# Task-Set Inertia, Attitude Accessibility, and Compatibility-Order Effects: New Evidence for a Task-Set Switching Account of the Implicit Association Test Effect

**Karl Christoph Klauer**
*Albert-Ludwigs-Universität, Freiburg*

**Jan Mierke**
*Rheinische Friedrich-Wilhelms-Universität, Bonn*

*Based on a task-set switching account of the Implicit Association Test (IAT), the authors predict a specific pattern of aftereffects as a consequence of working through IAT blocks. In Study 1, performance in an evaluative decision task, but not in a color-naming task, was decreased after working through the incompatible rather than compatible block of a flower-insect IAT. In Study 2, response latencies in an evaluative rating task, but not in a color-rating task, were analogously affected, whereas the ratings themselves were not a function of the compatibility of prior IAT blocks. The aftereffects demonstrate reactivity of the IAT; they bear on the mechanisms underlying the IAT and on compatibility-order effects.*

**Keywords:** *Implicit Association Test; task switching; attitude accessibility; indirect attitude measurement*

In recent years, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) has been widely used to measure a variety of psychological constructs. The IAT comprises five phases involving two tasks. The two tasks are a categorization task, in which exemplars of two target categories (e.g., flowers and insects) are to be classified according to their category membership, and an attribute task, in which stimuli are to be classified with respect to an attribute dimension (e.g., as either positive or negative). The critical phases are the Phases 3 and 5, in which both tasks are combined. In these phases, both tasks are mapped onto the same response keys, which can be done in two qualitatively different ways; for example, flowers and positive stimuli can share one of the two response keys and insects and negative stimuli the other one. Another possibility is that flowers and negative stimuli share the same response key and insects and positive stimuli the other one. The former response mapping usually leads to better performance than the latter. The mapping that leads to faster and more accurate responding is called the compatible mapping, the other one the incompatible mapping.

The IAT rests on the assumption that it is easier to make the same behavioral response to concepts that are strongly associated than to concepts that are weakly associated. In this view, a strong association is indicated by a large performance difference between Phases 3 and 5. This performance difference is called the IAT effect. For a variety of reasons, the IAT has instigated an enormous amount of research reviewed by Fazio and Olson (2003), among others.

As pointed out by Fazio and Olson (2003), considering the mechanisms by which indirect measures such as the IAT work is likely to further our understanding of what exactly is measured by a given indirect measure and of the interrelationships that may exist between different

**Authors' Note:** The research reported in this article was supported by Grants Kl 617/29-1 and Kl 617/29-2 from the Deutsche Forschungsgemeinschaft to the first author. The authors thank Jan De Houwer, Brian Nosek, and Klaus Rothermund for helpful comments on a first version of this article. Correspondence concerning this article should be addressed to K. C. Klauer at the Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, D-79085 Freiburg, Germany; e-mail: christoph.klauer@psychologie.uni-freiburg.de.

*PSPB*, Vol. 31 No. 2, February 2005 208-217
DOI: 10.1177/0146167204271416

indirect measures. Considerable progress has been made in improving methodological aspects of the IAT by means of large-scale exploratory studies (Greenwald, Nosek, & Banaji, 2003; Nosek, Greenwald, & Banaji, 2003); theory-driven research on the underlying mechanism also suggests ways in which the measurement procedure can be improved (Mierke & Klauer, 2003). For example, Mierke and Klauer (2003) predict, based on their account of the underlying mechanisms, that different IATs share some amount of method variance due to a cognitive-skill confound (McFarland & Crouch, 2002). They demonstrate that such method variance exists, and they propose and illustrate theory-based ways to control for it.

The mechanism that drives the IAT effect is unclear. As elaborated below, possible mechanisms that have received some attention are a random-walk model by Brendl, Markman, and Messner (2001); a figure-ground asymmetry model by Rothermund and Wentura (2001, in press); a stimulus-response compatibility model by De Houwer (2001); and a task-set switching account by Mierke and Klauer (2001, 2003).

The purpose of the present article is to test a number of nonobvious predictions of the task-set switching account. In addition to this primary theoretical focus, the predictions and findings offer an explanation of the effect of compatibility order that is frequently encountered in applied IAT research (e.g., Greenwald et al., 1998, 2003; Nosek et al., 2003): IAT effects tend to be more pronounced if the compatible phase precedes the incompatible phase rather than vice versa. Because compatibility order is a function of interindividual differences in many applied contexts and cannot be determined a priori, such compatibility-order effects constitute an undesirable confounding in the measurement of interindividual differences. Understanding the causes of compatibility-order effects and other kinds of reactivity in IATs suggests ways to control for them and thereby to improve the measurement by IATs as discussed below. Finally, the current experimental methods and results suggest a new and unobtrusive way to manipulate what Fazio (1989) has called attitude accessibility.

