. This seems to be indeed the case. With the traditional IAT, IAT scores in session 2 correlated positively with the thought index in session 1 ( $r = .244, p = .011$ ). With the personalized IAT, IAT scores in session 2 did not correlate significantly with the thought index in session 1 ( $r = .092, p = .372$ ). Although the difference between the two correlations was only marginally significant ( $z = 1.69, p = .091$ ), the directional results provide additional evidence for the likelihood of attitude change being captured more by the personalized than by the traditional IAT.

*Discussion*

The goal of this experiment was to investigate whether the personalized IAT would be more sensitive in detecting attitude changes than the traditional IAT. To do so, we developed negative and positive persuasion messages intended to produce changes in how people view luxury brands. Our preliminary analyses showed that the positive arguments looked more conducive to produce attitude change than the negative arguments. Indeed, when negative arguments were presented about luxuries, we observed no differences between the traditional or personalized IAT effects. However, when positive arguments were presented about luxuries, the personalized IAT showed a positive change but the traditional IAT did not. The relationships established between pre-exposure thought measures and IAT effects also supported the differential sensitivity of the two IAT variants to the induced positive attitude change. A tentative explanation for this latter finding could be the possibility that for some people, certain negative extrapersonal associations (e. g., the snobbishness of luxuries) remained activated after reading strong positive arguments about the quality of luxury brands. If so, then the salience of those extrapersonal associations could have obscured the detection of attitude change with the traditional IAT but not with the personalized IAT. Our second experiment was designed to more directly investigate this possibility. Also, our interpretation of the findings was based on the assumption that participants perceived the messages presented to them as strong. In Experiment 2, we manipulate the messages to create persuasion attempts of varying strength.

## Experiment 2

The attitude objects considered in this experiment were again luxury brands. We focused only on positive arguments because Experiment 1 suggested that this is where a difference between the traditional and personalized IAT effects occurred in the present context. We hypothesized that negative extrapersonal associations could be responsible for the traditional IAT

producing lower effects than the personalized IAT after a positive persuasion attempt. Our explanation was based on the likelihood of extrapersonal associations interfering with the traditional IAT, thereby limiting its capacity to capture real changes in personal attitudes. However, our first experiment did not provide direct evidence for extrapersonal associations being responsible for the effects observed. In our second experiment, we changed the experimental design and procedure so as to be able to more directly identify specific extrapersonal associations obscuring the detection of attitude change with the traditional IAT.

The design included a first condition with no persuasive attempt at all so as to assess the relative extent of attitude change observed in the experimental conditions. Experiment 1 relied upon a comparison of conditions in which positive or negative arguments had been presented. In contrast, Experiment 2 employs a baseline control condition. This will enable an examination of how scores on the traditional and personalized IATs differ in the absence of a persuasive communication.

In Experiment 1, a strong positive message appeared to induce positive attitude change to which the personalized IAT proved sensitive but the traditional did not. In the second condition of Experiment 2, we employed this same strong positive message, but sought to introduce contexts that would make negative extrapersonal associations more or less salient. Our argument is that the natural salience of negative extrapersonal associations led to decreased evidence of attitude change on the traditional IAT. To lessen the potential influence of such information, we introduced the strong positive argument with a clear statement that the particular interview had been selected because it was very representative of the collective opinion expressed within the interviewed sample. We expected both IAT variants to detect positive attitude change in this condition.



Experiment 2 also included a third condition in which a negative extrapersonal association was made very salient immediately after the presentation of the strong positive arguments. In this condition, a brief interview with another person, explicitly labelled as unrepresentative of the interviewed sample, introduced mention of a negative comment about the snobbishness of luxuries. We expected this weak negative comment to increase the salience of a negative extrapersonal association; however, we did not expect this comment to undermine the positive attitude change induced by the strong positive arguments. Thus, we expected the traditional IAT to be affected by this negative comment and evidence less positive attitude change than in the condition in which only the positive argument was presented. On the other hand, we expected the personalized IAT to remain impervious to the negative comment and evidence the same positive attitude change as in the positive argument alone condition.

