ARTICLE IN PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Experimental Social Psychology xxx (2005) xxx–xxx

Journal of Experimental Social Psychology

www.elsevier.com/locate/jesp

Do features of stimuli influence IAT effects? ☆

Matthias Bluemke *,[1], Malte Friese [1,2]

Psychological Institute, University of Heidelberg, Hauptstrasse 47-51, 69117 Heidelberg, Germany

Received 7 June 2003; revised 24 November 2004

Abstract

The Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) is a categorization task intended to measure the strength of associations between concepts. The present research investigated the influence of individual stimuli on IAT effects. Exploring implicit attitudes of East and West Germans, we systematically manipulated relatedness of target stimuli to the attribute dimension and, simultaneously, relatedness of attribute stimuli to the target dimension. Two experiments demonstrate the influence of stimulus associations as one source that drives IAT effects. Depending on the strength and the direction of these cross-category associations, the result was either stronger IAT effects or a decline of IAT effects. Implications for theoretical models as well as for the interpretation of IAT effects are discussed.

Keywords: IAT; Implicit associations; Attitudes; Cross-category associations; Internet

Introduction

In the space of a few short years, many psychologists and researchers all over the world have embraced a new response latency-based measure for the investigation of automatic associations, especially implicit attitudes—the Implicit Association Test (IAT) as first published by Greenwald et al. (1998). The IAT has been used to explore a wide array of topics: for example, evaluative differences between social groups (Greenwald et al., 1998), self-concepts of genders (Greenwald & Farnham, 2000), implicit ageism and racism (Ottaway, Hayden, & Oakes, 2001; Rudman, Greenwald, Mellott, & Schwartz, 1999), and implicit attitudes towards homosexuality (Banse, Seise, & Zerbes, 2001).

Apart from the measure's good psychometric properties (Cunningham, Preacher, & Banaji, 2001), its validity has been empirically demonstrated in a broad range of fields, from nonverbal discrimination in racial interactions (McConnell & Leibold, 2001), to fear-related automatic associations in phobia patients (Teachman, Gregg, & Woody, 2001), gender interactions of shy people (Asendorpf, Banse, & Muecke, 2002), and consumer brand choices (Friese & Wänke, 2005).

Assessing associations of concepts with the IAT

The purpose of the IAT is to measure associations between concepts of interest, namely two target categories (e.g., *Black* and *White*) and two attribute categories (e.g., *pleasant* and *unpleasant*). For each category, a researcher chooses several exemplars. For example, well-known people like Michael Jordan or Charles Manson stand for the target categories, while words like Freedom or Jail denote the pleasant and unpleasant attribute categories.

☆ The authors gratefully acknowledge helpful comments on an earlier draft of this paper from Herbert Bless, Klaus Fiedler, Jan De Houwer, Henning Plessner, and Christian Unkelbach. We especially thank Georg Johann for programming the IAT-Java applet.

* Corresponding author.
E-mail address: Matthias.Bluemke@psychologie.uni-heidelberg.de (M. Bluemke).
[1] Both authors contributed equally to the manuscript.
[2] Present address: Faculty of Psychology, University of Basel, Switzerland.

0022-1031/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2005.03.004

In the course of an IAT, the chosen exemplars serve as stimuli (items) that are presented on a computer screen to a subject who sorts them into their respective categories using two response keys. Initial training trials for the category-response key assignments are followed by two combined blocks in which stimuli of all four categories need to be sorted (double-discrimination task): (a) one compatible block, which allows a subject to use the same response key for Black + unpleasant items, and another key for White + pleasant items; (b) one incompatible block, in which response key assignments are switched to Black + pleasant on one key, and White+ unpleasant on the other key. By convention, the difference that remains after compatible mean response latencies are subtracted from incompatible ones is referred to as the IAT effect. From this difference one infers the strength of association between concepts. If subjects show faster sorting speed in the compatible block than in the incompatible block, a closer association between the concepts *Black* and *unpleasant* than between the concepts *Black* and *pleasant* can be assumed—in other words, a negative implicit attitude is inferred (see Greenwald et al., 1998, for further procedural details).

To date, it has been a matter of debate whether the general concepts (i.e., the categories *Black* and *White*) or the individual stimuli (exemplars like Michael Jordan or Charles Manson) are responsible for the magnitude of IAT effects (De Houwer, 2001; Govan & Williams, 2004; Mitchell, Nosek, & Banaji, 2003; Steffens & Plewe, 2001). The purpose of this article is to examine this methodological issue closely.

Influence of stimuli on IAT effects

The IAT seeks to measure associations between the general categories employed in the task, and not individual associations related to the specific exemplars that represent those categories. Although one might wonder why stimuli should not be important for the formation of IAT effects, no generic rule has been proposed for the selection of these stimuli. When talking about the relative influences of categories and stimuli, one has to clearly distinguish between these two components and their respective evaluations. For example, although a White person may prefer the category White to Black, this subject will evaluate the stimulus Charles Manson negatively because he is a murderer. Thus, the evaluations of the general categories and the individual exemplars representing them are conceptually distinct. In what follows, we will refer to associations between a target stimulus and its category as *congruent* if they share the same general evaluation (e.g., a positively evaluated target stimulus like Princess Diana belonging to the positively seen category White). An association is *incongruent* if the evaluations are at odds (e.g., a negative stimulus like Charles Manson representing a positively evaluated target category White). The logic is similar for the attribute dimension. If an attribute stimulus is associated with a target category, this association is *congruent* if the general evaluation of that target category corresponds to the stimulus' valence (e.g., the negative attribute stimulus poor relates to the negatively evaluated category Black). It is *incongruent* if the evaluations of attribute stimulus and related target category are inconsistent (e.g., the positive attribute stimulus athletic is associated with the negatively evaluated category Black).

Do these different stimulus associations somehow affect IAT scores? Let us briefly present some findings on this issue.

Category influences or stimulus influences?

The evidence for stimulus influence in the IAT is mixed. De Houwer (2001) examined implicit attitudes towards *British* and *foreign* people. Attribute stimuli were either *positive* or *negative* adjectives. However, he subdivided each target stimulus set into positive and negative exemplars, such as Princess Diana (British-positive), Margaret Thatcher (British-negative), Albert Einstein (foreign-positive), and Adolf Hitler (foreign-negative). Thus, for British participants target stimuli were associated in either a congruent (British-positive, foreign-negative) or an incongruent fashion (British-negative, foreign-positive). Consequently, these stimuli not only represented categories typically associated with a certain valence, but each stimulus itself contained valence information that was independent of the valence of the general concept it belonged to. In our terminology, De Houwer reasoned that stimuli holding congruent additional information should increase IAT effects, whereas incongruent stimuli should decrease them. Nevertheless, he found a null effect for target stimulus valence in this within-subjects design, which led him to the conclusion that in a typical IAT only the evaluation of the general categories is important for the magnitude of IAT effects (De Houwer, 2001; footnote 4, p. 450).

In contrast, Mitchell and others (2003, Experiment 2) found influence of target stimuli in two race IATs using the attribute categories *good* and *bad*. The target categories *Black* and *White* were represented by names of well-known exemplars. In one condition, the White exemplars were liked prominent people, whereas the Black exemplars were disliked people. In another condition the authors switched the valence of the exemplars used to represent both categories while keeping the stimuli of the attribute categories constant. After applying both IAT versions in a repeated measurement design, response latencies indicated a general preference for White exemplars over Black exemplars. However, a noteworthy weaker IAT effect ($d = 0.23$) emerged for liked Black (and disliked White) exemplars

compared to disliked Black (and liked White) exemplars ($d = 1.68$).

Govan and Williams (2004) even eliminated the IAT effect in a similar Black/White-IAT. Moreover, they replicated the seminal IAT by using evaluatively congruent target stimuli of the kind that were used by Greenwald et al. (1998) for the target categories *flowers* vs. *insects* and they compared it to an IAT with incongruent stimulus-associations. Specifically, in the latter condition they used positively valenced exemplars for insects (e.g., BUTTERFLY) and negatively valenced exemplars for flowers (e.g., POISON IVY). In the replicated version the original IAT effect was obtained, whereas incongruent exemplars even led to a reversed sign of the IAT effect.