## THE TASK-SET SWITCHING ACCOUNT

Mierke and Klauer (2001, 2003) assume that the IAT effect reflects differential costs of task switching. Task switching occurs in the critical IAT phases, in which participants are instructed to switch between two tasks on a trial-by-trial basis. Task switches are associated with costs (e.g., Allport, Styles, & Hsieh, 1994; Meiran, 1996; Rogers & Monsell, 1995), that is, with increased response latencies and error rates. Mierke and Klauer argue that the overall task-switch costs are smaller in the compatible than in the incompatible block. In the compatible block, participants can perform fast and accurately even if they do not switch from the attribute task to the categorization task and instead evaluate category exemplars on the attribute dimension. For example, in a flower-insect IAT, responding to "tulip" on the basis of its category membership as flower or on the basis of it being positively evaluated leads to the same response in the compatible phase. This means that participants do not have to perform each and every task switch to perform fast and accurately. In contrast, accurate responding in the incompatible phase requires performing each task switch, implying a larger overall task-switch cost for that phase.[1]

The task-set switching account predicts a subtle sequential effect. Trial $n$ is called a switch trial if it calls for another task than trial $n-1$; it is a repeat trial if the same task is to be performed as in trial $n-1$. The performance difference between switch and repeat trials reflects one component of task-switch costs. Mierke and Klauer (2001, 2003) consistently found larger response latencies and error rates for switch than for repeat trials in the IAT. Of importance, this sequential effect was more pronounced for incompatible than for compatible blocks. It thereby directly contributes to the overall IAT effect, although it does not account for it completely. Furthermore, in line with the task-set switching account, IAT effects were reduced, although not eliminated, when task cues preceded each trial and allowed participants to prepare for the upcoming task (Mierke & Klauer, 2001).

The present article focuses on another component of task-switch costs that is often considered responsible for remaining, so-called residual task-switch costs. Task switching requires the activation of appropriate task sets and the suppression of competing and interfering task sets. Allport et al. (1994) argue that task sets exhibit so-called task-set inertia: Once activated, task sets maintain a heightened state of activation or readiness for substantial amounts of time; conversely, if a given task set must be suppressed, it is subsequently more difficult to activate (cf. Mayr & Keele, 2000).

Task-set inertia in the form of enduring activation or inhibition can contribute to IAT effects: In incompatible blocks, exemplars of the target category are associated with different responses under the categorization task and under the attribute task. A correct response therefore requires suppressing the attribute–task set. Doing so is likely to be effortful and time-consuming given that this task set has only recently been activated in previous trials and may still be in a state of heightened activation. In fact, increased interference by irrelevant features (i.e., in the present case, the attribute information) that were relevant in a recently activated task set has consistently been found in the form of so-called reverse Stroop effects (e.g., Allport et al., 1994). Similarly, the inhibition of the attribute–task set may linger and make it more difficult to reactivate that task set when it is again appropriate, that is, in the presence of attribute stimuli. In fact, as pointed out by Brendl et al. (2001), performance decrements in the incompatible relative to the compatible blocks are also found for attribute stimuli even though these stimuli are typically not associated with conflicting category information. In contrast, in compatible blocks, a suppression of the attribute task is not necessary so that the just-described costs for the categorization task and the attribute task do not arise. To the contrary, the comparatively heightened level of activation of the attribute–task set may even facilitate responding in the categorization task through congruency effects. In summary, inertia effects are predicted that affect both attribute and category-exemplar stimuli and that directly contribute to IAT effects.

Aftereffects of this kind can be surprisingly long-lived (Mayr & Liebscher, 2001; Wylie & Allport, 2000). If they affect performance in the IAT, they contribute not only to the IAT effect itself but they also explain a couple of phenomena that suggest a reactivity of IAT procedures, in particular the above-mentioned order effects: Only in incompatible blocks is it regularly required to suppress the attribute–task set. If the inhibition is sufficiently long-lived, it should persist during subsequent compatible blocks and reduce performance in the attribute task. Performance in the categorization task also would be depressed because facilitation by congruent attribute information would be diminished in a subsequent compatible block. As a consequence, the IAT effect, that is, the performance difference between the compatible and the incompatible block, would be decreased relative to the case where the compatible block is presented first.

## STUDY 1

Study 1 is a first attempt to find the aftereffects discussed above in the IAT. Specifically, it is tested whether performance in the attribute task is differentially affected by working through compatible versus incompatible IAT blocks. Performance in the attribute task was assessed in special test blocks that succeeded regular compatible or incompatible IAT blocks.

In the IAT blocks of Study 1, the categorization task was to classify words as flowers or insects, whereas the attribute task was an evaluative decision task, in which words have to be evaluated as denoting something positive or something negative. In the test blocks, the attribute task was to be applied to the exemplars of the target categories, that is, flowers and insects had to be evaluated as either positive or negative. These trials were mixed with an equal number of trials of a new color-naming task in which the color (blue vs. red) of meaningless letter strings was to be named. To avoid trivial interferences by learned response-key mappings, responses in the test blocks were to be spoken out loud, whereas responses in the IAT blocks were key presses.