### *Method*

*Materials.* Materials in the positive arguments condition were similar to those of Experiment 1, with the only difference being a brief sentence introducing the arguments as representative of the global view of the sample on luxuries. We relied on the pretest of the previous experiment to develop experimental materials involving a negative argument. The focal topic of this short negative argument, presented as unrepresentative of most interviewees, was the apparent snobbishness of luxury brands. We expected the argument to be too weak to induce attitude change (i. e., we did not expect it to affect personal associations about the object). However, we expected that this negative comment would provide people with a negative extrapersonal association that would be salient when luxury brands were later categorized in the traditional version of the IAT.

*Participants and procedure.* Two-hundred and two undergraduate students participated in a 3 (message arguments: no arguments; positive arguments; positive arguments followed by a

negative argument)  $\times$  2 (IAT type: traditional, personalized) between-subjects design.

Participants were randomly assigned to the experimental conditions. Participants in the no arguments condition only completed either the traditional or personalized IAT about luxury brands vs. common brands. In the other two conditions, participants were told that the present research investigated how people formed impressions about unknown people on the basis of various interview excerpts. Participants in the positive arguments condition read the same positive interview excerpts used in Experiment 1 about the 50-year old Eric, preceded by the information that his opinions were representative of what most interviewed people thought (see Appendix). After reading the entire set of interview excerpts, they responded to a series of manipulation check questions comprising the ones about message valence used in Experiment 1. We also included three additional items measured on 1-7 scales intended to check the message strength manipulation (not interesting/interesting, weak/strong, not convincing/convincing). After this, participants completed either the traditional or personalized IAT used in Experiment 1. Participants in the positive arguments + negative argument condition followed a similar procedure, with the only difference that after reading and evaluating the representative positive arguments presented by Eric, they also read a short interview excerpt from another person, Michel, 35, commenting on the snobbery of luxury brands (see Appendix). Michel's comment was presented as unrepresentative of the sampled people. Identical manipulation checks were taken after this interview excerpt, too.

### *Results*

*Manipulation checks.* In the condition in which positive arguments were followed by the brief negative comment of another interviewee, participants evaluated the target's attitude toward luxuries as more positive after reading the positive ( $M = 5.34$ ) than the negative arguments ( $M = 2.37$ ),  $t(71) = 19.74$ ,  $p < 10^{-6}$ . They also evaluated the positive (vs. negative) arguments as more

interesting ( $M = 4.54$  vs.  $M = 3.57$ ,  $t(71) = 4.08$ ,  $p < .0001$ ), stronger ( $M = 3.67$  vs.  $M = 2.93$ ,  $t(71) = 3.41$ ,  $p < .001$ ) and more convincing ( $M = 3.90$  vs.  $M = 3.17$ ,  $t(71) = 3.03$ ,  $p = .003$ ).

These results suggest that the manipulation of message valence and strength occurred as intended.

*Descriptive IAT statistics.* D-measures for the IAT effect were calculated following the scoring algorithm developed by Greenwald et al. (2003). Mean error rates in both IATs were low (4% in the traditional IAT and 4.74% in the personalized IAT). Corroborating Experiment 1 findings, participants across experimental conditions revealed a strong preference for *luxury brands* (vs. *common brands*) on both IATs (traditional:  $M = .35$ ,  $t(102) = 8.76$ ,  $p < 10^{-6}$ ; personalized:  $M = .51$ ,  $t(99) = 12.50$ ,  $p < 10^{-6}$ ).