To explain their findings, Mitchell and colleagues (2003) stated that contextual variations determine the evaluation of target stimuli. A positive context, as conveyed by similarly positive target stimuli, may lead to the activation of positive associations, while the very same target-stimulus could be negatively evaluated in a negative context. Govan and Williams (2004) reasoned that restricting the selection of stimuli to atypical exemplars may lead to the activation of a different mental representation or a temporary re-definition of the category (e.g., "*nice* Blacks"). Although a subject may generally dislike Black people, the activation of exemplars belonging to a likeable subtype may lead to a temporarily different evaluation of the category Black.

Steffens and Plewe (2001) started a somewhat different approach. In two gender IATs, they kept stimuli constant for the target categories (*male* vs. *female* first names), but varied attribute stimuli (*pleasant* vs. *unpleasant* adjectives): in one condition, pleasant adjectives were associated with the female stereotype (e.g., BEAUTIFUL), whereas unpleasant adjectives were associated with the male stereotype (e.g., BRUTAL). In the second condition, the stereotypical association of the attribute items was reversed, such that pleasant adjectives bore male associations (e.g., LOGICAL), whereas unpleasant adjectives carried a female association (e.g., HYSTERICAL). If these references of the attribute stimuli, which Steffens and Plewe referred to as *cross-category associations*, did not contribute to the IAT effect, then both conditions should have produced equal IAT effects. Although the resulting implicit attitudes were positive toward women in both conditions, the different stimuli led to a stronger IAT effect for the evaluatively congruent feminine-positive and masculine-negative condition than for the incongruent feminine-negative and masculine-positive condition.

In sum, research has shown that properties of stimuli change the magnitude of IAT effects at least under some conditions, thus indicating that IAT effects are not driven solely by the association of concepts, but also by the stimuli representing these concepts. However, are both target and attribute stimuli generally capable of changing IAT effects? And what theoretical implications for the interpretation of IAT effects arise from these phenomena?

Our research systematically examines the effects of the cross-category associations of stimuli. Instead of locating stimulus effects in different *mental representations* of the participant, as the explanations of context variability (Mitchell et al., 2003) and category re-definition (Govan & Williams, 2004) assume, cross-category associations offer a rather procedural explanation for stimulus influence on IAT effects: depending on the conceptual overlap of IAT stimuli with another category, the relative *ease of the sorting process* changes.

Fig. 1 may illustrate the logic: the four objects in one rectangle symbolize the categories and their stimuli in an IAT—the circles are the targets, the squares the attributes. The size of objects symbolizes a conceptual association between two categories. For example, large objects may share positive valence (e.g., the categories White + pleasant, to stick to the example), whereas small objects may share negative valence (e.g., Black + unpleasant). In the upper panel, classification is facilitated if large objects (associated categories) are mapped onto the same response key (e.g., White + pleasant). This compatible assignment will lead to shorter response latencies than will an incompatible assignment when unassociated categories share one response key (e.g., White + unpleasant). This association of concepts is what the IAT intends to measure. In the lower panel, the shading information represents cross-

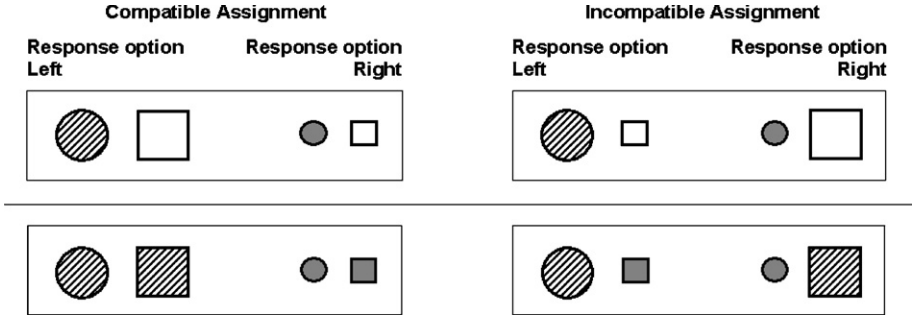

Fig. 1. Symbolic representation of IAT sorting task. The upper panel symbolizes a compatible and incompatible assignment of categories represented by circles and squares. The lower panel illustrates how cross-category associations of stimuli influence the sorting process in an IAT. (Figure adopted from Fiedler, Messner, & Bluemke, 2003).

category associations between stimuli and another category. Although it is not essential to solving the categorization task that size and shade (conceptual association and congruent cross-category associations) go hand in hand, if they do it should be even easier to find the correct responses than in the upper panel (e.g., the positively evaluated stimulus PRINCESS DIANA representing the target category White). In the incompatible assignment the solution should be inhibited, because size and shade do not correspond (incongruent cross-category associations; e.g., the negatively evaluated stimulus CHARLES MANSON representing the target category White).

Experiment 1

The first experiment was intended to directly test the influence of cross-category associations on IAT effects. Specifically, we manipulated the degree and direction of additional stimulus information between target stimuli and attribute categories, as well as between attribute stimuli and target categories in a between-subjects design.

To that end, we conducted an Internet-study to assess implicit attitudes of West German participants towards people in West and East Germany (the former German Democratic Republic). Recent research has revealed implicit preferences for one's own group among both East and West Germans (Kuehnen et al., 2001).

To test stimulus influence in the IAT, the concepts *East* and *West* served as target categories, while the concepts *positive* and *negative* constituted the attribute dimension. However, we created seven different versions of an evaluative East–West German IAT by using different stimulus sets while keeping procedural aspects constant.

Overview and hypotheses

We started out with a control condition, that is, an IAT that consisted entirely of stimuli without any cross-category associations with regard to both the target and attribute categories. This established a baseline for all further versions (see Appendix A for a complete list of stimuli). Hence, target stimuli were non-valenced (neutral, e.g., BAVARIA), and attribute stimuli were unrelated to the target dimension (e.g., FRIENDLY). Throughout this paper, we will indicate the direction of cross-categorical associations by indexing it to the category labels. Thus, the control condition consisted of $\text{West}_{\text{neutral}}$ and $\text{East}_{\text{neutral}}$ target stimuli besides $\text{positive}_{\text{unrelated}}$ and $\text{negative}_{\text{unrelated}}$ attribute stimuli. Our first hypothesis was that West German participants would show a typical ingroup preference in the IAT (Kuehnen et al., 2001). We expected participants to react faster in compatible blocks (when positive and West shared the same response key) than in incompatible blocks (when negative and West shared one key), resulting in a standardized effect size *d* above zero (Cohen, 1977).

Moreover, we conducted three IATs using item sets with congruent cross-category associations (from a West German perspective): (a) A first version used *pro-West attribute stimuli* ($\text{positive}_{\text{West}}$, $\text{negative}_{\text{East}}$), but kept the target stimuli the same as in the control condition. This version comprised positive adjectives that referred to positive aspects of the West German stereotype (e.g., SUCCESSFUL), and negative adjectives that referred to negative aspects of the East German stereotype (e.g., XENOPHOBIC). (b) A second version used *pro-West target stimuli* ($\text{West}_{\text{positive}}$, $\text{East}_{\text{negative}}$), but kept the attribute items unchanged in comparison to the control condition. Therefore, West German target stimuli were in addition positively connoted (e.g., NORTH SEA), while East German target stimuli conveyed a negative association (e.g., STASI—the former East German Secret Service). (c) A third version combined both *pro-West target and attribute stimuli* ($\text{West}_{\text{positive}}$, $\text{East}_{\text{negative}}$, $\text{positive}_{\text{West}}$, $\text{negative}_{\text{East}}$).

