Minimizing Method-Specific Variance in the IAT

A Single Block IAT

Sarah Teige-Mocigemba[1], Karl Christoph Klauer[1], and Klaus Rothermund[2]

[1]Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, Germany
[2]Institut für Psychologie, Friedrich-Schiller-Universität Jena, Germany

Abstract. The present paper introduces a new variant of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) called the Single Block IAT (SB-IAT). By eliminating the IAT's block structure, the SB-IAT is argued to solve the structural problem of recoding in the IAT and accordingly, its contamination by method-specific variance. In Study 1, a flower-insect SB-IAT, a task-switching ability SB-IAT, and a geometry SB-IAT showed reduced, but still significant effects. Zero correlations between the three SB-IATs indicated a substantially reduced amount of method-specific variance. Study 2 examined the SB-IAT's psychometric properties. A political attitude SB-IAT showed acceptable reliability, discriminated between liberal and conservative voters, and correlated with the corresponding attitude rating in the same magnitude as the standard IAT. Results indicate that the SB-IAT minimizes method-specific variance while retaining the IAT's satisfying psychometric properties. The discussion focuses on potentials and constraints of this newly developed measure.

Keywords: Implicit Association Test, method-specific variance, recoding, task-switching processes, speed-accuracy trade-offs

Introduction

In the 10 years since its publication, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) has received significant attention and has been widely used in diverse areas of research. The IAT has been shown to predict self-report data, behavior, group membership, and physiological responses, and has outperformed other response-time paradigms in terms of psychometric criteria and predictive validity (for a recent review of IAT research, see Nosek, Greenwald, & Banaji, 2006; Schnabel, Asendorpf, & Greenwald, 2008). Despite the IAT's widespread use, the processes underlying IAT effects are not yet sufficiently understood (e.g., Wentura & Rothermund, 2007). Consequently, there is still some ambiguity in the interpretation of IAT effects. According to its developers, the size and direction of IAT effects reflect the relative association strengths between target and attribute categories (Greenwald et al., 1998). However, a large body of research indicates that, besides associations between categories, nonassociative processes also contribute to IAT effects and cause additional systematic variance in IAT effects (e.g., De Houwer, 2003; Klauer & Mierke, 2005; McFarland & Crouch, 2002; Mierke & Klauer, 2001, 2003; Rothermund & Wentura, 2004). Although the accounts differ in many respects, they rely on the same fundamental idea: The IAT's block structure, more precisely, the comparison of performance in the incompatible vs. compatible block, is at the root of many of the identified confoundings (De Houwer, 2003). The consistent mapping of categories onto response keys across many trials in the incompatible vs. compatible block elicits qualitative and possibly strategic processing differences between the two blocks. These processing differences reflect unwanted sources of (systematic) nonassociative variance that contribute to the IAT effect and compromise an unequivocal interpretation.

The present paper focuses on a particular marker of such processing differences: method-specific variance. Extending prior research, we argue that method-specific variance in the IAT is largely the result of the IAT's block structure. The proposed solution is an IAT variant called the Single Block IAT (SB-IAT) that eliminates the block structure. In Study 1, we investigated whether the procedural modification of eliminating the block structure reduces method-specific variance. Study 2 examined the psychometric properties of the newly developed SB-IAT. Finally, potentials and constraints of the SB-IAT are discussed.

Implicit Association Test

The IAT comprises two categorization tasks that are performed in alternating order. In the concept task, stimuli of two target categories (e.g., flower vs. insect) are to be categorized according to their target category membership. In

DOI 10.1027/1015-5759.24.4.237

the attribute task, stimuli of two attribute categories (e.g., positive vs. negative) are to be categorized according to their attribute category membership. In the diagnostically relevant phases of the IAT, one target and one attribute category are assigned to one of two response keys, in two complementary mappings. The IAT rests on the assumption that if two categories are highly associated, categorization will be easier (i.e., faster and more accurate) when the two associated categories share the same response key (i.e., in the so-called "compatible" block) than when they require different responses, that is, when two nonassociated categories are mapped onto the same response key (i.e., in the so-called "incompatible" block). Thus, in a flower-insect IAT, better performance is found if the categories flower and positive as well as insect and negative share one response key than with the reversed mapping (flower and negative share one response key, insect and positive share the other key). The performance difference between these two kinds of mappings is called the IAT effect. Direction and size of the IAT effect are often interpreted as reflecting the relative association strengths between the target and attribute categories.

