Beyond structural problems of the Implicit Association Test (IAT): Approaches to reduce contaminations of IAT effects

Inaugural-Dissertation

zur
Erlangung der Doktorwürde
der Wirtschafts- und Verhaltenswissenschaftlichen Fakultät der
Albert-Ludwigs-Universität Freiburg

vorgelegt von

Sarah Teige-Mocigemba
geboren am 01.03.1979 in Berlin

SS 2008

Dekan: Prof. Dr. Dieter K. Tscheulin

1. Gutachter: Prof. Dr. Karl Christoph Klauer

2. Gutachterin: PD Dr. Andrea Bender

Datum des Promotionsbeschlusses: 02.10.2008

Table of Contents

Danksagung

Ich möchte „Danke“ sagen, und zwar – als Ausnahme in dieser englischsprachigen Dissertation – in meiner Muttersprache.

An erster Stelle danken möchte ich meinem Doktorvater Karl Christoph Klauer, von dessen Erfahrungsschatz und methodischen wie inhaltlichen Wissen ich unglaublich viel lernen durfte und immer noch lerne. Ich danke Ihnen für Ihre unbedingte Unterstützung auf dem Weg zur Erlangung der Doktorwürde und für die Zeit, die Sie immer umgehend gefunden haben, um Fragen und Probleme zu diskutieren. Ich danke Ihnen für Ihr Vertrauen und die Freiräume, die Sie mir ließen, um meine eigenen Forschungsideen zu verfolgen. Nicht zuletzt gebührt Ihnen Dank für die erstklassigen Arbeitsbedingungen und das unterstützende Umfeld in Freiburg.

Danken möchte ich auch Andrea Bender, die sich als Zweitgutachterin dieser Arbeit zur Verfügung stellte.

Klaus Rothermund gilt mein Dank für viele fruchtbare Diskussionen, die unter anderem in zwei gemeinsame Publikationen mündeten, welche in diese Dissertation eingegangen sind.

Andreas Voß war während der Promotionszeit nicht nur ein geschätzter Kollege, sondern wurde mir vor allem auch zum Freund, Laufkumpel und Ratgeber in allen erdenklichen beruflichen wie privaten Lebenslagen. Danke, Andreas!

Ohne Christoph Stahl hätte ich meinen Weg von Berlin nach Freiburg möglicherweise nicht gefunden. Dank Deiner herzlichen und unterstützenden Art habe ich mich hier von Beginn an wohlgefühlt. Danke für Deine Freundschaft, Christoph! Andreas und Christoph sei außerdem gedankt, dass sie sich bereit erklärt haben, Teile dieser Arbeit Korrektur zu lesen.

Rainer Leonhart danke ich für seine stete Unterstützung in zahlreichen, nicht nur organisatorischen Belangen und für die Bekanntmachung mit der badischen Lebensart.

Für die freundschaftliche (Arbeits-) Atmosphäre und gegenseitige Unterstützung bei methodischen, theoretischen, technischen, strategischen und manchmal ganz einfachen praktischen Fragen danke ich auch meinen weiteren (Ex-) Kolleginnen und (Ex-) Kollegen Kerstin Dittrich, Julia Herfordt, Mandy Nuszbaum, Mandy Raumschüssel, Florian Schmitz, Katrin Schornstein, Zengmei Zhao und nicht zuletzt Adriaan Spruyt, dessen Anwesenheit in Freiburg zwar viel zu kurz, dafür aber sehr bereichernd war.

Einigen „externen“ Kolleginnen und Kollegen möchte ich danken dafür, dass sie mich motiviert haben in den letzten Jahren, indem sie den (wissenschaftlichen) Austausch zu einem schönen Erlebnis werden ließen. Dass die Erfahrung wissenschaftlichen Arbeitens für mich eine sehr positive und freudvolle ist, verdanke ich unter anderem Rainer Banse, Matthias Blümke, Jan De Houwer, Fritzi Dislich, Malte Friese, Nina Hansen, Wilhelm Hofmann, Jana Lüdtke, Konrad Schnabel und Michela Schröder-Abé.

Zu großem Dank verpflichtet bin ich auch den wissenschaftlichen Hilfskräften unserer Abteilung, die mit großem Engagement meine Versuche testeten, meine Versuchspersonen betreuten, mich mit Literatur versorgten und sich von Zeit zu Zeit nach meinem Wohlergehen erkundigten. Ich danke Christine Alle, Dalia Amar, Uwe Bach, Carsten Bartel, Christoph Breuniger, Anne Flad, Tessa Hannemann, Cornelia Hauss, Fabian Hölzenbein, Jan Hundertmark, Lisa Hüther, Damila Karacic, Matislava Karacic, Maria Mustafina, Birgit Netz, Christine Örtl, Marijo Plavac, Robert Ripfl, Miriam Seel und Katharina Wilm. Auch den Versuchspersonen sei gedankt, die sich für meine Studien zur Verfügung stellten.

Meinen Freundinnen und Freunden danke ich für viele schöne gemeinsame Erlebnisse, Gespräche und Reisen, durch die ich immer wieder den nötigen Abstand zur Arbeit bekam. Ich danke hier insbesondere Lydia und David Beck, Nils Behrendt, Tilo Bernhardt, Alice Dechêne, Ulrike Eichhorn, Anja Hochschild, Liane Jablonski, Jan-Gerrit Keil, Ulrike und Lars Krumrey, Britta Pieper, Gerald Riechmann, Claudio Roller, Rahel und Jörg Rosowski, Gunni und Vero Sachs, Henja Semmler, Alexandra Voß und Nina Weckesser.

Mein besonderer Dank gilt meinen Eltern Elisabeth und Wolf-Rüdiger Teige sowie meinen Brüdern Daniel und Robin Teige. Euch danke ich für Eure bedingungslose Unterstützung, Eure „Aufbauanrufe“ und das Gefühl, ein liebenswerter Mensch zu sein, egal ob mit oder ohne Doktortitel.

Aus tiefstem Herzen danke ich meinem Mann, Freund und Weggefährten Dennis Mocigemba, dessen Neugierde auf das Leben und dessen Begeisterung für neue Erfahrungen uns den Weg von Berlin nach Freiburg ebneten. Ich danke Dir, dass Du nicht müde wurdest, meinen Ausführungen über den IAT, den Single Block IAT, über ermutigende und frustrierende Erfahrungen im Entstehungsprozess der Dissertation zu lauschen. Deine ausdauernde Unterstützung und Deine bedingungslose Zuneigung haben einen unschätzbaren Anteil am Gelingen dieser Arbeit. Ich bin glücklich und dankbar, an Deiner Seite zu sein, wie vor Jahr und Tag.