Likewise, we switched the direction of the cross-category associations, thus creating incongruent item sets from a West German perspective in three further IAT versions: (d) The fourth version contained *pro-East attribute stimuli* ($\text{positive}_{\text{East}}$, $\text{negative}_{\text{West}}$), that is, positive adjectives primarily associated with East Germans (e.g., SOCIABLE), and negative adjectives primarily associated with West-Germans (e.g., IMPERSONAL). (e) The fifth version included *pro-East target stimuli* ($\text{West}_{\text{negative}}$, $\text{East}_{\text{positive}}$), such as negative West German stimuli (e.g., RAF—RED ARMY FRACTION, a terrorist organization in the 1970s) as well as positive East German stimuli (e.g., BALTIC SEA). (f) The sixth version combined both *pro-East target and attribute stimuli* ($\text{West}_{\text{negative}}$, $\text{East}_{\text{positive}}$, $\text{positive}_{\text{East}}$, $\text{negative}_{\text{West}}$). For a summary of all seven conditions, see Table 1.

Our next hypotheses regarding the magnitudes of IAT effects were as follows: If the IAT exclusively measures the association between concepts (Greenwald et al., 1998), stimuli should have no influence on the IAT effect. Any IAT should be equally able to assess the implicit attitude as long as the categorization of stimuli to their respective category remains unambiguous (Greenwald & Nosek, 2001). However, if the cross-category associations of the stimuli do make a difference, we expected the IAT effects to increase in those versions where the stimuli were associated in a congruent manner for West German participants, that is, when either $\text{West}_{\text{positive}}$/$\text{East}_{\text{negative}}$, or $\text{positive}_{\text{West}}$/$\text{negative}_{\text{East}}$ stimuli, or both were applied. To the degree that congruent connotations enable a relatively fast correct response in compatible blocks, they enforce the response conflict in incompatible IAT blocks, thus leading to increased IAT effects compared to the control condition. The opposite

Table 1
Direction of cross-category associations determine IAT effects in a between-subjects design (Experiment 1): mean IAT effects in Z differences and standardized effect size d

IAT-Version	Additional association of				N	α	Mean IAT effect			Mean latencies in ms (SD)		Error rate (%)
	Target stimuli		Attribute stimuli									
	West	East	Positive	Negative			Z difference[a]	SD	d	Compatible	Incompitable	
Pro-East target and attribute stimuli	Negative	Positive	East	West	62	.87	-0.30^{***}	0.37	−0.81	1284 (209)	1148 (217)	7.5
Pro-East target stimuli	Negative	Positive	—	—	57	.74	−0.08	0.37	−0.23	1120 (186)	1089 (184)	7.0
Pro-East attribute stimuli	—	—	East	West	57	.85	0.27^{***}	0.41	0.65	993 (195)	1114 (182)	6.2
Control condition	—	—	—	—	59	.89	0.42^{***}	0.34	1.23	922 (185)	1087 (197)	5.6
Pro-West attribute stimuli	—	—	West	East	67	.92	0.59^{***}	0.34	1.73	922 (196)	1164 (210)	5.4
Pro-West target stimuli	Positive	Negative	—	—	65	.92	0.66^{***}	0.33	2.01	898 (149)	1164 (187)	4.5
Pro-West target and attribute stimuli	Positive	Negative	West	East	62	.94	0.77^{***}	0.28	2.70	931 (170)	1289 (217)	5.0

[a] $^{***}p < .001$

holds for incongruent connotations of stimuli, which is why IAT effects should decrease in those conditions that comprise either $West_{negative}/East_{positive}$, or $positive_{East}/negative_{West}$ stimuli, or both.

Consequently, we expected the strongest deviations from the control condition when both target and attribute stimuli were manipulated in the same direction at the same time, that is, either $West_{positive}/East_{negative}$ in conjunction with $positive_{West}/negative_{East,}$ or $West_{negative}/East_{positive}$ in combination with $positive_{East}/negative_{West}$. In both of these versions, every stimulus contained additional information, while in every other experimental condition only half of the stimuli carried additional associations (the other half remained the same as in the control condition). In the case of the twofold incongruent version, we predicted that the IAT effect would not only decrease, but also turn negative. In other words, if stimuli do have substantial impact, in this latter version West German participants should be faster in incompatible blocks (categories East and positive sharing one key) than in compatible blocks (categories West and positive sharing one key), resulting in a negative sign of the effect size d.

Furthermore, we predicted that the effect size d would deviate more strongly from the control condition when only target stimuli are related to valence as compared to when only attribute stimuli are related to the target dimension. This expectation was caused by the very high salience of the evaluative dimension in target stimuli, compared to the rather absent salience of the East–West dimension in attribute stimuli. To illustrate this difference, consider the following examples: the negativity of the target stimulus STASI (the former East German secret service) is extremely high for all Germans. The relationship of an attribute stimulus like SUCCESSFUL to the East–West dimension is far less salient (Pratto & John, 1991; Rothermund & Wentura, 2001, 2004). Taken together, we expected a continuous pattern of IAT effects as indicated by the order of IAT versions in the first column of Table 1.

Method

Participants

Five hundred and eight West German Internet users completed the experiment. After controlling for sufficient level of language comprehension, undisturbed execution of the IAT, multiple submissions, and informed consent after the experiment, we included 484 participants in our final sample. Participants either received partial course credit, or had the chance to win one of 10 book coupons (worth 15 Euro each). Mean age was 25.6 years ($SD = 6.2$). Participants differed remarkably from laboratory samples in experimental psychology in terms of education, occupation, and other demographic variables, indicating that our sample allows for inferences with greater external validity. After controlling for high error rates and outliers (see results section), we were left with 57–67 participants for each IAT version (see Table 1 for details).

Design and materials

In a between-subjects design, we conducted seven IATs, each of which carried the categories *East*, *West*, *positive*, and *negative*. Table 1 shows the specific manipulation of each IAT as described in the hypotheses section.

A pretest on the Internet with 379 West German participants had identified words that were unambiguous with respect to their category membership. Target stimuli (nouns and proper names) and attribute stimuli (adjectives) were chosen out of a pool of 760 words according to participants' evaluations on 9-point rating scales. After pretesting, we chose 10 stimuli for each category, thus creating stimulus sets with congruent, incongruent, or no cross-category associations (see the Appendix A for a complete list of stimuli and their English translations). To recapitulate, each stimulus was associated in the first place with its proper category, that is, the feature relevant for solving the sorting-task was always clearly recognizable.

We made several arrangements to minimize the chance of any confusion about the stimuli's relevant

dimension when additional associations were conveyed by a stimulus. First, all attribute items were adjectives, while all target items were nouns, names, or geographic locations.[3] Second, in line with Greenwald et al. (1998, Experiment 2), stimuli of the target dimension were presented in capital letters whereas stimuli of the attribute dimension were presented in lower case letters, as were the corresponding categories that were shown in the upper corners of the screen. Third, following De Houwer (2001, footnote 4), target stimuli and their categories were written in a different color than attribute stimuli and their categories. No participant complained about confusion in matching a stimulus with its relevant dimension. Thus, it is reasonable to assume that any emerging stimulus impact cannot be attributed to confusion.[4]

Procedure

Upon entering the URL http://knut.psi.uni-heidelberg.de, participants read a welcome page, which informed them about the duration of the experiment and asked for serious completion of the tasks. On the next page, participants answered several questions concerning their ethnic background, demographic variables, and personal opinions about West and East Germans. The third page informed participants that reaction times would be measured during a task in which they were to sort words appearing in the middle of the screen into one of several categories. Correct responses were to be given as quickly but also as reliably as possible by pressing one of two response keys. A red "X" would indicate incorrect reactions, and the next stimuli would only appear after the correct response had been given. The fourth and final introductory page informed participants that categories would remain in the upper left and upper right corners of the screen, and that stimuli in capital letters were to be sorted into one of the categories presented in capitals.