*Evidence of attitude change with the two IAT versions.* A  $3 \times 2$  ANOVA was conducted with IAT effect as a dependent variable and message arguments and IAT type as factors. The IAT effect as a function of message arguments and IAT type are presented in Figure 2. Significant main effects for both message arguments ( $F(2, 196) = 7.40$ ,  $p = 0.001$ ) and IAT type ( $F(1, 196) = 7.92$ ,  $p = 0.005$ ) emerged. Those main effects were qualified by a significant two-way interaction between message arguments and IAT type,  $F(2, 196) = 3.40$ ,  $p = 0.035$ . Simple contrasts showed that in absence of persuasion (i.e., in the no arguments condition) participants completing the personalized IAT exhibited more positive attitudes toward luxuries ( $M = .39$ ) than participants completing the traditional IAT ( $M = .21$ ),  $F(1, 196) = 3.95$ ,  $p = 0.048$ . This result is consistent with our hypothesis about possible negative extrapersonal associations affecting the traditional IAT, which has the likely consequence of producing a less positive baseline attitude score on the traditional IAT than the personalized.

Let us now examine how each of the IATs varies across the two conditions in which arguments were presented. Simple contrasts evidence that when only positive arguments are

presented, the traditional IAT reflects a positive attitude change ( $M = .58$ ) compared to the baseline ( $M = .21$ ),  $F(1, 196) = 14.89$ ,  $p = 0.0001$ . However, if the positive arguments are immediately followed by a negative argument, the traditional IAT measure is substantially reduced ( $M = .28$ ) compared to the positive arguments only condition ( $M = .58$ ),  $F(1, 196) = 10.27$ ,  $p = 0.002$ . Indeed, if the positive arguments are immediately followed by a negative argument, no attitude change is detected with the traditional IAT ( $M = .28$ ) compared to the baseline ( $M = .21$ ),  $F(1, 196) = .54$ ,  $p = 0.461$ . The pattern of results is different with the personalized IAT. When only positive arguments are presented, the personalized IAT reflects a positive attitude change ( $M = .54$ ) compared to the baseline ( $M = .39$ ),  $F(1, 196) = 2.77$ ,  $p = 0.09$ . Likewise, if the positive arguments are immediately followed by a negative argument, a positive attitude change is still detected ( $M = .59$ ) compared to the baseline ( $M = .39$ ),  $F(1, 196) = 4.08$ ,  $p = 0.048$ . These results support our prediction that the negative extrapersonal association activated through a weak message may mask an attitude change if the traditional IAT, but not if the personalized IAT, is used.

### *Discussion*

We hypothesized that strong positive arguments would induce changes in personal associations, leading to a globally positive attitude change. A relatively weak negative argument would not produce such an effect but would rather heighten the salience of negative extrapersonal associations about the snobbishness of luxury brands. The pattern of results observed in Experiment 2 supports these expectations. A strong positive message likely induced attitude change and this change was captured by both IATs. However, if this attitude change was followed by a weak counter message, the traditional IAT failed to evidence the original attitude change anymore and looked heavily affected by that message. Not so the personalized IAT, which exhibited a measurement outcome still reflecting attitude change. From the perspective of

the framework adopted in this paper, this pattern of results further evidences that the traditional IAT is highly sensitive to extrapersonal associations that are salient at the time the attitude assessment is made.

### General Discussion

Extant research on the IAT's validity has provided mixed evidence about the sensitivity of IAT measures to various persuasion attempts. In this research, we attempted to provide one possible explanation for these discrepancies from a methodological perspective. By relying on Olson and Fazio's typology of personal vs. extrapersonal associations, we proposed that the traditional IAT implemented in most prior research can sometimes fail to capture the full extent of attitude change in domains where extrapersonal associations are prevalent. The two experiments presented support the view that the traditional and personalized IATs are differentially sensitive to persuasion in such attitude domains.

The present findings are all the more meaningful when considered in conjunction with the Han et al. (2006) research. That work demonstrated that the traditional IAT can be sensitive to a rather trivial, unjustified and objectively incorrect remark offered by two 10-year old children. The remark did not produce attitude change as evidenced by its failure to affect either the personalized IAT or a priming measure, but it did serve as an attitudinally inconsistent extrapersonal association. As a result, the traditional IAT scores suggested that change had occurred in response to the boys' comment, when other measures, including assessments of the boys, suggested that it had not. The present research illustrates what might be considered the flip side: attitudes *did change* as a function of a strong positive message, and that change was consistently evident on the personalized IAT. However, in Experiment 1, the traditional IAT failed to reveal such change. In Experiment 2, a brief comment about the snobbishness of luxury brands, by an interviewee whose opinion was explicitly labelled as unrepresentative, was