Summary

This dissertation is concerned with an implicit measure that has received significant attention in the last decade: the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). Since its publication, the IAT has been applied to diverse areas of psychological research for the assessment of implicit attitudes, stereotypes, self-esteem, and personality traits. Numerous findings have shown that the IAT captures construct-related variance (see Nosek, Greenwald, & Banaji, 2006), but is also contaminated by several confounding factors that cause additional, construct-unrelated variance in the IAT effect (e.g., Fiedler, Messner, & Bluemke, 2006). Such contaminants can affect the size and the rank order of IAT effects. Therefore, they pose problems for both the interpretation of the absolute IAT effect and the interpretation of interindividual differences in the IAT effect. The focus of this dissertation lies in the exploration of techniques that prevent such contaminations as elaborated in the manuscripts presented in the Appendixes A to C.

The introduction of this dissertation provides background information about implicit measures in general and the IAT in particular. In this part, I briefly describe the research tradition that implicit measures emerged from. Then, I introduce the IAT methodology. In particular, research on the IAT's validity is reviewed, process models of the IAT are presented, and research on confounding factors of the IAT are discussed. I argue that the IAT effect reflects a conglomerate of different factors, some of which are construct-unrelated. Based on these considerations, I propose two approaches to reduce contaminations of IAT effects. Finally, possible extensions and limitations of these two approaches are discussed.

The empirical part of this dissertation investigates in how far the proposed approaches reduce contaminations of IAT effects. In one approach that is presented in the first manuscript (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007, Appendix A), it is suggested to extract the construct-related portion of variance from the IAT effect. Specifically, a diffusion-model analysis (Ratcliff, 1978) of the IAT is proposed. This analysis allows for the dissociation of distinct process components of the IAT effect and disentangles construct-related variance components and construct-unrelated variance components. Diffusion-model analysis therefore contributes not only to a deeper understanding of the processes underlying the IAT, but provides also a less contaminated measure of construct-related variance.

The other approach is presented in the second manuscript (Teige-Mocigemba, Klauer, & Rothermund, in press, Appendix B) and the third manuscript (Rothermund, Teige-

Mocigemba, Gast, & Wentura, in press, Appendix C). This approach is based on the assumption that many of the IAT's contaminations emanate from the IAT's block structure. It is therefore suggested to eliminate the IAT's block structure in order to prevent the confounds. Elimination of the block structure is realized in two paradigms, one called the Single Block IAT (SB-IAT; Appendix B) and the other called the IAT-recoding free (IAT-RF; Appendix C). Both procedures prove to be less contaminated by known confounding factors of the IAT, while the psychometric properties remain satisfactory. The findings therefore indicate that eliminating the IAT's block structure is sufficient to minimize the impact of several confounds on the IAT, thereby leading to less contaminated measures.

The two approaches proposed in this dissertation thus reduce confounding influences on IAT effects by different means. Future research will have to evaluate how far these approaches provide suitable measures of those "unconscious" parts of the self that self-reports cannot reveal.

Zusammenfassung

Diese Dissertation befasst sich mit einem impliziten Messverfahren, das in den vergangenen zehn Jahren große Aufmerksamkeit erregt hat: der Implizite Assoziationstest (IAT; Greenwald, McGhee & Schwartz, 1998). Seit seiner Publikation wurde der IAT in verschiedenen Bereichen der Psychologie zur Messung von impliziten Einstellungen, Stereotypen, Selbstwert und Persönlichkeitseigenschaften eingesetzt. Zahlreiche Studien zeigen, dass der IAT konstruktspezifische Varianz misst (siehe Nosek, Greenwald & Banaji, 2006), aber auch kontaminiert ist durch verschiedene konfundierende Faktoren, die zu zusätzlicher, nicht mit dem Konstrukt verwandter Varianz im IAT-Effekt führen (siehe Fiedler, Messner & Bluemke, 2006). Solche Konfundierungen können die Größe und die Rangreihe von IAT-Effekten beeinflussen. Daher erschweren sie sowohl die Interpretation von absoluten IAT-Effekten als auch die Interpretation von interindividuellen Unterschieden in IAT-Effekten. Die vorliegende Dissertation befasst sich in erster Linie mit der Erprobung von Techniken zur Vermeidung solcher Konfundierungen (siehe Anhänge A bis C).

Der einführende Rahmentext dieser Dissertation liefert Hintergrundinformationen zu impliziten Messverfahren im Allgemeinen und dem IAT im Besonderen. In diesem Teil beschreibe ich kurz die Forschungstradition, aus der implizite Messverfahren hervorgingen. Ich stelle dann die IAT-Methodik vor. Hierbei wird ein Literaturüberblick über Validitätsnachweise des IATs gegeben, bevor Prozessmodelle des IATs skizziert werden und schließlich jene Faktoren vorgestellt werden, die den IAT-Effekt kontaminieren. Es wird deutlich, dass der IAT-Effekt ein Konglomerat aus verschiedenen Variablen darstellt, wovon nicht alle konstruktvalide sind. Auf der Grundlage dieser Überlegungen schlage ich zwei Ansätze vor, um Konfundierungen im IAT-Effekt zu vermeiden. Abschließend diskutiere ich mögliche Erweiterungen und Grenzen beider Ansätze.

Der empirische Teil dieser Dissertation untersucht, inwieweit die vorgeschlagenen Ansätze die Konfundierungen im IAT-Effekt tatsächlich reduzieren. Ein Ansatz, der im ersten Manuskript vorgestellt wird (Klauer, Voss, Schmitz & Teige-Mocigemba, 2007, Anhang A), schlägt eine Diffusionsmodellanalyse (Ratcliff, 1978) des IAT vor, um konstrukt-spezifische Anteile aus dem IAT-Effekt zu extrahieren. Eine solche Analyse erlaubt die Unterscheidung verschiedener Prozesskomponenten des IAT-Effekts und trennt konstruktvalide Varianzkomponenten von konstruktinvaliden Varianzkomponenten. Diffusionsmodell-

analysen tragen damit nicht nur zu einem besseren Verständnis der Prozesse bei, die dem IAT zugrunde liegen, sondern liefern darüber hinaus weniger konfundierte Maße konstruktvalider Varianz.