Each participant completed one IAT consisting of five phases: (1) a practice block for target stimuli only; (2) a practice block for attribute stimuli only; (3) first double-discrimination task; (4) a practice block for target stimuli only, but with reversed key-assignments; (5) second double-discrimination task. The practice blocks consisted of 20 trials and presented each stimulus once. The double-discrimination blocks contained 80 trials, with every stimulus appearing twice. Words were drawn from a stimulus list without replacement in an individually randomized order. The inter-stimulus interval was set at 150 ms. Most procedural variables were held constant following Greenwald et al. (1998, Experiment 2): the order of combined blocks was varied between participants. Stimulus presentation alternated between the target and the attribute dimension in the combined blocks, which also controlled for specific task-set switching costs (Mierke & Klauer, 2001, 2003). The final pages asked for quality of data and provided a careful debriefing as well as some hyperlinks to related Internet pages.

Results

Data preparation

We excluded erroneous responses (7.8%) and values of less than 300 ms or more than 3000 ms (3.4%) from the analysis. The first 10 of the 80 trials in each combined block were considered training trials and dropped. We additionally excluded those participants whose results indicated error rates of more than 20% in either the compatible or the incompatible block (ranging from 2 to 12 participants per condition). To reduce the error variance between participants while conserving the relative differences between the compatible and incompatible block for each participant, we *Z*-standardized participants' data individually (Fiedler & Zogmaister, 2001).[5] Finally, we excluded as outliers those participants whose IAT effects were more than 1.5 interquartile ranges below the first quartile or above the third quartile of their IAT conditions (six participants; box-plots, Tukey, 1977). Additional analyses with absolute response latencies as well as logarithmic transformations did not yield any substantial differences. Also, dropping the first 2 or 20 trials revealed no substantial differences. The IAT procedure displayed rather high internal consistency (Cronbach's α ranging from .74 to .94, cf. Table 1).[6]

[3] This combination was derived from formerly published IAT applications. First, prominent exemplars were used, that is, well-known East or West Germans such as Boris Becker (Tennis star) and Erich Honecker (former president of the GDR). Second, we included names of geographical locations such as Erfurt or Bavaria, similar to successfully applied items in the East/West IAT by Kuehnen et al. (2001). The third class of target items also consisted of unambiguously associated aspects of West and East Germany, such as democracy or communism. A similar type of stimuli was used previously to exemplify heterogeneous concepts such as *smoking* (Swanson, Rudman, & Greenwald, 2001) or *negative/positive outcomes* in social situations (De Jong, Pasman, Kindt, & van den Hout, 2001). We thus intended the activation of a broad conceptual representation of both German groups.

[4] Although in Experiment 1 no participant reported any confusion in sorting the stimuli, we cannot completely rule out that the similarity of a few target and attribute stimuli (e.g., Arrogance and arrogant) might have augmented stimulus effects. Therefore, in Experiment 2 we employed a stimulus set with significantly reduced similarities.

[5] Each participant's latency data were transformed into standardized *Z*-scores item-wise by subtracting the participant's mean of combined blocks from an item response time, and dividing the result by the individual standard deviation of both combined blocks (Rosenthal & Rosnow, 1991). A similar procedure was recently proposed by Greenwald, Nosek, and Banaji (2003), who additionally recommend the use of calculating the standard deviations by including practice blocks, plus the use of error penalties.

[6] For each IAT in Experiment 1, Cronbach's α was determined by using a 7-item-scale. Items represented the mean differences between the incompatible and compatible blocks averaged across every 10 trials.

Stimulus influence

Fig. 2 shows the standardized effect sizes d for the response latency differences between compatible and incompatible blocks based on Z values. The predicted positive IAT effect resulted for the control condition comprising $West_{neutral}$, $East_{neutral}$, $positive_{unrelated}$, and $negative_{unrelated}$ stimuli (see also Table 1). A paired t test revealed a significant IAT effect in the expected direction, with a standardized effect size of $d = 1.23$.

We will first look at the analysis of IAT versions comprising congruent cross-category associations. When applying $positive_{West}$ and $negative_{East}$ attribute stimuli, the IAT effect reached a higher value ($d = 1.73$). As predicted, the IAT effect was even higher for $West_{positive}$ and $East_{negative}$ target stimuli ($d = 2.01$). The highest IAT effect was obtained when both target and attribute stimuli were associated in a way that favored West Germans and derogated East Germans ($d = 2.70$) (Fig. 2).

What if the incongruent cross-category associations favored East Germans instead? Would IAT effects still suggest ingroup favoritism or outgroup derogation, respectively? Compared to the control condition, a sharp decline in the IAT effect size occurred when $positive_{East}$ and $negative_{West}$ attribute stimuli were applied ($d = 0.65$). What is more, this decline was even stronger—and led to an inverted sign of the IAT effect—when $West_{negative}$ and $East_{positive}$ target stimuli were used ($d = -0.23$). In line with our predictions, when attribute plus target stimuli were simultaneously manipulated, there was a very sharp decline of the IAT effect ($d = -0.81$).[7] The differences in IAT effects between the seven IAT versions are reflected in a 2 (Order of Blocks) × 7 (IAT version) ANOVA that revealed a highly significant main effect for the seven IAT versions, $F(6,415) = 81.86$, $p = 10^{-67}$. Almost all planned contrasts revealed significant differences between the seven IAT versions, at least $p < .05$ (two-tailed). Two direct comparisons yielded results that only tended to be of the hypothesized directions: first, comparing the IAT version containing $West_{positive}$ and $East_{negative}$ target stimuli with the IAT version containing $West_{positive}$ and $East_{negative}$ target stimuli plus $positive_{West}$ and $negative_{East}$ attribute stimuli ($p = .11$), and second, comparing the IAT version applying $positive_{East}$ and $negative_{West}$ attribute stimuli with the IAT version applying $West_{positive}$ and $East_{negative}$ target stimuli ($p = .22$).

We also found a main effect for the control factor order of blocks ($F(1,415) = 13.09$, $p = .0003$). This effect reflects a tendency toward lower IAT effects in the compatible-first condition (ds range from −1.25 to 2.25), but higher values in the incompatible-first condition (ds range from −0.48 to 3.40), and might be attributable to the rather low number of training trials (Greenwald et al., 2003). This order effect, however, was not qualified by an Order × IAT version interaction ($F(6,415) = 1.42$, $p = .20$). Mean error rates did not exceed 7.5% for any IAT version (see Table 1 for mean response latencies and further details).

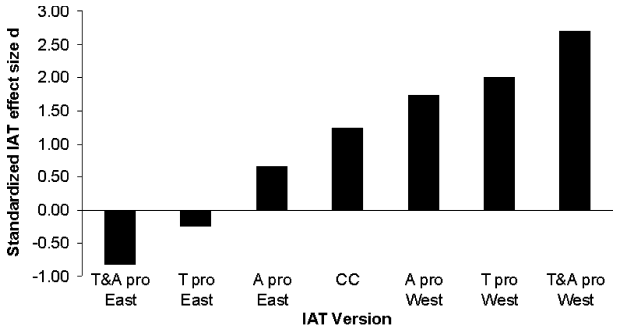

Fig. 2. IAT effect size d as a function of the direction of cross-category associations in a between-subjects design (Experiment 1). T-denoted and A-denoted versions refer to additional target or attribute item associations, respectively. Pro-West-denoted and pro-East-denoted versions refer to stimuli favoring West and East Germans, respectively. Positive d values indicate faster sorting-speed in compatible blocks than in incompatible blocks. Negative values indicate a reversed pattern of response latencies.

Discussion

The presented data clearly demonstrate the influence that individual stimuli can have on IAT effects. Associating stimuli systematically with cross-categorical references led to the predicted pattern of IAT effects within a wide variation of effect sizes. As predicted, the control condition resulted in a positive IAT effect, suggesting an implicit ingroup preference by West German participants. Moreover, stimuli with congruent cross-category associations led to significantly stronger effect sizes. By contrast, incongruent cross-category associations substantially reduced effect sizes, even down to a change of signs. Our expectations were met in that versions with cross-category associations of target stimuli deviated more strongly from the control condition than did versions with cross-category associations of attribute stimuli, while the combined versions produced the strongest divergence from the control condition.