Method-Specific Variance in the IAT

Numerous encouraging findings have demonstrated that the IAT reliably assesses construct-specific variance (for an overview, see Greenwald, Poehlman, Uhlmann, & Banaji, in press). However, IAT effects have also been shown to be contaminated by stable, but construct-independent, method-specific variance (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Mierke & Klauer, 2003). Method-specific variance in the IAT is indicated by correlations between content-unrelated IATs for which one would not expect any correlations on a priori grounds. Mierke and Klauer (2003), for example, developed a control IAT, the so-called geometry IAT, in which simple geometrical objects (rectangles, triangles, circles) are used as stimuli. In the concept task, participants have to categorize objects according to color (target categories: red vs. blue), whereas in the attribute task, they have to categorize objects that are colored other than red or blue according to size (attribute categories: small vs. large). Importantly, color is confounded with size in that all red objects are small and all blue objects are large (or vice versa), which artificially creates associations between target and attribute categories. Accordingly, participants performed better when the two confounded categories shared one response key (red and small vs. blue and large) than when the two nonconfounded categories shared one response key (blue and small vs. red and large). Mierke and Klauer (2003) found that the geometry IAT correlated significantly with a flower-insect IAT and (with the absolute size of) an extraversion IAT effect, with correlations ranging between .30 and .40.

Similar results were found by Back, Schmukle, and Egloff (2005). In their task-switching ability IAT (TSA IAT), the concept task requires participants to discriminate letters (e.g., C) from numbers (e.g., 7), whereas the attribute task requires discrimination of words (e.g., shirt) from calculations (e.g., 8 – 5 = 3). Because words are associated with letters, and calculations with numbers, participants performed better when these associated categories shared one response key in comparison to the reversed mapping. Back and colleagues (2005) reported correlations between the TSA IAT and an anxiety IAT of similar magnitude as found by Mierke and Klauer (2003). Similarly, McFarland and Crouch (2002) found significant correlations between two control IATs and a flower-insect IAT. Finally, Klauer et al. (2007) used a political attitude IAT, the geometry IAT, and the TSA IAT, and found correlations ranging between .32 and .50. How can these correlations between content-unrelated IATs be explained? Recent research has identified two factors that can account for method-specific variance, namely cognitive skills and speed-accuracy trade-offs.

Cognitive Skills as Reliable Contamination of the IAT Effect

Two factors of cognitive skills have been shown to contribute to IAT effects: task-set switching (see Mierke & Klauer, 2001) and inhibition (see the quad-model; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005). Stable interindividual differences in such factors can account for the IAT's method-specific variance as has been explained, for example, by means of the task-set switching account (Mierke & Klauer, 2003): The IAT in its standard format requires participants to apply the attribute and the concept task in alternating order. In order to follow the instruction to respond as fast and accurately as possible, participants may try to facilitate the complex categorization task by recoding it (see De Houwer, 2003; Rothermund & Wentura, 2004). Such recoding processes are particularly likely to occur in the compatible block, because here, participants can capitalize on response synergy. The consistent (compatible) mapping of categories onto response keys allows for saving costly switches from the attribute to the concept task. In order to respond correctly, participants do not need to identify and – if necessary – switch to the appropriate task-set. Instead, they can categorize both attribute and target stimuli on the attribute dimension (e.g., valence) or on another dimension shared by attribute and target stimuli (e.g., perceptual similarities, De Houwer, Geldof, & De Bruycker, 2005, or salience asymmetries, Rothermund & Wentura, 2004). For instance, in the compatible block of a flower-insect IAT, responding to a flower stimulus on the basis of target category membership (i.e., concept task) or on the basis of valence (i.e., attribute task) leads to the same response. Thus, categorization of all target stimuli (e.g., flower) according to valence (here: positive) allows for fast and accurate responses in the compatible block. Such a recoding strategy, however, cannot be applied to the incompatible block, where accurate responding requires performing each task switch. Performance costs associated with task-set switching (e.g., Meiran, 1996; Rogers & Monsell, 1995) thus affect both blocks asymmetrically, and contaminate the IAT effect. However, task-switching costs may reflect *stable* interindividual differences in executive control processes as indicated by method-specific variance. Method-specific variance in the IAT can thus be interpreted as reflecting interindividual differences in the participants' ability to solve the IAT task, a phenomenon called the "cognitive skill confound" of the IAT (McFarland & Crouch, 2002, p. 493).

Speed-Accuracy Trade-offs as Reliable Contamination of the IAT Effect

Recently, Klauer et al. (2007) proposed a diffusion model analysis of the IAT that allowed for the dissociation of distinct parameters for construct- vs. method-specific variance. The analysis revealed that interindividual differences in speed-accuracy trade-offs in the compatible vs. incompatible block also account for method-specific variance. The differently difficult blocks of the IAT obviously triggered differently chosen speed-accuracy settings. Taken together, the IAT's block structure elicits contaminations by cognitive skills and speed-accuracy trade-offs that distort both the size of the IAT effect and its rank order, because these sources of systematic variance are unrelated to the purpose of measurement.