Der andere Ansatz wird im zweiten (Teige-Mocigemba, Klauer & Rothermund, in press, Anhang B) und dritten (Rothermund, Teige-Mocigemba, Gast & Wentura, in press, Anhang C) Manuskript vorgestellt. Dieser Ansatz basiert auf der Annahme, dass viele der Konfundierungen im IAT-Effekt durch die Blockstruktur des Verfahrens zustande kommen. Um die Konfundierungen zu vermeiden, wird daher die Auflösung der Blockstruktur im IAT vorgeschlagen. Diese prozedurale Veränderung des IATs wird in zwei Paradigmen realisiert, dem sogenannten Single Block IAT (SB-IAT; Anhang B) und dem IAT-recoding free (IAT-RF; Anhang C). Beide Verfahren zeigen sich weniger anfällig gegenüber den bekannten Konfundierungen des IATs, wobei gleichzeitig die psychometrischen Eigenschaften zufriedenstellend bleiben. Die Befunde deuten demnach darauf hin, dass die Auflösung der Blockstruktur im IAT hinreichend ist, um den Einfluss konfundierender Faktoren zu minimieren, was zu konstruktvalideren Maßen führen sollte.

Durch beide Ansätze, die diese Dissertation vorschlägt, werden also auf unterschiedliche Arten Konfundierungen im IAT-Effekt reduziert. Zukünftige Forschung wird beurteilen müssen, inwieweit die durch diese Ansätze gewonnenen Maße geeignet sind, jene „unbewussten“ Teile des Selbst zu erfassen, die Selbstberichten verschlossen bleiben.

Introduction

> I cannot totally grasp all that I am … For that darkness is lamentable in which the possibilities in me are hidden from myself: So that my mind, questioning itself upon its own powers, feels that it cannot rightly trust its own report.
>
> (St. Augustine, trans. 1944)

In the early 5th century, St. Augustine described a phenomenon that is still subject to contemporary psychological research: the limits of self-understanding and, accordingly, the limitations of self-reports. So-called “explicit” measures such as questionnaires or interview methods suffer from two main problems, namely introspective limits (see Nisbett & Wilson, 1977) and susceptibility to self-presentation or socially desirable responding (Fazio, Jackson, Dunton, & Williams, 1995). In order to deal with these two key problems of explicit measures, research has concentrated on alternative methodologies, particularly so-called “implicit” response-time measures.

This dissertation focuses on such an implicit measure that has received significant attention in various areas of psychological research, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). This introduction will provide background information about implicit measures in general and the IAT in particular that were not conveyed in the published manuscripts (Appendixes A to C) constituting the main part of this work. First, the research tradition from which implicit measures emerged will be described briefly. Then, the IAT methodology will be introduced and the most important findings as well as process models of the IAT will be reviewed. It will be argued that due to its structure, the IAT captures not only desired, that is, construct-related systematic variance, but is also contaminated by undesired, that is, construct-unrelated (systematic) variance. The focus of this dissertation lies in the exploration of techniques that prevent such contaminations as elaborated in the manuscripts presented in the Appendixes A to C.

The first manuscript (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007, Appendix A) describes a diffusion-model analysis of the IAT that allows for the dissociation of distinct process components of the IAT (for diffusion-model analyses, see Ratcliff, 1978; Ratcliff & McKoon, 2008; Voss, Rothermund, & Voss, 2004). Importantly, construct-specific and method-specific variance clearly mapped on different components. Accordingly, diffusion-model analysis contributes not only to a deeper understanding of the processes underlying the

IAT, but also provides a less contaminated measure of construct-specific variance.

Both the second manuscript (Teige-Mocigemba, Klauer, & Rothermund, in press, Appendix B) and the third manuscript (Rothermund, Teige-Mocigemba, Gast, & Wentura, in press, Appendix C) explore a procedural change of the IAT. It is proposed that eliminating the IAT's block structure is sufficient to prevent those (recoding) processes that are argued to elicit confounding effects of construct-independent variables. Elimination of the block structure is realized in two paradigms. In the so-called Single Block IAT (SB-IAT), markers of method-specific variance are clearly reduced while the psychometric properties remain comparatively satisfying (Teige-Mocigemba et al., in press, Appendix B). In a similar vein, the so-called IAT-recoding free (IAT-RF) proves to be less confounded by other markers of recoding processes such as task-switch costs and biased selection of stimuli both of which have been found to contaminate the IAT effect (Rothermund et al., in press, Appendix C).

Within the scope of this introduction, the content of the three manuscripts will not be further reiterated. In particular, the last section of the introduction assumes that the reader is familiar with all three manuscripts. Final considerations in this last section address some possible extensions and limitations of the two proposed techniques for reducing contaminations of IAT effects.

New Approaches to an Old Problem

The story of the IAT is a story of an incredible boom. Ten years after its first publication, more than 200 papers report use of the method and hundreds of conference papers concerning the IAT have been held (Lane, Banaji, Nosek, & Greenwald, 2007). 38.300 hits of the term "Implicit Association Test" in a Google search (July 31, 2008) also point to its popularity, provided that such data can be regarded as meaningful. These figures pose the question: What has made the IAT so popular? In order to understand the explosion of research on the IAT, a closer look at past research traditions might be enlightening.

As the quotation in the beginning of this introduction makes clear, the desire to "grasp all that I am" (St. Augustine, trans. 1944) is an old phenomenon. For centuries, philosophy and psychology have been concerned with the idea that there might be more about ourselves than we *can* tell or *want* to tell. In the course of these considerations, researchers realized that information assessed by self-reports can only reveal a limited part of the self, namely the consciously accessible part that we are willing to communicate. Accordingly, the assessment of the other, hidden parts of the self that are thought to be "unconscious" and uncontrollable has always posed a fascinating challenge (e.g., see psychoanalysis, Freud, 1915).

Dual-Process Models

Contemporary psychology developed dual-process models to account for the assumption of different – consciously accessible and consciously inaccessible – parts of the self. Such dual-process models provide a theoretical framework integrating both controllable, rather conscious aspects of the self and uncontrollable, rather unconscious aspects of the self in terms of different information processing systems (see Chaiken & Trope, 1999; Smith & DeCoster, 2000). For example, the reflective-impulsive model developed by Strack and Deutsch (2004) distinguishes between two interacting systems of information processing, one reflective system that is based on propositional processes and one impulsive system that is based on associative processes. Propositional information processing corresponds to higher-order processes of reasoning and operates consciously, but slowly. Associative information processing corresponds to spread-of-activation processes and operates fast and effortlessly, but with limited conscious accessibility. These two kinds of reflective versus impulsive mental representations form what has been termed explicit versus implicit psychological constructs (for a thorough definition of the term "implicit" and its inflationary use, see De Houwer, 2006; De Houwer & Moors, 2007). Researchers have applied the distinction of explicit versus implicit to various constructs including attitudes (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Greenwald & Banaji, 1995), stereotypes (e.g., Devine, 1989), self-esteem (e.g., Greenwald & Farnham, 2000), and personality traits (e.g., Asendorpf, Banse, & Mücke, 2002).