To explain our results differently, one could argue that participants became aware of the manipulation. This could have resulted in a systematic dropping out of those participants who did not comply with the tone of their IAT version and would have left those who did. However, participants who completed the seven IAT versions did not differ in their explicit attitudes towards East and West Germans assessed prior to the IAT. Neither did they differ from those participants who dropped out of the experiment.

[7] All t tests aggregated over the control factor order of blocks (compatible vs. incompatible first) indicated statistically significant differences between compatible and incompatible blocks, $p < .001$, except for the IAT applying $West_{negative}$ and $East_{positive}$ target stimuli ($p = .08$).

Category labels or stimuli?

A reasonable amount of recent research about the IAT dealt with the question of what drives the IAT effect—the categories or the stimuli? Our results demonstrate the impact not only of the stimuli, but clearly that of the categories, as well. A neutral IAT with non-valenced target stimuli and target-unrelated attribute stimuli produced a strong effect, which is hard to explain by other properties of the stimuli alone. This "socket-effect" is evident in all conditions, as the absolute value of IATs favoring West Germans was much higher than in those versions favoring East Germans. Yet, in our experiment, stimulus influence was extremely powerful, overriding the impact of the categories and even changing the sign in two versions.

In keeping with the interpretation of IAT effects as an indicator of implicit attitudes, our results range from very strong ingroup preference to strong outgroup preference of West German participants. Obviously, this reasoning alone cannot explain the resulting pattern of IAT effects. The seven IATs were not equally able to assess participants' general implicit attitudes, despite the fact that the categories and procedural variables remained constant across all IAT versions.

How do our findings relate to De Houwer's (2001) results denying any influence of target stimuli's cross-category associations? Let us consider some respects in which the present research differs from De Houwer's and how these differences might explain his null-finding: (a) the power for testing stimulus influence with a three-way interaction and interpreting this null-finding may simply have been too low; (b) target items in his study were highly specific (only individual persons as category exemplars), while most of our items were related more generally to their respective concepts; (c) De Houwer used balanced item sets in a within-subjects design, that is, each participant was confronted with both positive and negative British exemplars (e.g., Queen Mother, Margaret Thatcher) as well as positive and negative foreign exemplars (e.g., Mahatma Gandhi, Adolf Hitler) in the same IAT. Given that we applied a between-subjects manipulation, it could be that cross-category associations of stimuli may be recognized only in IATs with consistently associated item sets.

The last point, especially, is theoretically promising and deserving of further study. To strengthen the conclusiveness of Experiment 1 regarding the influence of cross-category associations, we sought to replicate our results in a within-subjects design.

Experiment 2

Similar to De Houwer (2001), we used an IAT with both target and attribute stimuli being manipulated with congruent, incongruent, or no cross-category associations. Thus, target stimuli consisted in equal number of positive, neutral, and negative East and West German words. Likewise, positive and negative attribute items carried a West German, an East German, or no German connotation at all. To be clear, the target stimulus set was composed of $\text{West}_{\text{positive}}$, $\text{West}_{\text{neutral}}$, $\text{West}_{\text{negative}}$, as well as $\text{East}_{\text{positive}}$, $\text{East}_{\text{neutral}}$, and $\text{East}_{\text{negative}}$ items, and the attribute stimulus set was composed of $\text{positive}_{\text{West}}$, $\text{positive}_{\text{unrelated}}$, $\text{positive}_{\text{East,}}$ as well as $\text{negative}_{\text{East}}$, $\text{negative}_{\text{unrelated}}$, and $\text{negative}_{\text{West}}$ items. Note that the unit of analysis to test stimulus influence changes from different IAT versions in Experiment 1 to different subsets of stimuli in only one IAT in Experiment 2.

Our hypothesis can be summarized as follows: if either target or attribute stimuli carry cross-category associations in favor of West Germans, these stimuli should produce enhanced IAT effects, whereas cross-category associations in favor of East Germans should reduce them for the respective stimuli. Moreover, stimuli without additional associations should lie in the middle of the other conditions, thus forming a neutral point of reference. Note that balancing stimulus subsets should create an enormous amount of response latency variance because the type of cross-category association varies per trial. By contrast, in Experiment 1 these additional associations remained constant for all stimuli of one category during the complete IAT assessment. Therefore, we expected reasonably smaller effect sizes in Experiment 2.

Method

Participants

One hundred and twenty-two participants completed the experiment. After excluding 16 participants because of poor quality of data as a result of low language comprehension, multiple submissions, disturbances during participation, or a refusal to provide informed consent afterwards, we were left with 106 participants for analysis. Age ranged from 16 to 63 years ($M = 27.2$, $SD = 10.2$). After completing the experiment, participants were thanked and debriefed.

Materials and procedure

The Web pages were held constant to Experiment 1. The pages prior to the IAT contained background information on the study, assessment of socio-demographic data, explicit evaluations of East and West Germans, as well as the instructions for the IAT.

Twelve stimuli representing each category (four items per stimulus subset) were presented twice, adding up to 96 items in each combined block (bold items in the Appendix A). All relevant procedural variables were held constant with Experiment 1, except for practice blocks consisting of 24 items. Once again, target and attribute items were presented in distinct letter case and color.

Table 2
Mean IAT effects depending on direction of cross-category associations in a within-subjects design (Experiment 2), expressed as Z differences and standardized effect size d

Condition	N	Target stimuli				Attribute stimuli			
		Z difference[a]	SD	d	Error rate (%)	Z difference[a]	SD	d	Error rate (%)
Pro-East	70	0.05	0.62	0.08	15.0	0.12	0.63	0.19	13.7
Control condition	70	0.18**	0.47	0.39	14.5	0.18**	0.56	0.32	12.9
Pro-West	70	0.21***	0.44	0.48	17.0	0.37***	0.60	0.61	16.3

[a] $^{**}p < .01$, $^{***}p < .001$.

Results and discussion

Data preparation

We applied the same standardizing algorithm as in Experiment 1 (see Footnote 5). Erroneous responses (14.9%) and values below 300 ms or above 3000 ms (3.1%) were recorded as missing values. Due to randomly changing cross-category associations of target *and* attribute stimuli, choosing the correct response key was more difficult for participants in this design and led to a higher number of errors, whereas the amount of overly short and long response times remained comparable. We excluded participants with error rates of more than 20% in either the compatible or the incompatible block, leaving a total of 70 participants for the final sample.[8] The IAT displayed an internal consistency of Cronbach's $\alpha = .77$.[9]

Analysis of target stimuli

IAT effects (cf. Table 2) were computed as in Experiment 1. Unless indicated otherwise, IAT effects differed significantly from zero at $p < .01$, at the least. We analyzed IAT effects of target stimuli by conducting a 2 (Order of Blocks: compatible vs. incompatible first) × 3 (Stimulus Association: pro West vs. neutral vs. pro East) mixed ANOVA with repeated measurement on the last factor. A stimulus association main effect indicated different magnitudes of IAT effects for the three stimulus subsets, $F(2, 136) = 3.72$, $p = .03$ (see also Fig. 3). In line with our hypothesis, IAT effects were more extended in the *pro-West* condition ($d = 0.48$) than in the *neutral* control condition ($d = 0.39$). Moreover, a decline to almost zero occurred in the *pro-East* condition ($d = 0.08$, $p = .53$). The resulting difference between the *pro-West* and *pro-East* conditions was significant ($t(69) = 2.27$, $p = .03$). Whereas the difference between the *pro-East* condition and the *neutral* condition was clearly recognizable ($t(69) = 1.70$, $p = .09$), the difference between the *pro-West* condition and the *neutral* condition—although in the predicted direction—was too small to reach significance ($t < 1$). Thus, in contrast to De Houwer (2001), we clearly demonstrated the potential influence that cross-category associations of target stimuli can have on IAT effects even in a balanced within-subjects design.[10]

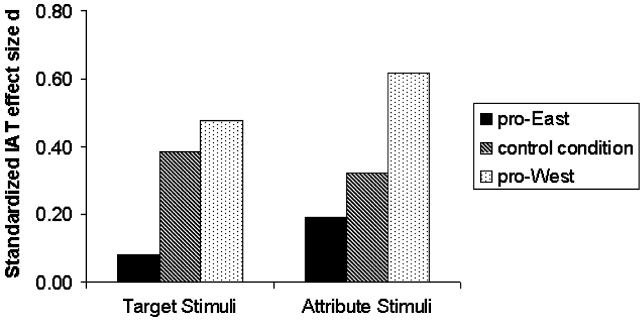

Fig. 3. IAT effect size d as a function of the direction of cross-category associations in a within-subject design (Experiment 2), separated for target and attribute stimuli. Cross-category associations favoring East Germans decrease IAT effects, while cross-category associations favoring West Germans extend IAT effects.