Participants first worked through an IAT block of 24 trials, followed by a test block. To maximize test power for detecting possible aftereffects of IAT blocks, the sequence of IAT block and test block was repeated 12 times. One group of participants worked through compatible IAT blocks, a second group through incompatible IAT blocks. Two control groups also worked with the compatible response mapping (Group 3) or the incompatible mapping (Group 4), but the 24 IAT trials were ordered so that task switches were minimized. Thus, either all 12 attribute stimuli were shown first, followed by 12 flowers and insects, or all 12 flower and insect names were shown first, followed by the attribute stimuli.

The predictions are straightforward. If task switching requires the suppression of the attribute–task set in incompatible blocks, and if the inhibition is sufficiently long-lived, performance in the attribute task should be reduced in test blocks following an incompatible IAT block, in which both tasks are mixed in a trial-by-trial fashion. Neither compatible IAT blocks nor blocks in which the number of task switches is minimized (Groups 3 and 4) should have this effect.

Introducing the color-naming task allows us to evaluate a simple alternative explanation: Brendl et al. (2001) have argued that incompatible blocks lead to the adoption of a more conservative response threshold; similarly, Mayr and Liebscher (2001) have argued that more effortful and conservative global control strategies may be implemented in difficult task-switching contexts such as the incompatible IAT block with mixed tasks. If these are still in force in subsequent blocks, performance should not only be decreased in the attribute task but also in the color-naming task. In contrast, the task-set switching account predicts an aftereffect that is restricted to the previously suppressed attribute task.

### *Method*

The two between-participants factors of Study 1 are manipulated in the IAT blocks that precede the test blocks. They are "response compatibility" (compatible vs. incompatible) and "task mixture" (mixed vs. blocked). In groups with mixed tasks, the trials of the IAT blocks were presented in random order; in groups with blocked tasks, all trials for one of the IAT tasks are presented first, followed by all trials of the other IAT task.

*Participants.* Participants were 49 University of Bonn students with different majors. They either received partial course credit or a monetary gratification of 5 Euro for their participation. Data from one participant were lost due to a technical failure in recording test-block performance.

*Stimuli.* The stimuli for the IAT blocks were words denoting 24 insects, 24 flowers, 24 positive objects, and 24 negative objects, taken from a previous study (Mierke & Klauer, 2001). The words were matched in 24 quadruples that were selected to be as similar as possible on three criteria: number of characters, word frequency as estimated through the CELEX lexical database (Celex, 1995), and a rating of the word's valence. Details on stimulus selection can be found in Mierke and Klauer (2001).

Stimuli for the test blocks were the 24 insect and the 24 flower words already used in the IAT blocks as well as 24 consonant strings presented in red color and 24 consonant strings presented in blue color. The consonant strings had a length of seven letters that were randomly sampled from a set of consonants excluding the letters q, x, y, and z.

*Block construction.* IAT blocks comprised 24 trials, presenting, respectively, six insects, six flowers, six positive objects, and six negative objects randomly sampled from the above-described pools without replacement. The 24 trials were presented in random order in the groups with mixed tasks and blocked by task (i.e., either all insects and flowers first or all positive and negative objects first) in the control groups with blocked tasks. Test blocks comprised six stimuli each of insects, flowers, red consonant strings, and blue consonant strings randomly sampled from the respective stimulus pools. The 24 stimuli were presented in random order.

*Procedure.* For the groups with compatible response mappings, flowers and positive objects shared the same response key in IAT blocks; insects and negative objects shared the other response key. For the groups with incompatible response mappings, flowers and negative objects shared one response key and insects and positive objects the other response key. The response keys were to be pressed with different hands and assignment of positive objects (and equivalently of flowers) to the dominant versus nondominant hand was counterbalanced over participants.

For the test blocks, participants were instructed to evaluate the word stimuli by saying "positive" or "negative" into a microphone that was part of a headset worn by participants and to name the color of consonant strings. Participants' responses were recorded with a sampling rate of 22048 Hz for a duration of 2 s following stimulus onset.

In both blocks, a trial began with the presentation of a fixation mark in the center of the screen. After 800 ms, the fixation mark was replaced by the stimulus. Upon a response, the next trial was initiated after 200 ms.