The differentiation of two information processing systems at the theoretical level is accompanied by two different approaches of measurement at the empirical level (Strack & Deutsch, 2004). It is hypothesized that direct, explicit measures (such as self-reports using Likert scales, semantic differentials, or thermometer scales) assess intentionally given information about the self and are thus particularly suited to capture explicit representations of the reflective system. In contrast, indirect, implicit measures aim at assessing information about the self that is not intentionally given and are thus suited to capture implicit representations of the impulsive system.[1] Accordingly, direct, explicit measures as well as indirect, implicit measures are argued to be useful and meaningful tools to predict behavior (Strack & Deutsch, 2004; see also Asendorpf et al., 2002; Hofmann, Rauch, & Gawronski, 2007).

Direct, Explicit Measures

Past research thus attempted to improve both explicit measures and indirect, implicit measures. Regarding explicit measures, different strategies have been explored to counteract self-presentational distortions. For instance, the so-called bogus pipeline technique makes participants believe that a machine – a kind of lie detector – can identify untruthful responding to explicit measures (Jones & Sigall, 1971). This belief should discourage participants from self-presentational responding. In support of this assumption, the bogus pipeline technique has been shown to reduce socially desirable responding to explicit measures (Roese & Jamieson, 1993). Other, less effortful techniques focused on instructions that stress the importance of truthful responding or emphasize the anonymity of the respondents (Bradburn, Sudman, & Wansink, 2004; see also Antonek & Livneh, 1995, for the so-called randomized response technique). Self-presentation tendencies themselves also became the object of investigation (Banse & Gawronski, 2003; Dunton & Fazio, 1997; Plant & Devine, 1998). The basic idea was to assess interindividual differences in traits related to self-presentation in order to control for their undesired influences on the measurement's outcome. Research on explicit measures also tried to minimize other distorting influences such as cognitive and communicative processes in question comprehension and judgment formation (Schwarz, 1999, 2007b; Sudman, Bradburn, & Schwarz, 1996).

Indirect Measures

Even if self-presentational distortions are prevented, the problem remains that an explicit measure is restricted by the respondent's introspective ability. This insight has promoted the development of several indirect measures. Although these indirect measures differ in many aspects, they share one central feature: Other than direct (i.e., explicit) measures, indirect measures do not directly ask participants to provide information about attitudes or stereotypes. Instead, it is assumed that the to-be-measured construct (e.g., an attitude) causally produces the outcome of the indirect measure (De Houwer, 2006; De Houwer & Moors, 2007; see also De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2008). Accordingly, the outcome of an indirect measurement should reflect the underlying construct. A crucial element of the usefulness of indirect measures is therefore a deeper understanding of its underlying processes, that is, of how a to-be-measured construct translates into observed responses (cf. Wentura & Rothermund, 2007; see also Borsboom, Mellenbergh, & van Heerden, 2004).

The benefit of indirect measures is thus twofold: First, indirect measures conceal the measurement's mechanisms and possibly also the purpose of measurement. Therefore, they should be less prone to self-presentational distortion (e.g., Fazio et al., 1995; but see Teige-Mocigemba & Klauer, in press; Ziegler, Schmidt-Atzert, Bühner, & Krumm, 2007). Second, indirect measures should not be subject to introspective limitations because their access to the to-be-measured construct does not rely on introspection.

Early developments of indirect measures included so-called projective tests (e.g., Thematic Apperception Test; Murray, 1943) and so-called objective tests (e.g., the Objective Test Battery; Häcker, Schmidt, Schwenkmezger, & Utz, 1975). Although such measures have been widely used in applied contexts, findings of unacceptable reliability and validity seriously question their usefulness (e.g., Bohner & Wänke, 2002; Lilienfeld, Wood, & Garb, 2000). More recently, however, technological progress has paved the way for a new class of indirect measures to emerge, namely implicit response-time measures (see Wittenbrink & Schwarz, 2007). Considering the conception of the impulsive system as an associative system (Strack & Deutsch, 2004), it makes sense that in search of suitable indirect measures, recent research has concentrated on response-time paradigms: Implicit response-time measures are expected to offer straightforward access to cognitive structures or processes because response-time patterns should reflect spread-of-activation processes appropriately. Due to highly accurate computer-based methods for recording response times, it thus appeared to be possible to tackle an old problem by using new approaches.

Implicit Response-Time Measures

Social cognition research applied experimental paradigms of cognitive psychology such as sequential priming (Neely, 1977) or response interference tasks (Kornblum, Hasbroucq, & Osman, 1990) to the assessment of attitudes, stereotypes, and self-esteem (for reviews, see Fazio & Olson, 2003; Wittenbrink & Schwarz, 2007). The most prominent examples of such implicit measures include the affective priming task (Fazio et al., 1986), semantic priming (Wittenbrink, Judd, & Park, 1997), the go-/no-go association task (GNAT; Nosek & Banaji, 2001), the extrinsic affective Simon task (EAST; De Houwer, 2003a), and the Implicit Association Test (IAT; Greenwald et al., 1998). A detailed discussion of similarities and differences between these measures is provided by De Houwer (2001; 2003b; in press).

For the new implicit measures, the known problems of early indirect measures initially seemed to reoccur. Implicit response-time measures proved to be useful tools to examine differences at the group level, but scarcely any of the implicit measures met the test-theoretical criteria required for the assessment of differences at the individual level. As for the early indirect measures, most implicit response-time measures suffered from low to, at best, moderate reliability (Bosson, Swann, & Pennebaker, 2000; for affective priming, see Fazio & Olson, 2003; for semantic priming, see Kawakami & Dovidio, 2001; for the GNAT, see

Nosek & Banaji, 2001; for the EAST, see Teige, Schnabel, Banse, & Asendorpf, 2004; for an overview, see Nosek, Greenwald, & Banaji, 2006). In 1995, Greenwald and Banaji argued that because such unreliable measures fail to detect interindividual differences, their application to the assessment of implicit attitudes, stereotypes, or self-esteem at the individual level is highly problematic. Three years later, the IAT was introduced as the hitherto first implicit response-time measure that proved to be reliable, at least in terms of internal consistency with estimates ranging between .70 and .90 (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; but see Nosek et al., 2006, for lower test-retest reliability estimates of median r = .56).[2] Thus, the great demand for *reliable* implicit measures helped to get the IAT widely accepted. Last but not least, the IAT's easy applicability and effective promotion might also have contributed to its popularity and widespread use in diverse sub-disciplines of psychological research such as social psychology (e.g., Nosek et al., 2006), personality psychology (Asendorpf et al., 2002; Egloff & Schmukle, 2002; Schmukle, Back, & Egloff, in press; for an overview, see Schnabel, Asendorpf, & Greenwald, 2008), developmental psychology (e.g., Baron & Banaji, 2006; Dunham, Baron, & Banaji, 2007), neuroscience (e.g., Richeson et al., 2003), market research (Friese, Wänke, & Plessner, 2006; Maison, Greenwald, & Bruin, 2004), clinical psychology (e.g., Teachman, Gregg, & Woody, 2001; Teachman & Woody, 2003; Wiers, van Woerden, Smulders, & De Jong, 2002), health psychology (e.g., Schröder-Abé, Rudolph, & Schütz, 2007; Swanson, Rudman, & Greenwald, 2001), and forensic psychology (Gray, MacCulloch, Smith, Morris, & Snowden, 2003).