sufficient to render the traditional IAT insensitive to the attitude change produced by the strong positive arguments. Only when that positive message had been identified as representative of all the interviewees sampled and not followed by the presumably rare comment to the contrary did the traditional IAT reveal the change that was apparent on the personalized IAT. Thus, due to its susceptibility to extrapersonal associations, the traditional can suggest change where none has occurred, as in Han et al. (2006), or fail to reveal change that has occurred, as in the present work.

Of course, further research should implement the two IAT variants across a variety of domains to experimentally ascertain our proposition that only in domains where extrapersonal associations are prevalent does a difference between the two IATs occur. We suspect, and the current data certainly suggest, that luxury brands are understood to be associated with snobbishness, but that for many people this understanding is irrelevant to their positive view of the quality of such brands and their liking for them. It is for this reason that the traditional and personalized versions of the IAT diverged. But, a very different pattern might emerge for attitude objects characterized by attributes that are homogenous in their evaluative implications. The likelihood of extrapersonal associations would be minimal in such cases. It is when objects involve a variety of attributes with mixed evaluative implications, some of which might form the basis of an individual's attitude and some of which might have been deemed irrelevant or even rejected, that extrapersonal knowledge exists.

In a broader perspective, implicit cognition research may benefit from a deeper investigation of the nature of associations captured by different implicit measures. Such an enterprise is not without difficulties, however. Indeed, Gawronski and Bodenhausen (2006) propose that one possible perspective on the distinction between personal and extrapersonal associations implies that the source of the association (e. g., personal experience or cultural

knowledge) must be an essential part of the association. They argue however, that this implication is somewhat far-stretched on the basis of extant research indicating that source and information are stored independently in memory (Johnson, Hashtroudi, & Lindsay, 1993; Kumkale & Albarracin, 2004). We do agree with Gawronski and Bodenhausen (2006) that the personal/extrapersonal distinction would not have those unlikely implications, and for this reason, does not provide explanations about the origins of a particular association. Olson and Fazio's (2004) typology of personal-extrapersonal associations merely implies that some accessible associations do not contribute to the individual's own evaluation of the object. The typology actually does not theorize either about the exact source of associations or the availability of associations to introspection. Indeed, extrapersonal associations can be very different from one person to another; those associations may originate from a variety of sources including cultural stereotypes, media information, friend's opinion, etc.; and people may or may not be aware of the exact source of a particular association. In addition, as individual attitudes form and evolve over time, the status of an association can change from personal to extrapersonal and vice-versa, depending on its chronically stronger or weaker link to the attitude node. Thus, determining the origins of personal/extrapersonal associations is a complex issue in domains people interact with on an everyday basis.

Despite the complexity of the distinction between personal and extrapersonal associations, its relevance to the IAT is further substantiated by the present findings. As a result of the influence of an extrapersonal association, the IAT as traditionally implemented can obscure the presence of attitude change or imply that change has occurred when it has not. Thus, if one's interest is in assessing people's current attitudes toward a category of objects, there appears to be considerable value to personalizing the IAT.

## Appendix

## Materials in Experiment 1

*Negative Arguments*

Please carefully read the following excerpts from an interview with Eric, 50, living in Paris:

“In the super luxury boutiques, I wouldn’t feel at ease there, it’s not my environment. I don’t know why I associate it with the desire to appear to be something, something that’s hollow. I don’t wish to identify with people who live like this, no, not at all. In posh neighborhoods, there are some chic ladies and I find them unsympathetic. It’s negative, show-off, these people are not aware of the value of their money. If you systematically go to upmarket shops, that’s a proof of snobbism”