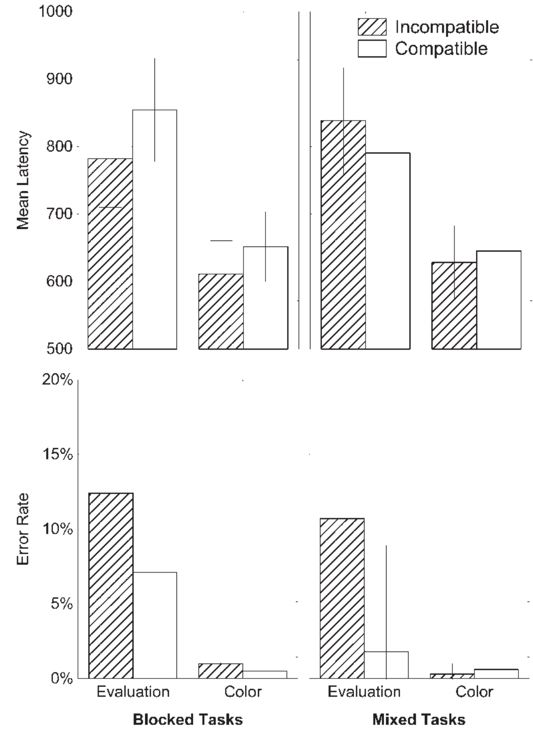


**Figure 1 Mean response latencies (upper panels) and error rates (lower panels) as a function of response compatibility and task mixture for the evaluative decision task and the color-naming task.**

NOTE: Error bars show 95% confidence intervals.

Participants were allowed to rest after each pair of IAT block and test block. There were 12 such pairs, and one session required approximately 30 min.

*Results*

Data from three participants were excluded because these participants consistently classified flowers as negative and insects as positive contrary to our a priori assumptions. Response compatibility in the IAT blocks may not have been as intended for these participants. The same pattern of results emerged, however, when the data from these participants were included. Test-block trials with utterances outside the response set (blue, red, positive, negative; 0.4%) and without response within the 2 s recording interval (0.7%) were excluded from the analyses, thereby leaving out a total of 1.1% of the data.

Figure 1 shows the average latencies (upper panels) and error rates (lower panels) as a function of task mixture (blocked vs. mixed) and response compatibility (compatible vs. incompatible) for the evaluative decision task and the color-naming task. For the sake of comparability with Study 2, latencies include trials with false responses, but the same pattern of results emerges if these are excluded.

Turning first to the latency data, an analysis of variance with factors task, response compatibility, and task mixture revealed the expected three-way interaction, $F(1, 41) = 5.27$, $p = .03$. Separate analyses of variance showed that the two-way interaction of task and response compatibility was significant given IAT blocks with mixed tasks, $F(1, 20) = 6.37$, $p = .02$, but not given blocked tasks, $F(1, 21) < 1$, as expected. Given mixed tasks, performance in the evaluative decision task is decreased following an incompatible IAT block relative to the compatible block, but color-naming performance is not decreased as can be seen in Figure 1 (upper right panel). The just-described simple effect on the evaluative decision task does not, however, reach significance in the latency domain, $t(20) = 1.24$, one-tailed $p = .11$; as expected, there is no effect of response compatibility on the color-naming task, $t(20) = –0.55$, $p = .58$.

In the error data, the expected three-way interaction was not significant, $F(1, 41) < 1$, but the two-way interaction of task and response compatibility was significant given mixed tasks, $F(1, 20) = 6.37$, $p = .02$, and it was not significant given blocked tasks, $F(1, 21) < 1$, as expected. Given mixed tasks, significantly more errors are made in the evaluative decision task after an incompatible than after a compatible IAT block, $t(20) = –2.33$, $p = .03$, but performance in the color-naming task is not affected by response compatibility, $t(20) = 0.86$, $p = .40$ (see lower right panel of Figure 1).

### *Discussion*

Study 1 provides initial evidence for differential aftereffects of working through compatible versus incompatible IAT blocks. The aftereffects followed the very specific form predicted by the task-switch analysis: Given mixed tasks in the IAT blocks, performance in the evaluative decision task, but not in the color-naming task, was poorer after incompatible than after compatible IAT blocks. When the number of task switches in the IAT blocks was minimized, no effects of compatibility order emerged at all.[2] The two-way and three-way interactions implied by this pattern of results were significant with few exceptions (there was no three-way interaction of task, response compatibility, and task mixture in the error data).

This pattern of effects permits a couple of first conclusions. Working through an IAT block has aftereffects on subsequent task performance that remain sizeable for a duration of at least 24 trials. Differential aftereffects of compatible versus incompatible IAT blocks were found in the groups with mixed tasks but not in the groups with blocked tasks, underlining the causal role of task switches for the aftereffects. The differential aftereffects are restricted to the attribute task, suggesting that shifts in global control strategies or global response criteria are not responsible.