Implicit Association Test

Now that important factors have been identified which contributed to the IAT's popularity, how does the IAT work? The IAT is believed to assess the strengths of associations between target categories arranged on a bipolar dimension (such as Black persons vs. White persons) and attribute categories arranged on a bipolar dimension (such as positive vs. negative) by comparing the response latencies for two differently combined categorization tasks. Participants are instructed to categorize stimuli that represent the four categories (e.g., names typical for Blacks vs. Whites and positive vs. negative words). For categorization, participants use two response keys, each assigned to two of the four categories. The IAT's basic assumption is that if two concepts are highly associated, categorization will be easier when the two associated categories share the same response (in the so-called *compatible* block, De Houwer, 2003b) than when they require different responses (in the so-called *incompatible* block, De Houwer, 2003b). For example, individuals with implicit prejudices against Blacks should respond faster and more accurately when Black names and negative attributes are assigned to the same response key (and White names and positive attributes to the other key) compared to the reversed configuration (Black and positive are assigned to one key, White and negative to the other key). The difference in performance between these two kinds of mappings is called the IAT effect.

It is important to note that IAT effects thus rely on the comparison of performances in two different blocks of trials. Specifically, the compatible and incompatible block usually comprise about 72 trials each, in which the categories are consistently mapped onto the same response keys. In other words, response key assignments are the same within one block, but differ between the two blocks. Direction and size of IAT effects are often interpreted as reflecting the relative association strengths between the target and attribute categories.

Validity

While the IAT's reliability estimates are encouraging, does the IAT also prove to be valid, that is, does it capture meaningful variance of the to-be-measured construct? Research has examined the IAT's validity both at the group level and at the individual level.

Group Level

Universal attitudes. At the group level, a priori assumptions have been used to examine the IAT's validity. For example, normative studies and a priori arguments suggest that there are objects towards which most people have relatively uniform attitudes (e.g., most people prefer flowers over insects). Accordingly, such universal attitudes should be reflected in the IAT effect. Indeed, flower-insect IATs have repeatedly been found to show more positive attitudes towards flowers than towards insects (for the first demonstration of this effect, see Greenwald et al., 1998) indicating validity of the IAT.

Known-group approach. The so-called known-group approach contrasts groups that are assumed a priori to differ regarding the construct of interest. For some domains, the IAT proved to be valid as it revealed such differences (for an overview, see Greenwald & Nosek, 2001). For instance, Americans of Korean versus Japanese origin held mutually negative implicit attitudes towards each other as assessed by a racial attitude IAT (Greenwald et al., 1998). Also, White and Black individuals differed in their racial attitude IAT effects (Nosek, Banaji, & Greenwald, 2002). Last but not least, a homosexuality attitude IAT distinguished between homosexuals and heterosexuals (Banse, Seise, & Zerbes, 2001). In other domains, particularly those related to addictive behavior, the IAT failed to differentiate between groups (such as smokers vs. non-smokers, Swanson et al., 2001; but see Perugini, 2005).

Experimentally manipulated attitudes. Assuming that the IAT effect reflects the construct in question, experimental manipulation of this construct should influence the IAT effect in the expected manner. Olson and Fazio (2001), for instance, drew on this assumption and found evidence for the IAT's validity. Novel attitudes were formed by pairing previously unknown stimuli with other, clearly positive or negative stimuli. Results showed that IAT effects reflected these new attitudes, even when participants did not notice that the attitudes resulted from the stimulus pairings.

It is important to note that this validation approach has its limits when it comes to the assessment of constructs that are expected to be stable over time. This is because per definition, stable constructs (e.g., the personality trait of anxiousness) should not be affected by short-term manipulations. For instance, Schmukle and Egloff (2004) experimentally induced state anxiety by a public speaking task and found no effects on anxiety IATs. They interpreted their findings not in terms of the IAT's invalidity, but in terms of the IAT's validity as a measure of *trait* anxiety.

Against the background of such a state-versus-trait debate, it is also difficult to interpret the findings of experimental studies that showed the malleability of implicit attitudes and stereotypes as assessed by the IAT (for a review, see Blair, 2002). On the one hand, such studies argue for the IAT's validity because they demonstrate its sensitivity to experimental manipulations of the construct in question. On the other hand, however, such findings argue for the IAT's invalidity as they question the IAT's ability to capture temporally stable implicit constructs (e.g., personality traits or stable attitudes) and/or question the existence of such invariant, trait-like cognitive structures (see Schwarz, 2007a, for the latter position).[3]

Also, the conclusiveness of the approaches by universal attitudes and known-groups is limited: It cannot be ruled out that there are other, uncontrolled variables confounded with universal attitudes (e.g., stimulus selection) or group membership (e.g., cognitive abilities) that also account for the findings (see Banse et al., 2001, for controlling for one such possibly confounded variable).

Individual Level

Most research on the IAT's validity adopted the correlational approach. This validation approach, however, is also limited (for a recent discussion, see Borsboom et al., 2004). Correlations between IAT effects (e.g., an aggressiveness IAT effect) and criterion variables (e.g., aggressive behavior) might emerge because of a third variable (e.g., impaired cognitive skills) that influences both the IAT effect and the criterion variable (for evidence of a cognitive skill confound on the IAT, see below). For example, the finding that the IAT predicts aggressive behavior does not necessarily attest to the IAT's validity as a measure of implicit aggressiveness.

All in all, however, the evidence provided by correlational studies is to a large extent in line with the assumption that IAT effects can capture meaningful construct-related variance. Specifically, the IAT's validity has been investigated in terms of its correlations with (a) explicit measures, (b) other implicit measures, and finally, (c) by its predictive validity for behavioral measures.

Correlations with explicit measures. Most studies have concentrated on implicit-explicit consistency. Meta-analyses over various content domains (including attitudes, stereotypes, and self-concept) revealed low to moderate correlations between IATs and explicit measures of .24 (Hofmann, Gawronski, et al., 2005) and .37 (Nosek, 2005). Two reasons might account for the somewhat higher correlations yielded by Nosek's meta-analysis. First, Nosek's data refer to attitude domains for which higher implicit-explicit consistency is expected. Second, Nosek exclusively used relative thermometer scales as explicit measures which may better correspond to the IAT in that they more directly tap into an affective component.