We obtained a marginally significant main effect for the factor order of blocks, $F(1, 68) = 3.69$, $p = .06$, reflecting higher IAT magnitudes in the *incompatible-first* condition compared to the *compatible-first* condition. As the differences between the stimulus subsets were more pronounced in the *compatible-first* ($ds = -0.27$, 0.26, 0.42) than in the *incompatible-first* condition ($ds = 0.38$, 0.49, 0.58), an Order of Blocks × Target Stimulus Association interaction resulted, $F(2, 136) = 3.87$, $p = .02$.

[8] In line with McFarland and Crouch (2002), we attribute higher error rates to the increased cognitive load due to longer stimulus lists and increased task-set switching (Mierke & Klauer, 2001, 2003). Note that our participants encountered more stimulus variation than did the participants in the De Houwer (2001) study, in which only target stimuli, but not attribute stimuli were manipulated. Moreover, while De Houwer applied only two types of stimuli, we applied three (*pro-West*, *pro-East*, *neutral*).

[9] Cronbach's α was calculated by using a 7-item-scale averaging across every 14 trials. Keeping the number of averaged trials instead of scale length comparable to Experiment 1 led to the same result.

[10] For a comparable analysis of our results with De Houwer's (2001), only those target stimuli which bore cross-category associations were submitted to a 2 (Block: compatible vs. incompatible) × 2 (Order: compatible vs. incompatible first) × 2 (Concept: East vs. West German) × 2 (Stimulus Association: positive vs. negative) mixed ANOVA. This analysis, again, yielded a clear impact of target stimulus associations which was reflected in a three-way-interaction between Block, Concept, and Stimulus Association ($F(1, 68) = 11.95$, $p = .0009$)—a result that failed to show up in De Houwer's (2001) analysis ($F(1, 26) < 1$).

Analysis of attribute stimuli

Conducting an identical 2 (Order of Blocks: compatible vs. incompatible first) × 3 (Stimulus Association: pro West vs. target-unrelated vs. pro East) mixed ANOVA on IAT effects of attribute stimuli (cf. Table 2), we found the expected main effect of stimulus association, $F(2,136) = 5.01$, $p = .008$ (cf. Fig. 3). IAT effects were higher in the *pro-West* condition with positive attribute stimuli conveying West German associations and negative attribute items bearing East German connotations ($d = 0.61$) than in the *target-unrelated* control condition ($d = 0.32$). In contrast, a lower IAT effect resulted in the *pro-East* condition when stimuli favored East Germans ($d = 0.19$, $p = .11$). A significant difference emerged for the *pro-West* vs. *pro-East* condition ($t(69) = 3.23$, $p = .002$), as it did for the comparison of the *pro-West* vs. the *target-unrelated* condition ($t(69) = 2.54$, $p = .01$), but not for the comparison of the *pro-East* vs. the *target-unrelated* condition ($t < 1$).

The order of blocks main effect, $F(1,68) = 16.82$, $p = .0001$, revealed higher IAT effects in the *incompatible-first* (mean $d = 0.71$) than in the *compatible-first* condition (mean $d = 0.02$), which—in contrast to the target stimuli analysis—was not qualified by a two-way interaction with the factor stimulus association ($F < 1$).

The results of Experiment 2 clearly demonstrate the influence of cross-category associations in a within-subjects design. Both target and attribute stimulus associations modulated the magnitude of IAT effects in the predicted way.[11]

[11] One could suspect that the effects depend on our reduced sample, because we lost one-third of the participants due to high error rates. Therefore, we redid the same analysis with all participants included, regardless of error rates. The results were essentially confirmed. We analyzed IAT effects of target stimuli by conducting a 2 (Order of Blocks: compatible vs. incompatible first) × 3 (Stimulus Association: pro West vs. neutral vs. pro East) mixed ANOVA with repeated measurement on the last factor. Again, a stimulus association main effect emerged, $F(2,208) = 3.14$, $p = .046$. The difference between the *pro-West* and *pro-East* conditions was significant ($t(105) = 2.22$, $p = .03$), and the difference between the *pro-East* condition and the *neutral* condition was marginally significant ($t(105) = 1.95$, $p = .054$). The difference between the pro-West condition and the neutral condition was too small to reach significance ($t < 1$). Thus, the pattern perfectly matches the one we obtained with the reduced sample. An identical 2 (Order of Blocks: compatible vs. incompatible first) × 3 (Stimulus Association: pro West vs. neutral vs. pro East) mixed ANOVA of attribute stimuli yielded a main effect of stimulus association that only approached significance, $F(2,208) = 2.24$, $p = .109$. However, as in the reduced sample, a significant difference in IAT effects emerged for the *pro-West* vs. *pro-East* condition ($t(105) = 1.98$, $p = .05$). The comparison of the *pro-West* vs. the *target-unrelated* condition tended to reach conventional levels of significance ($t(105) = 1.77$, $p = .08$). Replicating the effects of the reduced sample, the comparison of the *pro-East* vs. the *target-unrelated* condition remained non-significant ($t < 1$). All in all, the obtained pattern with the reduced sample does not depend on the exclusion of participants with very high error rates. Even if these participants enter the analysis and introduce additional error variance, the pattern of effects remains almost unchanged.

General discussion

Summary

To date, researchers have been relatively unaware of the influence that selected stimuli can exert on IAT effects. Whenever a stimulus representing a category is related to another category being measured, it contains a so-called cross-category association. Two experiments, applying one between-subjects manipulation and one within-subjects manipulation, fully corroborated our assumptions that these cross-category associations affect the resulting IAT effect. In Experiments 1 and 2, target stimuli of the concepts East and West Germans bore associations with the categories of the attribute dimension (positive and negative). Positive and negative attribute stimuli were related to the target categories East and West Germans. The manipulation of the direction of the cross-category associations led to the predicted pattern with huge differences in IAT effects in Experiment 1 and substantial differences in Experiment 2. Associations favoring the ingroup augmented IAT effects, whereas associations favoring the outgroup reduced IAT effects. Thus, we demonstrated that both target and attribute stimuli are capable of changing IAT effects within the same experimental design.

As a side note, we want to mention the results of East German participants who took part in the same studies. We do not report any detailed analyses for the East German samples because of small sample sizes ($N = 132$ and 35 for Experiments 1 and 2, respectively), and because the stimuli were drawn entirely from a West German pretest sample. Nevertheless, almost the same configuration of IAT effects resulted, but the reversed compatibility assignment from an East German perspective led to upside-down patterns (for what is a compatible block for a West German subject is incompatible for an East German subject, and vice versa). Thus, our expectations were by and large confirmed in two independent but complementary samples.