The differential aftereffects were found even though the response modality changed from keypresses in the IAT block to vocal responses in the test blocks. This suggests that the attribute connotations of the category-exemplar stimuli, rather than specific stimulus-response associations, were differentially affected by prior compatible versus incompatible IAT blocks. In the task-set switching literature, there is considerable debate about the precise nature of such aftereffects. One possibility is that the entire attribute–task set is affected, suggesting that aftereffects also should be found when stimuli are evaluated that were not seen in IAT blocks (Allport et al., 1994; Mayr & Keele, 2000). Another possibility is that the aftereffects are confined to the specific stimuli that were processed in episodes in which inappropriate task sets had to be suppressed and appropriate task sets activated (Wylie & Allport, 2000), that is, to the category-exemplar stimuli of the prior IAT blocks in the present case. In this latter view, the accessibility of these stimuli's attribute connotations is affected and the differential attribute accessibility is responsible for the aftereffects discussed under the label of task-set inertia. Further possibilities are considered in the General Discussion. For the time being, it is clear that the aftereffects affect the speed with which the evaluations of the category-exemplar stimuli can be retrieved. Although the effect may be more general, working through IAT blocks of an attitude IAT thus appears to affect the attitude accessibility of the used category exemplars.

## STUDY 2

Study 2 was a conceptual replication and extension of Study 1, using different test tasks and yet another response modality, namely, mouse clicks. The effects on attitude accessibility raised the possibility that attitude extremity also might be affected. The tasks used in this study's test blocks were (a) to rate the category-exemplar stimuli as quickly as possible on a scale from *strongly negative* to *strongly positive* and (2) to rate the color of the consonant strings as quickly as possible on a scale from *saturated red* to *saturated blue*. We expected to replicate the pattern of effects on attitude accessibility (i.e., on the response latencies of the evaluative ratings) but there were no theoretical reasons to expect effects on attitude extremity (i.e., on the ratings themselves).

### *Method*

The method closely followed that of Study 1 unless otherwise mentioned. In particular, the same stimuli were used as before, and assignment of positive objects (and equivalently of flowers) to the dominant versus nondominant hand in IAT blocks was counterbalanced across participants. The saturation of the color (red or blue) of the consonant strings was, however, varied in eight steps. The only other difference was that different tasks were to be performed in the test block: Participants were asked to evaluate flowers and insects on a 6-point rating scale ranging from –3 to 3, with left endpoint labeled *strongly negative* and right endpoint labeled *strongly positive*. They were to rate the color of colored consonant strings on a 6-point rating scale ranging from –3 to 3 with endpoints labeled *saturated red* and *saturated blue*. Both rating scales did not comprise a neutral point. Upon stimulus onset, the appropriate rating scale appeared along with a mouse cursor in the shape of an arrow positioned initially under the middle position of the rating scale. Participants were instructed to use the mouse to click on the point on the rating scale best representing their assessment as fast as possible. They had a 4-s response window for doing so.

The 80 participants were mostly University of Bonn students and a couple of nonstudent volunteers recruited from various sources by the experimenters. Students participated for partial course credit or a small monetary gratification (Euro 5); nonstudent volunteers received Euro 5 for participating.

*Results*

Two participants were excluded from the analyses because they rated flowers to be more negative on average than insects contrary to our a priori assumptions. Response compatibility in the IAT blocks may not have been as intended for these participants. The same pattern of results emerged, however, when the data from these participants were included. There were 0.9% trials without response in the 4-s response window. Figure 2 shows the average response latencies (upper panels) as a function of task mixture and response compatibility for the evaluative and the color ratings. An analysis of variance of the response latencies with these factors and the within-participants factor rating task (evaluative vs. color ratings) revealed the predicted three-way interaction, $F(1, 74) = 2.97$, $p = .09$. Because the interaction had the expected shape (see Figure 2) and because of the equivalence of $t$ tests and $F$ tests with 1 degree of freedom in the numerator (Maxwell & Delaney, 1990), the interaction can be considered significant by a one-tailed test ($p < .05$). Separate analyses of variance revealed that the two-way interaction of rating task and response compatibility was significant given mixed IAT tasks, $F(1, 37) = 4.95$, $p = .03$, but not significant given blocked tasks, $F(1, 37) < 1$, as expected. Separate $t$ tests in the groups with mixed IAT tasks showed the expected longer latencies after incompatible than after compatible IAT blocks in the evaluative ratings, $t(37) = 1.76$, one-tailed $p = .04$, but not in the color ratings, $t(37) = .30$, $p = .77$.

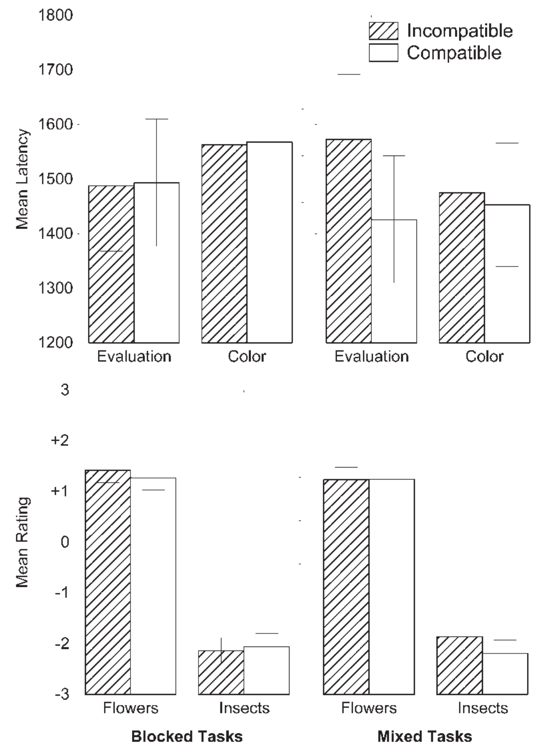


**Figure 2 Mean response latencies for the evaluative rating task and the color-rating task as a function of response compatibility and task mixture (upper panels); mean evaluative ratings of flowers and insects as a function of response compatibility and task mixture (lower panels).**

NOTE: Error bars show 95% confidence intervals.