There is still considerable controversy, whether such low to moderate correlations between the IAT and explicit measures should be interpreted as indices of discriminant validity or convergent validity (cf. Payne, Burkley, & Stokes, 2008; see also Nosek & Smyth, 2007). The core of this debate traces back to the question of differences and similarities of the cognitive structures that underlie implicit and explicit measures (see also Footnote 1). Some researchers postulate independent representations of implicit versus explicit constructs and thus interpret implicit-explicit correlations as indices of discriminant validity (e.g., Devine, 1989; Wilson, Lindsay, & Schooler, 2000). Other researchers postulate only one representation that can be tapped differently (i.e., using implicit or explicit measures) and consequently, interpret implicit-explicit correlations as indices of convergent validity (e.g., Fazio, 1990; Fazio & Olson, 2003; Fazio & Towles-Schwen, 1999; Nier, 2005).

It is beyond the scope of this dissertation to discuss this still unresolved issue in-depth (for a recent discussion, see the special issue of the journal Social Cognition denoted to the question "What is an attitude?", Gawronski, 2007). It is to be noted, however, that recent research advised caution in interpreting implicit-explicit correlations as evidence for underlying cognitive structures. For instance, Payne et al. (2008) argued that measures differ with regard to several (structural) features and showed that structural fit has a strong impact on implicit-explicit correlations. The more similar the task demands of implicit and explicit measures were, the higher both measures correlated, even when controlling for common method-specific variance. Implicit-explicit correlations might therefore rather reflect (structural) fit of the underlying measures than of the underlying cognitive structures (for further moderating factors of implicit-explicit consistency, see Hofmann, Gschwendner, Nosek, & Schmitt, 2005).

Correlations with other implicit measures. Considering the interpretation problems of implicit-explicit correlations, it has been suggested to concentrate on implicit-implicit correlations. Assuming that implicit measures capture the same (i.e., implicit) construct, the IAT's correlations with other implicit measures should reflect the IAT's convergent validity (e.g., Banaji, 2001). Interestingly, correlations between IATs and other implicit measures typically have been found to be weak (e.g., Bosson et al., 2000; Fazio & Olson, 2003; Marsh, Johnson, Scott-Sheldon, 2001; Olson & Fazio, 2003, Rudolph, Schröder-Abé, Schütz, Gregg, & Sedikides, in press; Sherman, Presson, Chassin, Rose, & Koch, 2003; Teige et al., 2004). Low implicit-implicit consistency, however, is often not attributed to the IAT's invalidity, but to two other factors: First, as discussed above, implicit measures other than the IAT often show unacceptable reliability estimates (but see the AMP, Payne et al., 2005). Because reliability sets upper limits on the to-be-expected correlation, implicit-implicit relations might necessarily be underestimated (Nosek et al., 2006; Teige et al., 2004; see Cunningham, Preacher, & Banaji, 2001, for an approach to correct for such measurement error through latent variable analysis).

Second, not only implicit-explicit consistency, but also implicit-implicit consistency might be influenced by the structural fit of the measures (Payne et al., 2008; see also De Houwer, in press). Empirical evidence for this assumption is provided by studies that approximated formerly dissimilar features of implicit measures and indeed found higher implicit-implicit correlations. For example, Olson and Fazio (2003) argued that – as a result of different task demands of the IAT and affective priming – the IAT reveals evaluations of *superordinate categories*, whereas affective priming reveals evaluations of *specific category exemplars* used as stimuli. When affective priming was made more similar to the IAT by encouraging the primes' categorization in terms of the superordinate category, the correlation between affective priming and IAT was increased. In a similar vein, Steffens, Kirschbaum, and Glados (in press) equated the IAT and a response-window priming task with regard to stimulus selection (i.e., both tasks used only the concept categories as stimuli). Again, IAT effects and priming effects correlated significantly. Last but not least, findings by Schnabel, Banse, and Asendorpf (2006b) provide indirect evidence for the assumption that implicit-implicit correlations are influenced by similarities and differences in measurement methods. Schnabel et al. employed the newly developed Implicit Association Procedure (IAP) that is methodologically very similar to the IAT and found unusually high correlations of .50 between IAT and IAP. Note however, that common method-specific variance might have enhanced the correlation (for the problem of method-specific variance, see Mierke & Klauer, 2003).

Taken together, the IAT's low correlations with other implicit measures do not necessarily indicate the IAT's invalidity, but can be accounted for by (a) low reliability estimates of implicit measures other than the IAT and (b) structural differences of the measures. As Rudolph et al. (in press) put it, "the devil may be in the procedural details not the underlying construct" (p. 15).

Predictive validity for behavioral measures. Most convincing in light of the above discussions are correlational studies that have demonstrated the IAT's ability to predict behavior over and above explicit measures. In this regard, Perugini (2005) distinguished between three different models of predictive validity of implicit and explicit measures: the additive, the multiplicative, and the double dissociation model. Research on the IAT found evidence for all three models. As proposed by the *additive model*, the IAT and explicit measures explained separate portions of relevant criterion variance (e.g., Schnabel, Banse, & Asendorpf, 2006a). As suggested by the *multiplicative model*, the IAT and explicit measures interacted in predicting relevant behavioral criteria (e.g., Schröder-Abé et al, 2007). Finally, as proposed by the *double dissociation model*, only the IAT predicted spontaneous behavior whereas only explicit measures predicted controlled behavior (e.g., Asendorpf et al., 2002; Egloff & Schmukle, 2002; McConnell & Liebold, 2001; cf. Friese, Hofmann, & Wänke, 2008; Hofmann et al., 2007).

Evidence for the predictive validity of IATs across various behavioral domains is also provided by a recent meta-analysis (Greenwald, Poehlman, Uhlmann, & Banaji, in press). In socially sensitive domains such as stereotypes and prejudice, the IAT showed better predictive validity than explicit measures. This might have been expected, given that particularly in these domains, socially desirable responding biases explicit measures. In contrast, the meta-analysis revealed lower predictive validity for IATs than for explicit measures in studies that explored brand preferences or political attitudes.

Importantly, in domains related to health behavior, the IAT has been shown to have weak predictive validity. For example, an IAT designed to assess preferences for apples vs. candy bars did not predict the subsequent choice between an apple and a candy bar (Karpinski & Hilton, 2001; Olson & Fazio, 2004; Spruyt, Hermans, De Houwer, Vandekerckhove, & Eelen, 2007). The IAT's insufficiency in such domains has been argued to be due to the IAT's sensitivity to so-called "extra-personal" knowledge (Olson & Fazio, 2004), that is, societal views that do not necessarily correspond to the personal view as is discussed below (see section on confounding factors of the IAT effect).