How to Deal with Method-Specific Variance

Different techniques have been suggested in order to decrease the confounding impact of method-specific variance in the IAT effect. Mierke and Klauer (2003), for example, proposed to remove task-switching trials from the analysis or to partial out method-specific variance, as was also suggested by Back et al. (2005). The use of the improved scoring algorithms recommended by Greenwald, Nosek, and Banaji (2003) has proven to be even more effective. Several studies have shown that correlations between substantive IATs and control IATs as markers of method-specific variance are reduced, although not consistently eliminated, when the D-scores are used instead of the conventional latency measure (e.g., Back et al., 2005; Klauer et al., 2007; Mierke & Klauer, 2003). Finally, the diffusion model allows for the dissociation of distinct parameters for construct- vs. method-specific variance (Klauer et al., 2007) and the quad model allows for the discrimination of four components including one for inhibition (Conrey et al., 2005).

Indeed, the aforementioned techniques all address the symptoms of the IAT's reliable contamination by cognitive abilities and speed-accuracy trade-offs, but they do not tackle the root of the problem of method-specific variance: the IAT's block structure. We therefore explored a small, but effective structural change within the IAT paradigm and eliminated the source of method-specific variance, namely, the IAT's block structure[1].

Solving the IAT's Problem of Method-Specific Variance: The Single Block IAT

The product of this structural change is the SB-IAT. The basic principle of the SB-IAT is that the mapping of categories onto response keys may change from trial to trial instead of blockwise. An otherwise irrelevant stimulus feature, namely word position, determines the valid response mapping (compatible vs. incompatible) for each trial. All stimuli are randomly presented above or below a dashed line that divides the screen into an upper and a lower half. If a stimulus appears in the upper half, the compatible mapping is valid (i.e., compatible categories share one response key). If a stimulus appears in the lower half, the incompatible mapping is valid (i.e., incompatible categories share one response key). Importantly, for attribute stimuli, word position is irrelevant, because attribute stimuli always have to be assigned to the same response keys irrespective of word position (e.g., positive stimuli to the right key, negative stimuli to the left key). For target stimuli, however, word position is highly relevant. For instance, if target stimuli in a flower-insect SB-IAT appear above the dashed line, flower stimuli have to be assigned to the right key, whereas insect stimuli have to be assigned to the left key. Conversely, if target stimuli appear below the dashed line, flower stimuli have to be assigned to the left key and insect stimuli have to be assigned to the right key. As a reminder, category labels are presented throughout all trials.

The main difference between the SB-IAT and the standard IAT is that the SB-IAT compares performance on compatible vs. incompatible trials *within the same* (i.e., a single) *block*, whereas the standard IAT compares performance on compatible vs. incompatible trials *between two different* (i.e., compatible vs. incompatible) *blocks*. Thus, in the SB-IAT, the response mapping (compatible vs. incompatible) may randomly change from trial to trial and is not consistently blocked anymore. This should impede any

[1] At present, several research groups test such IAT variants that focus on different aspects and accordingly, show clear differences in their buildup (e.g., Eichstaedt, 2007; Rothermund, Teige-Mocigemba, Gast, & Wentura, in press). We believe that the IAT variant introduced in the present paper is especially suitable for the assessment of interindividual differences because of its clear buildup.

Table 1. Single Block Implicit Association Test (SB-IAT): Task sequence

Block	*N* of trials	Task	Left key [A]	Right key [5]
1	26[a]	Target discrimination in the upper screen	insect	flower
2	26[a]	Target discrimination in the lower screen	flower	insect
3	26[b]	Combined target discrimination in the upper and lower screen	insect	flower
			flower	insect
4	26[a]	Attribute discrimination in the upper and lower screen	negative	positive
5–8	52[c]	Combined discrimination of target and attribute stimuli in the upper and lower screen	insect	flower
			negative	positive
			flower	insect

Note. [a]2 warm-up-trials + 24 practice trials; [b]2 warm-up-trials + 24 practice trials (2 [position] × 2 [category] × 6 [stimuli]); [c]4 warm-up-trials + 24 compatible test trials (6 stimuli per category) + 24 incompatible test trials (6 stimuli per category). Note that with the exception of the geometry SB-IAT in Study 1 all SB-IATs presented each stimulus once above and once below the dashed line in each test block, respectively. In the TSA SB-IAT, the category labels "number" and "letter" substituted the labels "insect" and "flower," and "calculation" and "word" substituted "negative" and "positive." In the geometry SB-IAT, "red" and "blue" substituted "insect" and "flower," and "small" and "large" substituted "negative" and "positive." In the political attitude SB-IAT, only target category labels changed: "black" and "red" substituted the labels "insect" and "flower."

kind of recoding strategies, because recoding processes rely on a consistent mapping of categories onto response keys (Strayer & Kramer, 1994).