The lower panels of Figure 2 show the average evaluative ratings for flowers and insects as a function of task mixture and response compatibility. An analysis of variance with these factors and the within-factor valence (flowers vs. insects) revealed a trivial main effect of valence, $F(1, 74) = 1033.43$, $p < .01$, as the only significant effect; all other $F(1, 74) < 1.85$, all $p \geq .19$). Taken together, the same positions on the evaluative rating scales were clicked on after compatible as after incompatible IAT blocks, but given mixed tasks and incompatible IAT blocks, participants took longer to click on these positions.

*Discussion*

Study 2 replicated the results of Study 1 with different tasks and a different response modality. Participants were slower to rate the valence of category-exemplar stimuli after working through incompatible IAT blocks than after compatible IAT blocks, given mixed IAT tasks.

Again, the effect was restricted to the evaluative measure, whereas the color-rating latencies were not affected. The findings thereby suggest that task switches were causally responsible for these aftereffects of IAT blocks and that global shifts in response criteria or control strategies cannot explain them. At the same time, there was little evidence for effects on the evaluative ratings themselves, suggesting that the aftereffects were restricted to effects on attitude accessibility, but there was no substantial effect on attitude extremity.

## GENERAL DISCUSSION

Previous research suggests that working through IAT blocks can affect participants' performance in subsequent IAT phases. Reactivity of this kind is suggested by the well-known effect of compatibility order and by the finding that test-retest reliability of IATs is typically lower than internal consistency (e.g., Asendorpf, Banse, & Mücke, 2002; Bosson, Swann, & Pennebaker, 2000; Dasgupta & Greenwald, 2001; Egloff & Schmuckle, 2002; Greenwald & Farnham, 2000; Steffens & Buchner, 2003).

In two studies, responses in the attribute task and in a closely related rating task, both applied to category-exemplar stimuli, were delayed after incompatible blocks relative to compatible blocks. The effect was obtained even though the response modality was changed from the IAT blocks to the test blocks, suggesting that it reflects differential accessibility of the attribute information of category-exemplar stimuli rather than interference by specific stimulus-response associations established during the IAT blocks. Furthermore, the effect only occurred after IAT blocks with mixed tasks, not when task switches were minimized in conditions with blocked tasks. Thus, task switching is causally involved in the genesis of the aftereffects. The results thereby follow the specific pattern predicted on the basis of the task-set switching account. Although attribute accessibility was affected, there was no effect on ratings of attribute extremity. Thus, the present experimental procedures may be seen as a new and unobtrusive method of manipulating attitude accessibility.

The aftereffects reported in this article are predicted by the task-set switching account. Can alternative accounts explain them as well? Aftereffects are readily predicted from the account by Brendl et al. (2001). Brendl et al. assume, among other things, that incompatible IAT blocks induce a more conservative global response criterion than compatible IAT blocks. To assess this possibility, a color-naming control task was introduced in the present studies. If aftereffects reflect differences in global response criteria or control strategies (Mayr & Liebscher, 2001), aftereffects also should be seen in the control task. Contrary to this expectation, aftereffects were restricted to the attribute task in the present studies, rendering a role of global control strategies or response criteria unlikely. De Houwer (2001) proposed an account by stimulus-response compatibility. In his view, response keys are unambiguously associated with an evaluative or semantic meaning in compatible IAT blocks but carry conflicting evaluations or meanings in incompatible IAT blocks. The ensuing differences in stimulus-response compatibility cause the IAT effect according to De Houwer's account. Such differences cannot explain the present findings, however, because response modality changed from IAT block to test block, rendering previously acquired response associations irrelevant. Finally, the account by figure-ground asymmetry (Rothermund & Wentura, 2001, in press) does not suggest any specific expectation for aftereffects.

The present findings immediately elicit a couple of new questions, some of which are considered in the following paragraphs.

### *What Is the Role of the Response Modality?*

Aftereffects were found despite changes in the response modality, demonstrating some amount of generality of the effects. Yet, it seems likely that such effects would be even more pronounced within the same response modality as Mayr (2001) found for aftereffects in another task-switching context. If so, aftereffects might be even more important for the explanation of reactivity effects of IATs than suggested by the present studies because reactivity is typically observed within the same response modality. Testing this expectation is, however, far from trivial because care has to be taken to avoid trivial carry-over effects and confoundings stemming from the response mapping realized in the IAT block and its precise relation to the response mapping realized in the test block.