Summary and Discussion

In summary, the IAT has been proven to capture valid construct-related variance with regard to both the group level and the individual level. Although these findings are encouraging and indicate the IAT's validity, several studies have seriously challenged the assumption that IAT effects are driven primarily by the to-be-measured associations (for general criticism of the IAT, see Fiedler, Messner, & Bluemke, 2006; see also Arkes & Tetlock, 2004; Blanton, Jaccard, Gonzales, & Christie, 2006). First and foremost, researchers criticized that in contrast to the flourishing application of the IAT, the processes underlying the IAT are still unclear. The notion of the absence of a comprehensive and testable process model came along with a growing body of research that identified several confounds of the IAT effect. In the following, process models and possible confounds of the IAT will be reviewed. It will become apparent that much of the criticism of the IAT revolves around the IAT's block structure (i.e., the comparison of performance between the two IAT blocks each of which consistently maps categories onto response keys across many trials).

Criticism of the IAT

Fazio and Olson (2003) criticized that "despite incredible activity, research concerning implicit measures has been surprisingly atheoretical. It largely has been a methodological, empirically driven enterprise." (p. 301). Research on the underlying processes of the IAT indeed did not keep up with the explosion of studies that already applied the IAT to diverse psychological areas. This is problematic because particularly for indirect, implicit measures, the processes of how the to-be-measured construct (e.g., an attitude) translates into observed responses have to be clarified (Wentura & Rothermund, 2007). Identifying the underlying processes of the IAT is even more important, as several factors have been identified that contribute to the IAT effect independent of the to-be-measured construct and thus, cause additional, but construct-unrelated, variance in the IAT effect. Among these factors are task-switching costs (Mierke & Klauer, 2001, 2003), salience asymmetries (Rothermund & Wentura, 2001, 2004), unintended effects of the stimuli (Bluemke & Friese, 2006; Govan & Williams, 2004; Steffens & Plewe, 2001), strategic processing and faking (Fiedler & Bluemke, 2005; Fiedler et al., 2006; Steffens, 2004), and extra-personal associations (Olson & Fazio, 2004) as elaborated below. Thus, although the construct in question may be sufficient to cause an IAT effect, other factors might also lead to IAT effects independently of this construct.

Unfortunately, there is no comprehensive process model that takes all these factors into account and allows for disentangling their relative influences on the IAT effect. However, process models have been proposed that account for several, albeit not all factors that can cause systematic variance in IAT effects. A brief overview of the models proposed by Brendl, Markmann, and Messner (2001), De Houwer (2001, 2003b), Rothermund and Wentura (2001, 2004), and Mierke and Klauer (2001, 2003; Klauer & Mierke, 2005) is given in the next section.

Process Models of the IAT

Random-walk model. According to Brendl et al. (2001), the IAT effect reflects the result of a random-walk process in which evidence is accumulated on a joint response-related decision dimension. The time required before a response criterion is reached depends on whether all incoming information pushes an internal counter in the same direction. It is hypothesized that both information of the target categories (i.e., category membership such as Black vs. White) and information of the attribute categories (e.g., valence) drive the counter. Therefore, stimuli of the target categories (e.g., Black names vs. White names) should have a lower net accumulation rate in the incompatible than in the compatible IAT condition, as information on the category membership (i.e., Black vs. White) and valence of a stimulus (i.e., negative vs. positive) disagree in the former, but not in the latter, condition.

To illustrate the process, let us draw on the racial attitude IAT as introduced above. For individuals with implicit prejudices against Blacks, Black stimuli (e.g., typical names) do not only belong to the category Black, but are also negatively evaluated. If a Black stimulus has to be categorized in the compatible block (here: Black/negative vs. White/positive), both sources of information, that is, the membership of the category Black as well as the negative valence, push the accumulation process toward the same response (i.e., the common response for Black names and negative words). In contrast, in the incompatible block (here: White/negative vs. Black/positive), the two sources of information move the accumulation process in opposite directions, because now, Black names and negative words are to be mapped onto different responses. Usually, category membership will have the stronger impact resulting in correct responses in most trials. However, all in all, the net evidence accumulation rate for Black stimuli should be lower in the incompatible block than in the compatible block, thus leading to slower responses in the former than in the latter task.

Brendl et al. (2001) predict that differences in net accumulation rate are accompanied by a shift in the response criteria in the incompatible block of an IAT. The authors assume that because the incompatible block is perceived as more difficult, participants adopt a more conservative response criterion leading to slower responses in the incompatible block compared to the compatible block. Accordingly, Brendl et al. suggest two mechanisms by which IAT effects are produced, namely different rates of information accumulation and different response criteria (see also Klauer et al., 2007, Appendix A).

Stimulus-response compatibilities. De Houwer (2001, 2003b) proposes that the IAT effect is based on stimulus-response compatibility. The basic assumption in this model is that response keys acquire the meaning of the stimulus category they are assigned to. Compatibility between the meaning of a response key and stimulus features then facilitates responses. This mechanism can explain the IAT effect, as compatibility between stimulus and response is consistently given in the compatible block of the IAT, but not in the incompatible block.

Again, let us take a look at the racial IAT. By asking participants to press one key for negative words and another key for positive words, the a priori neutral keys become associated with negative and positive valence, respectively (cf. Eder & Rothermund, 2008). Hence, for prejudiced individuals who like White persons but dislike Black persons, stimuli and responses are compatible (i.e., associated with the same valence) when the “negative” key has to be pressed for Black names and the “positive” key has to be pressed for White names (as is the case in the Black/negative – White/positive block). When the same individuals are asked to press the “negative” key for White names and the “positive” key for Black names (as is the case in the White/negative – Black/positive block), stimuli and responses are incompatible. Because stimulus-response compatibility varies between the different blocks of an IAT, De Houwer (2001, 2003b) hypothesized that IAT effects are due to the activation of responses by (relevant or irrelevant features of) the presented stimuli (see also De Houwer, in press).

Figure-ground asymmetry. According to Rothermund and Wentura (2001, 2004), the IAT measures differences in the salience of stimulus categories. Figure-ground asymmetries within the target (e.g., Black vs. White) and attribute (e.g., negative vs. positive) dimensions are the central explanatory concept of this account. The authors assume that participants simplify the compatible block in which the salient categories are mapped onto one response key by recoding both categorization tasks as figure-ground discriminations. This way, all salient stimuli (i.e., figure) are assigned to one key and all non-salient stimuli (i.e., ground) to the other so that the salient stimuli constitute the figure against the background of the less salient stimuli. Such a figure-ground recoding might be based on strategic processes or on non-strategic, spontaneous processes (Rothermund & Wentura, 2004; Rothermund, Wentura, & De Houwer, 2005). Importantly, recoding is impossible in the incompatible block in which the salient categories are mapped onto different response keys. Accordingly, performance differences between both blocks are argued to be the result of salience asymmetries.