If the structural change in the SB-IAT really prevented recoding, one would expect the SB-IAT to show still significant, but reduced, compatibility effects because the contribution of recoding processes to IAT effects should be minimized. Even more importantly, one would predict the SB-IAT to show clearly reduced method-specific variance as a direct marker of cognitive abilities and speed-accuracy trade-offs. This assumption was tested in Study 1.

Study 1

Using SB-IATs instead of standard IATs, we adapted the design of Klauer et al.'s (2007) Study 3. As elaborated above, Klauer et al. administered a political attitude IAT and two control IATs, the geometry IAT and the TSA IAT. They found that correlations between the political attitude IAT, on the one hand, and the geometry IAT and the TSA IAT, on the other hand (r = .32 and r = .38, respectively) were somewhat lower than correlations between both control IATs (r = .50). Klauer et al. argued that the more systematic construct-specific variance a measure contains the less is the proportion of method-specific variance in the total variance of IAT scores. Because we were especially interested in markers of method-specific variance, we refrained from using a political attitude SB-IAT in Study 1. Instead, we administered a flower-insect SB-IAT, because we expected much less variability in the student participants' preference for flowers over insects than in their political attitudes. Note that the flower-insect IAT has indeed been found to correlate at .53 with both the geometry IAT and the TSA IAT (Schmitz & Klauer, personal communication, November, 2005).

Method

Participants

Participants were 32 students from the University of Freiburg (20 female, 12 male) with various majors. Mean age was 23 years, ranging from 18 to 31 years. Compensation for participation was 3.50 EUR.

Overall Procedure

Participants first completed a flower-insect SB-IAT. This was followed by a task-switching ability SB-IAT (TSA SB-IAT) and a geometry SB-IAT. The order in which the latter two tasks were administered and the nature of the contingency realized in the geometry SB-IAT (red = small vs. red = large) were balanced across participants. Finally, participants were asked to report personal data (age, sex, handedness, and major), speculate about the true purpose of the experiment, and were then debriefed. In all studies of this paper, tests were presented on a computer with a 43 cm VGA color monitor with a resolution of 1280 pixels × 1024 pixels, and data were recorded using Inquisit software (2005).

SB-IATs

All SB-IATs consisted of eight blocks of either 26 or 52 trials. In Table 1, specifics of each block are summarized for the flower-insect SB-IAT. Note that the TSA SB-IAT and the geometry SB-IAT followed an analogous format. Participants started out practicing the concept and the attribute tasks. First, they were to categorize target stimuli in the upper half of the screen (e.g., left key for insect stimuli and right key for flower stimuli). Then, they were to categorize the same target stimuli in the lower screen half (e.g., left key for flower stimuli and right key for insect stimuli).

The tasks of the first two blocks were combined in the third block, in which participants had to correctly assign target stimuli depending on word position. In the fourth block, participants were to categorize attribute stimuli (e.g., left key for negative stimuli and right key for positive stimuli). In four ensuing test blocks, the tasks of the third and the fourth block were combined. For example, in the flower-insect SB-IAT, participants were to discriminate insect and negative from flower and positive (upper screen) or flower and negative from insect and positive (lower screen) depending on word position. All blocks were preceded by additional warm-up trials using stimuli that were reserved for the warm-up trials, one trial and one stimulus per category that appeared in the block. Single-task blocks were thus preceded by two warm-up trials; blocks combining both tasks were preceded by four warm-up trials. Participants used the left key "A" and the right key "5" on a standard computer keyboard to respond.

Target and attribute stimuli were presented in randomized order. Each trial started with the presentation of a fixation star in the center of the upper or lower screen indicating the valid mapping for the respective trial. After 500 ms the star was replaced by a stimulus, which remained on the screen until the correct key was pressed. In case of a false response, a red "X" was shown in the center of the screen until the correct response was given. The intertrial interval was 500 ms. It took participants approximately 10 min to complete an SB-IAT.

The flower-insect SB-IAT and the TSA SB-IAT used six stimuli per attribute and target category, which were presented in dark gray and black, respectively. The geometry SB-IAT presented circles, triangles, and squares in six different sizes and with outlines colored in one of six colors. Stimuli of the latter two SB-IATs were the same as in Klauer et al. (2007; Study 3). Analogous to analyses in the standard IAT, SB-IAT scores were calculated as the difference between the mean response latencies in the 96 incompatible trials and the mean response latencies in the 96 compatible trials.