“Luxuries are a little negative to me, this word has a negative side, meaning show-off. There are lots of people who are very conscious about the image luxury brands can give them. Well, you can see it from afar and it’s always the same thing, those people are “nouveaux riches”. Those people don’t know how to truly enjoy life. Normally, if you are OK, you don’t need to open your jacket to exhibit where it comes from. Only people who want to impress others open their jacket to say; “well, have you seen my jacket, have you seen my shoes, these are Weston shoes ...” It’s just stupid. They do it because they want to belong to a social class, they want to say “You see, I belong to high society, I’m an important guy because I have an Yves Saint-Laurent suit. This is all about show-off ”

“It would almost amount to provocation to drive a Jaguar, compared to people who don’t have the means to buy a Jaguar. There are people who have a 500 m<sup>2</sup> house and 5 Mercedes, this is provocation because you can’t forget about the world that lives around you. For me, Mercedes is the car for parvenu people if you like. You have a Mercedes, you have a status, that’s what it means”

“To only live in luxury and to be constrained to choose only big name things, big brands, because you think other things are not worthy of you, well, that must be hell. I don’t envy those people because all the things that relate to luxury, well, the social side of it is related to it as well. Plus a certain kind of people. If you only have luxury things, this means that you only have luxury friends too. And if someone feels constrained to say, I only like luxuries, if this person only chooses something because it’s Chanel, well, that’s stupid.”

“Hermes scarves and Vuitton handbags are ostentatious luxuries. Really, a woman with a Hermes scarf and a Vuitton bag. This is ultimate superficiality ... This kind of thing associated with luxuries is strongly negative for me. Luxury is also big cars. Luxury in the sense of showing off your wealth, strength or success, that is a big engine, a huge car etc... this type of thing has no hold on me. What I said about Hermes and Vuitton is quite characteristic of what I think of luxury brands. Mercedes for example, really, someone who has a Mercedes and there are some I know of, those people are ridiculous ... These are negative connotations, because all this is a matter of external aspects”



“Regarding glasses, I don’t care about brands. For sure I would avoid buying a frame with the name Dior, St-Laurent, Nina Ricci or whatever written on it ... written in big letters, I refuse to do that. Someone who is fond of upscale things would be someone who cares about brands in the first place, that is, buys a brand because it’s a famous brand ...Someone who goes to a hotel because it’s a five-star hotel, who will buy Mercedes. So someone who will only care for the brand’s reputation to show off. It’s a little bit about showing off. Dior, for example, it’s about couture, fragrance, it’s about international show-off. People buy Dior because it’s Dior. Walking around with a Dior bag, no, I’d find it totally ridiculous, I’d certainly avoid doing that.”

### *Positive Arguments*

Please carefully read the following excerpts from an interview with Eric, 50, living in Paris:

“For some things, I will take luxury. If it’s for a gift, you have to buy a name brand. I can buy a prestige champagne if it’s necessary, for very memorable events or as a gift. If I’m about to present a gift to someone, well, I will not buy an unknown brand. To simplify the purchase, I will go to a great department store, I’d for example only go to the Bon Marché in Paris”

“I have memories of exquisite, mysterious hotels, giving the impression of entering a cozy, closed world, and I’m very sensitive to the charm of that kind of hotel. If we take the Mamounia as an example, I was 15 when I first came into contact there with luxury hotels, the thing that really impressed me was that it was a hotel for heads of state. This was where Churchill had spent much of his time during the war. And in the end, what people enjoy about luxury hotels, it’s being on the other side of the wall. Finally, passing to the other side of the wall is essential. At the San Regis hotel in Paris, a coffee costs 10 euros but you feel truly happy, you don’t regret those 10 euros”

“Now if I go out for dinner with a friend, really to go out, in this case it will depend on the person, on what he or she will like But I may say, Well, why don’t we go to the Tour d’Argent or something like that ... If I know they would appreciate, then why not !”