### *What Stimulus Classes Are Affected and What Is the Nature and Locus of the Aftereffects?*

Another question concerns the locus of aftereffects. Are the aftereffects restricted to "old" stimuli, that is, to the stimuli that were presented in the inducing IAT blocks, suggesting a role for episodical learning (Wylie & Allport, 2000) modifying specific stimulus-attribute associations? Or is the attribute information of all category exemplars affected—old ones as well as new ones that were not seen in the IAT blocks, suggesting that category-attribute associations are modified? Even more generally, attribute information per se might be suppressed (Goschke, 2000; Meiran, 2000), implying that aftereffects also should affect performance for attribute stimuli. All of these possibilities suggest that effects should be seen on the category exemplars used in the inducing IAT blocks. This is one reason why we focused on these stimuli in the present first demonstration of aftereffects, but it would be interesting to investigate in a series of appropriate experiments whether the aftereffects extend to new category exemplars and to old and new attribute stimuli. If the aftereffects are restricted to stimuli that appear in a prior IAT block, a simple way of mitigating, for example, compatibility-order effects would be to use different stimuli in subsequent blocks.

The nature and locus of inhibitory aftereffects has been an important question in the so-called negative priming paradigm (Fox, 1995). Negative priming research has demonstrated that attending to a particular feature (e.g., the category membership) of a stimulus while having to ignore another feature of that same stimulus (e.g., its valence) in trial $n-1$ can have a detrimental effect on processing the previously ignored feature in the subsequent trial $n$. There is evidence that this sequential effect generalizes to all stimuli that share the ignored feature and that there is an additional specific deficit in processing the particular feature of the particular stimulus that had been previously ignored (De Houwer, Rothermund, & Wentura, 2001).

*How Long-Lived Are the Effects and How Quickly Do They Build Up?*

The present data suggest that the effects survive 24 test-block trials; in fact, when we split the test block into the first 12 and the second 12 trials, the factor test-block half (first vs. second) did not enter into significant effects when the above-reported analyses of variance were repeated. Similarly, contrasting the first two pairs of (a) IAT block and (b) test block with the last 10 such pairs, no evidence for a gradual build-up was found, suggesting that the aftereffects were already in force after two short IAT blocks of 24 trials each.

*Conclusion*

What are the implications of the present theory and findings for practical work with IATs? First of all, the findings offer an explanation of undesirable compatibility-order effects and related reactivity effects. If the aftereffects contain a stimulus-specific component, as suggested by the work on negative priming, a simple way to mitigate reactivity is to use different stimuli in different blocks. Furthermore, the present theory and empirical findings contribute to one's understanding of what is measured by IATs. We argued in the introduction that accessibility of the attribute connotations of target-category exemplars is an important determinant of performance in compatible and incompatible IAT blocks. High accessibility should raise performance in the compatible block, but it should constitute a hindrance in incompatible blocks, leading to large IAT effects. Conversely, low accessibility should simply level IAT effects. The present work demonstrated that working through IAT blocks has specific effects on the accessibility of the attribute connotations of the exemplars of the target categories but little or no effect on ratings of attribute extremity. Taken together, these arguments and findings suggest that the size of the IAT effect is relatively closely tied to the accessibility of the attribute information. The IAT effect may therefore be sensitive to differences in the accessibility of the attribute information much more than to differences in the attribute extremity of the target categories or their exemplars. This line of reasoning explains (a) less than perfect correlations with explicit measures of attribute extremity even when measurement error is taken into account (another cause being the already-mentioned component of method variance) and (b) better success of IATs at predicting spontaneous rather than controlled aspects of behavior (Fazio, 1995).

A reviewer suggested that there might be different groups of IATs. One group, of which attitude IATs are a member, allows for recoding of the attribute and the categorization task into a single dimension such as valence. Another group of IATs does not permit a simple recoding of the two tasks (e.g., IATs with the attribute task to discriminate words from nonwords, i.e., with lexical decisions as the attribute task). In this group, salience asymmetries (Rothermund & Wentura, 2001, in press) play a major role according to the reviewer. It is an open and empirical question whether aftereffects also would be observed for this second kind of IAT. Note, however, that the present account assumes that aftereffects, if they occur, affect the accessibility of that dimension or feature that is shared by target concepts and attribute stimuli and used to simplify the IAT tasks in the compatible block. This will usually be the dimension that is relevant in, and made salient by, the attribute task; for example, in attitude IATs, it is valence. Depending on the attribute and categorization task, the relevant dimension can be almost any other dimension, such as the size of the stimuli or their color (Mierke & Klauer, 2003). For example, to perform speeded lexical decisions between words and nonwords, participants rely on the large difference in familiarity between words and nonwords (e.g., Ratcliff & McKoon, 1988). If lexical decisions constitute the attribute task (e.g., Brendl et al., 2001; Rothermund & Wentura, 2003), it is therefore likely that familiarity will become the dimension along which a simplification of the IAT tasks is attempted. That is, participants might attempt to discriminate not only words from nonwords but also exemplars of the more familiar target category from exemplars of the less familiar target category on the basis of familiarity differences. In this case, the present account and the account by Rothermund and Wentura (in press) concur to predict that asymmetries in familiarity (or salience) drive the IAT effect. In particular, in cases where salience asymmetries are exploited to simplify the IAT tasks in compatible blocks, we expect aftereffects for the speed with which judgments of salience or familiarity can be made.