Results

Following the conventional scoring procedure (Greenwald et al., 1998), analyses were based on log-transformed response latencies of correct responses, and latencies smaller than 300 ms or greater than 3000 ms were recoded to 300 ms or 3000 ms, respectively. Mean response latencies and error rates across the three SB-IATs (M = 764 ms, SD = 136 ms, and M = 8%, SD = 4%, respectively) were comparable to those known from prior IAT research, thereby indicating the feasibility of the SB-IAT task. Note that one participant's flower-insect SB-IAT data were excluded from all analyses because mean latency in the flower-insect SB-IAT was an extreme outlier in the distribution of the total sample according to Tukey (1977; mean latency was above the third quartile plus three times the interquartile range).

As expected, all SB-IAT effects differed significantly from zero, for the flower-insect SB-IAT, M = 29 ms, SD = 59 ms; for the TSA SB-IAT, M = 40 ms, SD = 125 ms; for the geometry SB-IAT, M = 51 ms, SD = 52 ms; all ts > 2.25, $p < .03$, $d > .40$. In order to calculate internal consistencies, we computed Cronbach's α for the four IAT scores of the four test blocks of each SB-IAT (flower-insect SB-IAT: $\alpha = .58$, TSA SB-IAT: $\alpha = .88$, geometry SB-IAT: $\alpha = .59$). Internal consistencies were somewhat lower than for the standard IAT, which typically range from .70 to .90 (see Nosek et al., 2006), but still higher than for other recently developed response-time paradigms such as the Extrinsic Affective Simon Task (e.g., Teige, Schnabel, Banse, & Asendorpf, 2004), affective priming (see Fazio & Olson, 2003), or the go/nogo association task (Nosek & Banaji, 2001). As predicted, there were no significant correlations between SB-IATs. The flower-insect SB-IAT did not correlate with the TSA SB-IAT, $r = .03$, $p = .86$, or with the geometry SB-IAT, $r = .10$, $p = .58$, nor were the correlations between the TSA SB-IAT and the geometry SB-IAT significant, $r = -.02$, $p = .91$. Note that inspection of the scatter plot revealed one extreme outlier sensu Tukey (1977) on the TSA SB-IAT score, which drove a nonsignificant correlation between the TSA SB-IAT and the geometry SB-IAT ($r = .18$, $p = .33$). This outlier was excluded from the correlational analysis. Note that with the present sample size, the power to detect the medium ($r = .32$) to large ($r = .50$) effects of Klauer et al. (2007; Study 3) was $1 - \beta = .59$ and .94, respectively (posthoc power analyses were conducted with G*Power3; Faul, Erdfelder, Lang, & Buchner, 2007)[2].

Discussion

As expected, all SB-IATs showed significant, but somewhat smaller effects than standard IATs. Zero-correlations between all SB-IATs indicated a clearly reduced contribution of method-specific variance in the SB-IATs as compared to the standard IATs in Study 3 of Klauer et al. (2007), which used highly comparable procedures and participant samples. Indeed, the geometry SB-IAT – TSA SB-IAT correlation, $r = -.02$, differed significantly from Klauer et al.'s geometry IAT – TSA IAT correlation, $r = .50$, $z = 2.05$, $p < .05$. Although the small sample size of Study 1 limits the explanatory power of the present findings, it may be concluded that the structural change of eliminating the block

[2] Note that adapting the D2-score (Greenwald et al., 2003) also revealed nonsignificant correlations: The flower-insect SB-IAT did not correlate with the TSA SB-IAT, $r = .17$, $p = .37$, or with the geometry SB-IAT, $r = .29$, $p = .12$, nor were the correlations between the TSA SB-IAT and the geometry SB-IAT significant, $r = .24$, $p = .21$.

structure minimizes contamination of the IAT effect by method-specific variance.

Importantly, lower internal consistencies of the present SB-IATs are not surprising given that reliability estimates depend on the amount of interindividual variability. The IAT's satisfactory reliability is thought to stem from two systematic, but conflated sources of variance (i.e., construct- and method-specific variance; Mierke & Klauer, 2003), whereas the SB-IAT's reliability should just stem from one systematic source of variance (i.e., construct-specific variance). Thus, if variability in the construct of interest is rather low, as should be the case for the associations assessed in Study 1, reducing method-specific variance in the SB-IAT should be accompanied by reduced reliability estimates. If, however, participants' variability in the construct of interest is high, reliability estimates for the SB-IAT should also be higher. Consequently, Study 1 might have underestimated the SB-IAT's reliability. Therefore, and because we were also interested in the SB-IAT's validity (here, with regard to implicit-explicit consistency), Study 2 examined the SB-IAT's psychometric properties.