Alternative explanation

One could argue that, in Experiment 1, IATs bearing stimulus sets with uniform cross-category associations did not tap the superordinate concepts of East and West Germans, but that certain subtypes of these concepts became more accessible and were subject to measurement. Implicit attitudes associated with subtypes do not necessarily equal those associated with the superordinate categories the subtypes belong to. Other researchers have referred to related processes to explain the influence of stimuli on IAT effects. Govan and Williams's (2004) proposed a "category re-definition" process, whereas Mitchell and others (2003) referred to "contextual variations" of the attitude assessment. Common to

these related suggestions is the reasoning that the highly differing stimulus sets in our IAT versions activated different mental representations of East and West Germans. In some conditions the evaluated mental representation of East Germans was negative, while in others it was rather neutral or even positive (likewise for the mental representation of West Germans). The shared connection to a general concept of East Germans was rather unimportant, because this superordinate concept was not tapped exclusively.

We recognize subtyping and related processes as an alternative explanation for different target stimulus sets. But we do not see how this reasoning could be easily applied to the manipulated attribute stimulus sets. It requires that the rather subtle East or West German connotations of the strong evaluative concepts positive and negative would have to activate distinct subconcepts of these very basic and fundamental dimensions, that is, a $\text{positive}_{\text{West}}$, a $\text{negative}_{\text{West}}$, a $\text{positive}_{\text{East}}$, and a $\text{negative}_{\text{East}}$ subconcept. In our view, it seems rather unlikely that there could exist a specific "Western positive" or "Western negative" subtype (and an "Eastern positive" or "Eastern negative" subtype). These subconcepts would have to be very distinct from each other and from their superordinate concepts positive and negative in order to produce highly dissimilar IAT effects. Yet, the influence of cross-category associations of attribute stimuli is clearly evident in Experiment 1 (and also in Experiment 2). To us, it seems quite unlikely that the alternative reasoning could explain the complete predicted pattern of seven different IAT effects that we obtained in Experiment 1.

Although the alternative explanation underlines the power of the stimuli as well, we do not believe it is at the heart of our findings. It implicitly requires the stimuli of one category to share the same evaluative tone to form a homogeneous subconcept. Otherwise the stimuli will represent the superordinate concept in a differentiated way, thus making the activation of a specifiable subconcept unlikely. In Experiment 2, we employed an IAT whose categories were represented by stimuli with heterogeneous cross-category associations. In addition, these stimuli were applied in a completely randomized order. Nevertheless, we found clear-cut evidence for the impact of these associations on the magnitude of IAT effects. In our opinion, the results are in line with a rather parsimonious explanation of facilitation and inhibition in individual IAT trials when cross-category associations are entangled within single stimuli.

Note, however, that we do not deny the possible effects that subtyping might have on IAT effects. When consistently confounded stimulus sets were applied (Experiment 1), the degree of bias was much higher than in a balanced version (Experiment 2). Therefore, we assume that uniformity of cross-category associations in stimulus sets adds to the effects that individual items initially produce.

Theoretical models

To date, several researchers have published ambitious models, which try to explain the cognitive processes underlying an IAT (e.g., Brendl, Markman, & Messner, 2001; De Houwer, 2001, 2003; Greenwald et al., 1998; Mierke & Klauer, 2001, 2003; Rothermund & Wentura, 2001, 2004; Steffens et al., 2004). In only a few short years, the respective research and debate in the literature have greatly improved our understanding of what exactly leads to the emergence of an IAT effect. Unfortunately, until now no model is able to explain empirical results in IAT research in their entirety. What implications do our findings have for theoretical work concerning the IAT? Our data show that the stimuli exert influence on the magnitude of IAT effects that can even be drastic under certain circumstances. A model aiming to give a comprehensive account of IAT effects needs to take this source of influence explicitly into account. Future models of IAT effects should be able to distinguish between the influence of the categories, the target stimuli, the attribute stimuli, their interactions and other sources that lead to the emergence of reaction time differences in compatible and incompatible blocks.

The IAT intends to measure implicit attitudes. Since its introduction in 1998, there has been an ongoing debate about this claim, and inferring an implicit attitude from a given IAT effect may be more problematic than initially assumed. For example, to a certain extent the IAT seems to assess what kind of stimuli a researcher puts into it, not only the associations of concepts that a subject holds. It is reasonable to assume that not all seven IAT versions in Experiment 1, although intended to assess the same construct, are equally valuable in drawing conclusions about a particular individual. The present research as well as that of other researchers shows that not only drastic, but even minor manipulations of the material or the procedure of the IAT can lead to notably altered results (Govan & Williams, 2004; Mierke & Klauer, 2001, 2003; Mitchell et al., 2003; Steffens & Plewe, 2001; Rudman, Greenwald, & McGhee, 2001). Conversely, even critics have to acknowledge the growing body of research showing that some IATs predict (mostly spontaneous) behavior, sometimes even better than explicit measures of attitudes (Asendorpf et al., 2002; Egloff & Schmukle, 2002; Friese & Wänke, 2005; McConnell & Leibold, 2001). Whatever the theoretical construct that allows for this validity, a wide-ranging theoretical model has to identify its basis in the cognitive processes that occur during an IAT. Consequently, an IAT-model should be able to specify conditions under which a researcher can expect to maximize her chance of obtaining outcomes that are not essentially due to material or procedural variables.

Conclusions

Let us consider Fig. 1 again. Above we described how additional information in stimuli facilitates (or inhibits) the sorting process in an IAT. In two experiments we systematically manipulated the relations of target stimuli with the attribute dimension, and vice versa. It is obvious that this reasoning applies not only to the specific categories we chose, that is *East* and *West Germans* as well as *positive* and *negative*, but that it is based on a much broader principle. It extends to target categories like *Black*/*White*, *Germans*/*Turks*, *male*/*female,* as well as attribute categories like *powerful*/*weak*, *math*/*arts*, and so forth.

But the rationale is still more general and not limited to the mutual associations of target and attribute categories. We assume that any feature of a stimulus could add shared information that potentially affects the performance in an IAT (e.g., salience, self-reference, facial cues in picture-based IATs, perceptual similarity; De Houwer, Geldof, & De Bruycker, in press). Let us use the feature salience as an example. In their figure-ground model, Rothermund and Wentura (2001, 2004) identify salience of categories and stimuli as the determinative feature which causes IAT effects. They state that in a typical IAT, a negative target category and a negative attribute category and their respective stimuli will pop out as salient "figures" against the background (the two less salient categories with their respective stimuli). Response latencies will be shorter when the two figures are mapped onto a single response key than when they are mapped onto different response keys.

Applied to Fig. 1, size of objects again represents a certain degree of conceptual association. Shade of the objects corresponds to the salience information. In the lower panel, hatched objects are always large and small objects are always gray. A compatible mapping of the categories sharing the same salience information (shade) facilitates the sorting process and will lead to stronger IAT effects than a conceptual association on its own. Note that even if there were no conceptual association at all, salience differences between the target and attribute categories would produce an IAT effect.

Let us assume that there was a conceptual association of the categories Black with negative (and White with positive) for a certain subject. This association equals the feature size in Fig. 1 and is sufficient to produce an IAT effect. But since Black and negative will also constitute the two figures in such an IAT (shade), the sorting process will be much easier than it would be on the basis of the conceptual association alone, and the IAT effect will be larger. The same holds if Black target stimuli differ in valence from White target stimuli such that Black items are less favorable than White items. This cross-category association between the Black stimuli and the negative attribute category (and between the White stimuli and the positive attribute category) will facilitate the sorting process and give rise to a stronger IAT effect.

Concerning the generality of our findings, one could argue that the IATs in Experiment 1 rested on extreme selections of stimuli bearing cross-category associations that will not occur in regular IATs reported in the literature. We agree with this view as it pertains to the double-manipulated IAT versions. Concerning the versions with only target or attribute categories carrying cross-category associations, however, we refer to De Houwer (2001), who recognized that the target stimuli are often confounded with valence, stating that "in a typical IAT, there is a perfect confounding between the relevant and irrelevant feature of target concept stimuli" (p. 446). This point is even more important given that some theoretical accounts—like the task-set switching model (Mierke & Klauer, 2001, 2003) and the random walk model with fixed response threshold (Brendl et al., 2001)—implicitly rely on this confounding as a prerequisite for their explanations of IAT effects. Therefore, the upshot of our findings is the call for very careful pretesting of the stimulus material prior to conducting an IAT. Cross-category associations between target and attribute categories are irrelevant for the conceptual association to be measured, but they modify IAT effects nonetheless and need to be controlled for (see also Fiedler, Messner, & Bluemke, 2003).