A number of recent studies have shown that IAT effects are malleable (e.g., Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001). For example, IATs indicated reduced implicit stereotyping and prejudice as a consequence of exposure to, or practice with, counterstereotypical material. The present studies looked at the opposite direction of these cause-effect relationships by assessing whether working through IAT blocks itself acts as an experience that modulates performance in other tasks tapping stimulus-attribute associations. In particular, the IAT blocks operationalized the independent variable rather than the dependent variable in the present work. A specific aftereffect was predicted and found that can tentatively be described as an effect on the accessibility of the attribute information: Working through an attitude IAT affected the ease of retrieving one's very own attitudes. Turning back to the malleability effects, the size of the IAT effect itself was argued to depend on the accessibility of the attribute information. Although some of the malleability effects probably reflect effects of context on the construal of the stimuli (Mitchell, Nosek, & Banaji, in press), we believe it likely that attitude accessibility can be an important mediator in other malleability effects. In concluding, we submit the speculation that the common mediator of many malleability effects as well as the present aftereffects is attribute accessibility of the category-exemplar stimuli.

## NOTES

1. This account assumes that category exemplars frequently share the attribute connotation of the superordinate category. For example, most flower exemplars are usually positively evaluated, as is the category of flowers itself. We acknowledge, however, that the construal of the category, and thereby its attribute connotation, may depend on the subset of used category-exemplar stimuli (Nosek, Greenwald, & Banaji, 2003). Conversely, the immediate construal of the category-exemplar stimuli is influenced by their (highly activated and task-relevant) category memberships so that their immediate attribute connotation in the IAT-task context may be assimilated to that implied by their category membership (Mitchell, Nosek, & Banaji, in press). Both of these influences will converge on rendering the above assumption fulfilled in most contexts.

2. In the groups with blocked tasks, response latencies are elevated in both tasks following compatible blocks relative to incompatible blocks according to Figure 1, whereas error rates are reduced, but the effect of compatibility order did not approach significance in either case, $F(1, 21) = 1.27$, and $F < 1$, respectively.

## REFERENCES


Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421-452). Hillsdale, NJ: Lawrence Erlbaum.

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

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.

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

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

Celex. (1995). *The Celex lexical database, Release 2.* Nijmegen: Center for Lexical Information.

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.

De Houwer, J. (2001). A structural and process analysis of the implicit association test. *Journal of Experimental Social Psychology, 37*, 443-451.

De Houwer, J., Rothermund, K., & Wentura, D. (2001). Stimulus-feature specific negative priming. *Memory & Cognition, 29*, 931-939.

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

Fazio, R. H. (1989). On the power and functionality of attitudes: The role of attitude accessibility. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 153-179). Hillsdale, NJ: Lawrence Erlbaum.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247-283). Mahwah, NJ: Lawrence Erlbaum.

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

Fox, E. (1995). Negative priming from ignored distractors in visual selection: A review. *Psychonomic Bulletin and Review, 2*, 145-173.

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task-set switching. In S. Monsell & J. Driver (Eds.), *Attention & performance XVIII: Control of cognitive processes* (pp. 333-355). Cambridge, MA: MIT Press.

Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022-1038.

Greenwald, A. G., 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.

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

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analysing data: A model comparison perspective.* Pacific Grove, CA: Wadsworth.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging, 16*, 96-109.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129*, 4-26.

Mayr, U., & Liebscher, T. (2001). Is there an age deficit in the selection of mental sets? *European Journal of Cognitive Psychology, 13*, 47-69.

McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, *20*, 481-506.

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

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research*, *63*, 234-249.

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.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (in press). Contextual variations in implicit evaluations. *Journal of Experimental Psychology: General*.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2003). *Understanding and using the implicit association test: II. Methodological issues*. Unpublished manuscript, University of Virginia at Charlottesville.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385-408.

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

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the implicit association test (IAT). *Zeitschrift für Experimentelle Psychologie*, *48*, 94-106.

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

Steffens, M. C., & Buchner, A. (2003). Implicit association test: Separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology*, *50*, 33-48.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs." *Psychological Research*, *63*, 212-233.

Received March 8, 2004
Revision accepted May 6, 2004