Study 2

For two reasons, the domain of political attitudes appeared to be suitable for evaluating the SB-IAT's psychometric properties. Firstly, considerable variability in the participants' political attitudes should allow for fair reliability estimates. Secondly, the moderate implicit-explicit correlations that are usually found in this domain (Greenwald et al., in press) should allow for validity estimates of the SB-IAT. Following Klauer et al.'s (2007) Study 2, a political attitude SB-IAT and explicit political-attitude ratings contrasted a red vs. black political attitude that is associated with the left vs. right political spectrum in Germany.

Method

Participants

Participants were 40 students from the University of Freiburg (25 female, 15 male) with various majors. Mean age was 22 years, ranging from 19 to 27 years. Again, compensation for participation was 3.50 EUR. Data of one participant were excluded from all analyses because her mean error rate in the SB-IAT of 35% was an extreme outlier in the distribution of the total sample. Thus, the final sample consisted of 39 participants.

Overall Procedure

Adapting the design of Klauer et al.'s (2007) Study 2, participants first completed self-report measures of political attitude before they worked through a political attitude SB-IAT. Finally, participants were asked to report personal data, speculate about the true purpose of the experiment, and were then debriefed.

Self-Report Measures

Self-report measures were as follows: (a) a 10-point Likert scale for the personal political attitude on a red vs. black dimension, (b) separate 10-point thermometer ratings for the red and the black political standpoint, and (c) 10-point Likert scales for the valence of each of the target stimuli used in the political attitude SB-IAT. The last two sets of ratings were averaged per person, with reverse scoring for ratings pertaining to black political attitudes (and categories). All three measures were then z-transformed, and the average of the three z-scores was the explicit measure of political attitude (Cronbach's α = .91). Participants were also asked to rate their interest in political issues and events and whether they would vote for the red or black political spectrum if elections were held next Sunday.

Political Attitude SB-IAT

The political attitude SB-IAT used the same format and parameters as the SB-IATs of Study 1 except that target and attribute trials were presented in alternating order. Stimuli of the target categories red vs. black and the attribute categories positive vs. negative were the same as in Klauer et al. (2007; Study 2).

Results

Response latencies were preprocessed as in Study 1. Participants needed a little more time to complete the political attitude SB-IAT (M = 998 ms, SD = 270 ms) as compared to the mean response latencies for the SB-IATs in Study 1, whereas error rates were in the same range (M = 8%, SD = 7%). As expected, the political attitude SB-IAT showed acceptable internal consistency when computed as in Study 1, Cronbach's α = .74. The SB-IAT proved to be valid both on the group level and on the correlational level. It discriminated between participants who indicated an intention to vote for the red political spectrum and participants who indicated an intention to vote for the black political spectrum[3], $t(37) = 4.77$, $p < .001$, and correlated to a moderate

[3] The mean SB-IAT effect was 65 ms (SD = 93 ms), $t(38) = 4.61$, $p < .001$, indicating a general preference for the red political spectrum. This corresponds to the finding that 62% of the subjects indicated an intention to vote for the red political spectrum, if elections were held next Sunday.

degree with the attitude rating, $r = .43$, $p < .01$, as has also been found for the highly comparable IAT of Klauer et al. ($r = .42$)[4]. Importantly, the SB-IAT's prediction of voting intention (red vs. black political spectrum) was mediated by the attitude rating: If the SB-IAT and the attitude rating entered a logistic regression separately, they both predicted voting intention, $B = 2.10$, $SE = .73$, $p < .01$, and $B = 7.29$, $SE = 3.11$, $p = .02$, respectively. Also, the SB-IAT predicted the attitude rating, $\beta = .43$, $p < .01$. If, however, the SB-IAT and the attitude rating entered the logistic regression simultaneously, only the attitude rating ($B = 7.20$, $SE = 3.43$, $p = .04$), but not the SB-IAT ($B = .76$, $SE = .77$, $p = .33$) predicted voting intention.