“On an evening out, or an a day of invitation or reception, or something like that, I think you have to wear a certain attire ... If we go to a gala dinner or something like that, banquets by friends, if we see someone who is not dressed according to our criteria, we say to each other, this one, well, this person could have made an effort to dress properly. If you wear luxury clothes, well-designed ones, you don’t have to pay attention to a certain number of things. For me, luxury means basically taking a cab instead of taking the underground, eating out in restaurants often, staying in great hotels when I’m on a trip”

“Regarding great wines, we have some fine bottles but we only open them if we invite connoisseur friends that we want to make a special honor to. I sometimes drink fine wines with friends who are really knowledgeable about them. We sometimes order a thing like that in a restaurant, it may happen. If you are with people who share the same preferences as you or if you are celebrating an event, then you can make an even bigger effort, you can buy something very very good. To drink a bottle of good wine alone or with people who don’t care makes no sense.

It's something rare, something that's exceptional. So something that needs exceptional moments with people who appreciate that. It is for quality guests only, in the sense of people who will like it. I can offer wine as a gift, I sometimes take a bottle or two out of the cellar to offer it to people who will really like them"

"If I buy cognac for my guests, in this case I will not buy the average quality, I will buy something better, of superior quality, that is, VSOP, something better than average. I have already given great cognacs to some people as a gift. I may choose it depending on the packaging, on the case ... It's true that visual appearance counts a lot for the person who receives it ... and also for one who is buying it. It must give pleasure to the recipient if it's nicely presented ... The way it looks like counts a lot"

## Materials in Experiment 2

### *Positive Arguments*

Please carefully read the following excerpts from an interview with Eric, 50, living in Paris. His thoughts are actually quite representative of how most people think in the sample we interviewed.

Remainder of materials identical to Experiment 1

### *Positive Arguments Followed by a Negative Argument*

Same as above, followed by:

Please carefully read the following excerpts from another interview with Michel, 35, who also lives in Paris. His thoughts are not really representative of how most people think in the sample we interviewed.

"I don't know, maybe I would avoid buying a frame with the name Dior, St-Laurent, Nina Ricci or whatever written on it ... I'm not very familiar with those brands. Someone who is fond of luxury things would be someone who cares about brands in the first place, that is, buys a brand because it's a famous brand ... Someone who goes to a hotel because it's a five-star hotel, who will buy Mercedes. Dior, for example, it's about couture, fragrance, it's a little bit about show-off"

## References

- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: the moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828-841.
- Briñol, P., Petty, R. E., & McCaslin, M. J. (in press). Changing attitudes on implicit versus explicit measures: what is the difference? In R. E. Petty, R. H. Fazio & P. Briñol (Eds.), *Attitudes: insights from the new implicit measures*. New York: Psychology Press.
- Fazio, R. H., Eiser, R. J., & Shook, N. J. (2004). Attitude formation through exploration: valence asymmetries. *Journal of Personality and Social Psychology, 87*, 293-311.
- 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.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692-731.
- Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40*, 535-542.
- 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., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: 1. an improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.

- Greenwald, A. G., Poehlman, A. T., Uhlmann, E. L., & Banaji, M. R. (in press). Understanding and using the Implicit Association Test: III. meta-analysis of predictive validity. *Journal of Personality and Social Psychology*.
- Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*, 1-20.
- Han, A. H., Olson, M. A., & Fazio, R. H. (2006). The influence of experimentally created extrapersonal associations on the Implicit Association Test. *Journal of Experimental Social Psychology*, *42*, 259-272.
- Johnson, M. K., Hashtroudi, S., & Lindsay, S. D. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.
- Kumkale, T. G., & Albarracín, D. (2004). The sleeper effect in persuasion: a meta-analytic review. *Psychological Bulletin*, *130*, 143-172.
- Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV What we know (so far) about the method. In B. Wittenbrink & N. Schwartz (Eds.), *Implicit measures of attitudes* (pp. 59-102). New York: The Guilford Press.
- 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.
- Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.
- Nosek, B. A., & Hansen, J. J. (in press). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition and Emotion*.

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.

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

Figure Captions

*Figure 1. IAT Effect as a Function of Message Arguments and IAT Type in Experiment 1*

*Figure 2. IAT Effect as a Function of Message Arguments and IAT Type in Experiment 2*