Pretesting still leaves problems for the assessment of interindividual differences of implicit attitudes. The associations of the participants who pretested the stimulus material do not necessarily go along with the associations of a specific participant performing an IAT. Individually perceived associations might modify IAT results without the awareness of the researcher. If pretested associations happen to vary strongly between subjects, applying an individualized selection of stimuli for each subject may be recommended. Yet, more research is necessary to substantiate these considerations.

In spite of the remaining issues that must be solved, we believe that comprehensive pretests will strengthen the conclusiveness of every empirical result and are therefore needed in future research. We expect the IAT to prove its usefulness as a research device and a tool for predicting behavior the more the unintended influences on IAT effects are controlled for.

Appendix A

German stimuli used in Experiments 1 and 2 (bold) including their English translations.

Target stimuli	Positive		Neutral		Negative	
East	**Demokratischer Aufbruch**	**(early democratic movement)**	Begrüßungsgeld	welcome money	**Bespitzelung**	**surveillance**
	Karl Marx	(philosopher, historian)	Christa Wolf	(writer)	Erich Honecker	(former socialist leader)
	Montagsdemo	**(famous demonstration in 1989)**	Cottbus	(city)	**Kommunismus**	**communism**
	Neubeginn	restart	**Erfurt**	**(city)**	Nationale Volksarmee	(former East German army)
	Ostsee	**Baltic Sea**	Jena	(ciy)	**Plattenbau**	**(typical architecture)**
	Potsdam	(city)	**Ost-Mark**	**(former East German currency)**	SED	(former socialist party)
	Rügen	(island)	**Rostock**	**(city)**	Selbstschußanlagen	(fully automated weapons)
	Unter den Linden	(famous promenade)	Sachsen	(state)	**Stasi**	**(former Secret Service)**
	Weimar	(city)	Thüringen	(state)	Todesstreifen	(former death zone at border)
	Wende	**communist turn down in 1989**	**Trabant**	**(East German car)**	totalitäres System	totalitarian system
West	Alpen	Alps	**Kapitalismus**	**capitalism**	**Arroganz**	**arrogance**
	Demokratie	**democracy**	Boris Becker	(famous tennis player)	Bundeswehr	German armed forces
	Freiheit	**freedom**	**Perfektionismus**	**perfection**	**Egoismus**	**egoism**
	Hamburg	(city)	Ruhrgebiet	(region near the Ruhr river)	Ellenbogengesellschaft	selfish society
	Individualität	individualism	Wuppertal	(city)	**Geldgier**	**greed**
	Karriere	**career**	Bayern	(state)	Materialismus	materialism
	Nordsee	**North Sea**	Hessen	(state)	**Mobbing**	**mobbing**
	Unternehmergeist	entrepreneurism	**Bremen**	**(city)**	RAF	(fraction of terrorists)
	West-Berlin	(city)	Harald Juhnke	(actor)	Solingen	(city)
	Wohlstand	prosperity	**Daimler**	(car facturer)	Überheblichkeit	arrogance
Attribute stimuli	West		Neutral		East	
Positive	**dynamisch**	**dynamic**	begabt	talented	bescheiden	modest
	eigenständig	self-contained	ehrlich	honest	einfach	simple
	erfolgreich	**successful**	**freundlich**	**friendly**	**familiär**	**family-loving**
	flexibel	flexible	friedlich	peaceful	gastfreundlich	hospitable
	individuell	**individualistic**	hoffnungsvoll	hopeful	geduldig	patient
	multikulturell	multicultural	lebensfroh	full of the joys of life	**gemeinschaftlich**	**sociable**
	optimistisch	optimistic	leidenschaftlich	passionate	**genügsam**	frugal
	selbständig	self-dependent	**musikalisch**	**musical**	hilfsbereit	**helpful**
	selbstbewußt	**self-confident**	**nett**	**nice**	idealistisch	idealistic
	weltoffen	cosmopolitan	**tierlieb**	**animal-loving**	**natürlich**	**natural**
Negative	arrogant	arrogant	feindselig	hostile	abhängig	dependent
	geldgeil	greedy	**gefühllos**	**emotionless**	**arbeitslos**	**unemployed**
	hektisch	hectic	geizig	stingy	**ausländerfeindlich**	**xenophobic**
	hochmütig	haughty	**krank**	**ill**	**hoffnungslos**	**hopeless**
	karrieregeil	**strong need for career**	nörgelig	peevish	neidisch	envious
	machohaft	macho	taktlos	tactless	pessimistisch	pessimistic
	selbstgefällig	self-satisfied	unbeherrscht	uncontrolled	rechtsradikal	right wing extremist
	überheblich	**presumptuous**	**untreu**	**unfaithful**	trist	sad
	unkollegial	**uncooperative**	verantwortungslos	irresponsible	unproduktiv	non- productive
	unpersönlich	**impersonal**	**verlogen**	**lying**	**unzufrieden**	**dissatisfied**

References

Asendorpf, J. B., Banse, R., & Muecke, 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.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48*, 145–160.

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.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*, 163–170.

De Jong, P. D., Pasman, W., Kindt, M., & van den Hout, M. A. (2001). A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research and Therapy, 39*, 101–113.

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. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Lawrence Erlbaum.

De Houwer, J., Geldof, T., & De Bruycker, E., in press. The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology*.

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.

Fiedler, K., Messner, C., & Bluemke, M. (2003). *Unresolved problems with the "I," the "A"and the "T": Logical and psychometric critique of the Implicit Association Test*. Manuscript submitted for publication.

Fiedler, K., & Zogmaister, C. (2001). *Self-referent accentuation as a process underlying the IAT effect*. Unpublished research, University of Heidelberg, Germany.

Friese, M., & Wänke, M. (2005). *Implicit consumer preferences attitudes and their influence on product choice*. Manuscript submitted for publication.

Govan, C. L., & Williams, K. D. (2004). Reversing or eliminating IAT effects by changing the affective valence of the stimulus items. *Journal of Experimental Social Psychology, 40*, 357–365.

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. (2001). Health of the Implicit Association Test at Age 3. *Zeitschrift für Experimentelle Psychologie, 48*, 85–93.

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

Kuehnen, U., Schiessl, M., Bauer, N., Paulig, N., Poehlmann, C., & Schmidthals, K. (2001). How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. *Zeitschrift für Experimentelle Psychologie, 48*, 135–144.

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

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

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. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455–469.

Ottaway, S., Hayden, D. C., & Oakes, M. O. (2001). Implicit attitudes and racism: Effects of word familiarity and frequency on the implicit association test. *Social Cognition, 19*, 97–144.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*, 380–391.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research* (2nd ed.). New York: Methods and Data Analysis. McGraw-Hill.

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

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

Rudman, L. A., Greenwald, A. G., & McGhee, D. E. (2001). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin, 27*, 1164–1178.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437–465.

Steffens, M., Lichau, J., Still, Y., Jelenec, P., Anheuser, J., Goergens, N. K., et al. (2004). Individuum oder Gruppe, Exemplar oder Kategorie? Ein Zweifaktorenmodell zur Erklärung der Reaktionszeitunterschiede im Implicit Association Test (IAT) [Individual or group, exemplar or category? A two-factor model for the explanation of reaction time differences in the Implicit Association Test (IAT)]. *Zeitschrift für Psychologie, 212*, 57–65.

Steffens, M., & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 123–134.

Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behavior consistency for stigmatized behavior. *Cognition and Emotion, 15*, 207–230.

Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology, 100*, 226–235.

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