Discussion

The SB-IAT reliably assessed interindividual differences in political attitudes and proved to be valid, both in terms of discriminating between red- vs. black-voters and in terms of implicit-explicit correlations. However, it did not show incremental validity in predicting voting intention over and above the attitude rating, as has also been shown for the IAT: In socially insensitive domains (e.g., political attitudes), explicit measures outperformed the IAT with regard to predictive validity (see Greenwald et al., in press). Interestingly, the SB-IAT's impact on voting behavior was mediated by its impact on the attitude rating, a finding consistent with recent models that suggest a (default) bottom-up influence of associative processes on propositional/reflective processes of evaluation (e.g., see Gawronski & Bodenhausen, 2006). One might have expected that reducing the amount of method-specific variance should be accompanied by an increase of construct-specific variance and,thus, higher implicit-explicit correlations for the SB-IAT as compared to the standard IAT. The SB-IAT's correlation with the attitude rating, however, was of the same magnitude as the IAT's correlation with the attitude rating in Klauer et al.'s (2007) Study 2. Importantly, this finding corresponds to recent research showing that even partialing out method-specific variance by means of a control IAT only slightly increases implicit-explicit correlations (see Back et al., 2005; Mierke & Klauer, 2003) as can be easily calculated using a formula provided by Mierke and Klauer (2003, p. 1188).

General Discussion

The present paper introduces a newly developed IAT variant called the Single-Block IAT. By eliminating the IAT's block structure, the SB-IAT is argued to solve the structural problem of recoding in the IAT and, accordingly, reduce its contamination by systematic method-specific variance. Study 1 provided first evidence for this assumption: A flower-insect SB-IAT, a TSA SB-IAT, and a geometry SB-IAT showed reduced, but still significant, compatibility effects. Zero-correlations between the three SB-IATs indicated a reduced amount of method-specific variance relative to the standard IAT (see Klauer et al., 2007). As method-specific variance is usually interpreted as a marker of the IAT's contamination by cognitive skills (e.g., Mierke & Klauer, 2003) or speed-accuracy settings (Klauer et al., 2007), this finding suggests that the SB-IAT is affected by these unwanted sources of variance to a smaller degree. In order to demonstrate the SB-IAT's ability to reliably assess meaningful construct-variance, Study 2 examined the SB-IAT's psychometric properties in the domain of political attitudes. The political attitude SB-IAT showed acceptable reliability, discriminated between red- and black voters, and correlated with the corresponding attitude rating in the same magnitude as the standard IAT. This finding is remarkable insofar as other recently developed response-time paradigms suffered from unsatisfying reliabilities (Nosek et al., 2006) and thus could not compete with the IAT in terms of reliability. Improving the IAT by focusing on a structural change within the IAT paradigm thus appears to be a promising approach that seems to reduce method-specific variance without compromising reliability and validity.

Importantly, elimination of the IAT's block structure should have further advantages. First, the IAT has been shown to be affected by compatibility-order (e.g., Nosek et al., 2006): IAT effects tend to be larger if the compatible block precedes the incompatible block than vice versa. Klauer and Mierke (2005) suggested that differences in the accessibility of attribute information in the compatible vs. incompatible block of the IAT may account for this effect. Because compatibility is a function of interindividual differences in the attitude of interest and cannot be determined a priori in many applied contexts, such compatibility-order effects constitute an undesirable confounding in the IAT and might influence both the magnitude and the rank order of individual IAT effects. By eliminating the block structure, the SB-IAT cannot be affected by confounding compatibility-order effects[5].

Secondly, Olson and Fazio (2004) showed that IAT effects are confounded by "extrapersonal associations," that is, culturally shared assumptions (e.g., apples are healthy and thus are positive) that do not necessarily correspond to personal evaluations (e.g., I don't like apples). Although not explicitly stated by the authors, one might hypothesize that this confounding is based on differences in the extent to which participants *strategically* use extrapersonal knowledge "when solving the mapping problem posed by

[4] Note that using the D2-score (Greenwald et al., 2003) led to the same results. The SB-IAT showed satisfactory reliability ($\alpha = .81$), discriminated between red- vs. black-voting participants, $t(37) = 3.47$, $p = .001$, and correlated with the attitude rating, $r = .44$, $p < .01$.

[5] Note that Nosek, Greenwald, and Banaji (2005) recently proposed a technique for reducing effects of compatibility-order.

the IAT" (Olson & Fazio, 2004, p. 661). Importantly, strategy use requires the consistent mapping of categories onto response keys across many consecutive trials, such as in the critical IAT blocks (see Fazio & Olson, 2003; Strayer & Kramer, 1994). Inasmuch as the IAT's contamination by extrapersonal associations is based on its block structure, we would expect the SB-IAT to be less affected by this confound, although of course, this assumption needs empirical testing.

Last, but not least, Govan and Williams (2004) demonstrated the crucial role of stimulus selection in the IAT. They showed that the affective valence of the chosen stimuli can determine the interpretation of the IAT's category labels, which influences size and direction of IAT effects. Again, one might expect that such processes of redefining the category labels require the consistent mapping of the IAT's block structure and thus might not occur in the SB-IAT. Very recently, Rothermund et al. (in press) provided first evidence for this assumption: Changing the affective valence of stimuli influenced the IAT, whereas an SB-IAT variant was unaffected by such changes.

Of course, the SB-IAT does not provide a solution for each and every problem of the IAT, and has some potential shortcomings itself. For instance, the complicated structure of the SB-IAT might be seen as a disadvantage. On the other hand, mean response latencies and error rates in the present studies indicated that respondents from a student population had little difficulty responding fast and accurately. For other populations, it may prove useful to simplify the SB-IAT task by changing some presentational parameters (e.g., longer presentation of each trial's fixation star in order to facilitate preparation for the valid response mapping). Another criticism might be that in addition to task-set switching, the SB-IAT introduces another type of switching, namely switching between the compatible and incompatible response mappings. Note, however, that contrary to task-switches in the IAT, the two types of switches in the SB-IAT contribute to the compatible and incompatible response mappings to the same extent. Thus, switching might increase error variance, but does *not* contaminate the SB-IAT effect as the reduced method-specific variance in the SB-IAT shows.

Two anonymous reviewers raised the question whether recoding based on valence may, in part, reflect construct-related variance and whether any attempts to reduce method-specific variance might thus lower the IAT's validity. In some domains (e.g., aggression), method-specific variance may, indeed, reflect construct-relevant information, at least inasmuch as interindividual differences in task-switching abilities are related to impulse control (which in turn might predict specific behavior). However, recent research did not confirm the assumption that method-specific variance systematically contributes to the IAT's validity: Partialing out method-specific variance did *not* reduce the IAT's validity (Klauer & Mierke, 2003) and a diffusion model analysis dissociated *distinct* parameters for construct- vs. method-specific variance (Klauer et al., 2007). Finally, even if method-specific variance contained construct variance, it nevertheless would appear to be worthwhile to design a task that impedes any type of recoding, because differences in the chosen recoding strategy and in the extent to which people recode the IAT task would still contaminate its effects.

Doubtlessly, the present paper only provides first evidence for the suitability of the SB-IAT, and future research is needed to clarify under which circumstances the SB-IAT might be superior to other response-time paradigms. Compared to the IAT, we would expect the SB-IAT to be less susceptible to effects of cognitive skills, compatibility order, extrapersonal associations, and stimulus influences, inasmuch as influences of these variables result from the IAT's block structure.

Acknowledgments

We would like to thank two anonymous reviewers for thoughtful comments on a previous draft of this manuscript. The research reported in this paper was supported by Grant Kl 614/31-1 from the Deutsche Forschungsgemeinschaft to Karl Christoph Klauer.

References

Back, M.D., Schmukle, S.C., & Egloff, B. (2005). Measuring task-switching ability in the Implicit Association Test. *Experimental Psychology, 52*, 167–179.

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

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50*, 77–85.

De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology, 59*, 228–239.

Eichstaedt, J. (2007). *Choice prediction with a balanced IAT.* Unpublished manuscript, Helmut Schmidt University, Hamburg.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

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

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.

Govan, C.L., & Williams, K.D. (2004). Changing the affective valence of the stimulus items influences the IAT by redefining 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., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197–216.

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

Inquisit 2.0.50401 [Computer software]. (2005). Seattle, WA: Millisecond Software LLC.

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

Klauer, K.C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93*, 353–368.

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

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1423–1442.

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

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

Nosek, B.A., & Banaji, M.R. (2001). The go/no-go association task. *Social Cognition, 19*, 625–666.

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

Nosek, B.A., Greenwald, A.G., & Banaji, M.R. (2006). The Implicit Association Test at age 7: A methodological and conceptual review. In J.A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.

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

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.

Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (in press). Minimizing the influence of recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF). *Quarterly Journal of Experimental Psychology.*

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

Schnabel, K., Asendorpf, J.B., & Greenwald, A.G. (2008). Assessment of individual differences in implicit cognition: A review of IAT measures. *European Journal of Psychological Assessment, 25,* 210–217.

Strayer, D.L., & Kramer, A.F. (1994). Strategies and automaticity: I. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 318–341.

Teige, S., Schnabel, K., Banse, R., & Asendorpf, J.B. (2004). Assessment of multiple implicit self-concept dimensions using the extrinsic affective Simon task (EAST). *European Journal of Personality, 18*, 495–520.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Wentura, D., & Rothermund, K. (2007). Paradigms we live by. A plea for more basic research on the IAT. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 195–215). New York: Guilford.

Sarah Teige-Mocigemba

Department of Social Psychology and Methodology
University of Freiburg
D-79085 Freiburg
Germany
Tel. +49 761 203-2474
Fax +49 761 203-2417
E-mail teige@psychologie.uni-freiburg.de