Applying the figure-ground asymmetry account to the racial IAT, it can be argued that Black names and negative words are more salient than White names and positive words. Black names are salient because they are unfamiliar, whereas negative words are salient because of the attention-grabbing power of negative information (Pratto & John, 1991). Accordingly, participants should respond faster and more accurately, if the salient categories Black and negative share one response key (as is the case in the Black/negative – White/positive block) than if the salient categories Black and negative are mapped onto different response keys (as is the case in the White/negative – Black/positive block). This is because in the former, but not in the latter case, participants can reduce the complex 4-to-2 categorization task to a single binary decision of whether the stimulus belongs to the salient category (i.e., figure) or to the non-salient category (i.e., ground). Unlike other process models, the figure-ground asymmetry account thereby assumes that associations between categories play a rather subordinate role for the IAT effect compared to salience asymmetries (but see Kinoshita & Peek-O'Leary, 2006). A wide range of IAT findings can be explained by assuming that asymmetries in salience are paralleled by asymmetries in valence or familiarity, even though, in principle, salience is dissociable from these latter constructs (Rothermund & Wentura, 2004; Rothermund et al., 2005; but see Greenwald, Nosek, Banaji, & Klauer, 2005).

Task-switching. According to Mierke and Klauer (2001, 2003; Klauer & Mierke, 2005), task-switching costs contribute to the IAT effect because they affect the two crucial blocks of the IAT asymmetrically. Thus, the central assumption of the task-switching account is that the IAT involves executive control processes (i.e., identifying and switching to the appropriate task set). Specifically, it is argued that in the compatible block of an IAT, the structure of the task provides participants with an overlapping attribute. Again, think of prejudiced individuals who like White persons but dislike Black persons. For these individuals, negative words and Black names share the attribute negativity, whereas positivity is shared by positive words and White names. In the Black/negative – White/positive block (here, the compatible block) of the racial IAT, categories that share an attribute, namely valence, are thus mapped onto one response key. Categorizing a Black or White stimulus according to valence (negative or positive) or according to category membership (Black or White) should thus lead to the same response (Mierke & Klauer, 2001, 2003). Consequently, the task-switching account assumes that participants derive their responses from the attribute shared by the target category in the compatible block. Because the process of deriving responses is thereby simplified, responses should be faster in this condition.

In contrast, responses cannot be derived from an overlapping attribute in the incompatible IAT condition. For instance, if the same prejudiced individuals complete the White/negative – Black/positive block of an IAT, responding to a Black name on the basis of its valence (here: negative) would lead to an incorrect response. In the incompatible block, attribute-related information thus needs to be ignored for stimuli of the target categories (Black vs. White), but has to be processed for stimuli of the attribute categories (negative vs. positive). Hence, in the incompatible block participants are required to perform each and every task-switch, whereas the compatible block can be completed without such task-switches. Task-switching is associated with performance costs (e.g., Rogers & Monsell, 1995). Because costly task-switches affect both blocks asymmetrically, task-switching ability should contribute to the IAT effect.

Summary. Despite the absence of a comprehensive, testable process model, fruitful proposals have been put forward about the processes by which variables may cause variations in IAT effects. Research has confirmed some, albeit not all predictions of the respective process models. For instance, a shift in response criterion has been found to contribute to the IAT effect as proposed by Brendl et al.'s random-walk model (Klauer et al., 2007, Appendix A). Salience has been shown to influence and contaminate IAT effects as proposed by Rothermund and Wentura's figure-ground asymmetry account (Rothermund & Wentura, 2001, 2004). Maybe the strongest support has been obtained for Mierke and Klauer's task-switching account, as task-switching has repeatedly been found to substantially contribute to IAT effects (e.g., Back, Schmukle, & Egloff, 2005; Klauer & Mierke, 2005; McFarland & Crouch, 2002; Mierke & Klauer, 2001, 2003).

The conclusions drawn in this dissertation have been strongly stimulated by the task-switching account, and one merit of the present work is the identification of different process components involved in the IAT (see Klauer et al., 2007, Appendix A) that can be explained by the task-switching account, but are not entirely consistent with other accounts. An exhaustive evaluation of strengths and weaknesses of all process models, however, is beyond the scope of this introduction. Instead, the following section concentrates on a shared assumption underlying all process models, namely that IAT effects are not only influenced by the to-be-measured associations between categories, but also by other, construct-unrelated factors.

Confounding Factors of the IAT Effect

Cognitive abilities. Several studies have shown the confounding influence of cognitive abilities on the IAT effect. For instance, McFarland and Crouch (2002) observed a correlation between overall response speed and the size of IAT effects on a variety of IAT tasks. Because overall response speed is associated with cognitive abilities, the results of McFarland and Crouch suggest that IAT effects are at least partially determined by the participants' cognitive abilities. Further evidence comes from studies showing larger IAT effects for older individuals compared to younger individuals (e.g., Hummert, Garstka, O'Brien, Greenwald, & Mellott, 2002). Given that cognitive abilities tend to decline with age, such findings also suggest that IAT effects are influenced by general cognitive abilities (see also Sherman et al., 2008). This assumption is also corroborated by analyses using the quad model that was recently proposed by Conrey, Sherman, Gawronski, Hugenberg, and Groom (2005). Analyzing accuracy data of the IAT via multinomial models, the quad model disentangles four components that drive the IAT effect, including one for inhibition that can be interpreted as reflecting the influence of cognitive abilities on the IAT effect.

In support of their task-switching account, Mierke and Klauer (2001, 2003; Klauer & Mierke, 2005) provided further evidence for the confounding influence of cognitive abilities on the IAT effect. Several studies (Back et al., 2005; McFarland & Crouch, 2002; Mierke & Klauer, 2003) found correlations between IATs that were supposed to capture different, unrelated constructs and therefore, should not be inter-correlated. The finding of substantial correlations indicates that a general factor such as cognitive ability influences all IAT effects, regardless of the to-be-measured constructs (but see Klauer et al., 2007, Appendix A, for an account by speed-accuracy trade-offs). This assumption is consistent with the task-switching account: Participants who show good performance in switching between tasks should be less affected by whether the response assignments require task-switches or not. Hence, regardless of the construct that an IAT is supposed to measure, participants with higher task-switching abilities are expected to reveal smaller IAT effects than participants with comparatively lower task-switching abilities. Thus, because the two crucial blocks of an IAT make different demands on the participants' cognitive abilities, the IAT effect should reflect those cognitive abilities to some extent. Accordingly, the IAT's block structure fosters the contamination by cognitive abilities.

Unlike for other confounding factors, for which empirical evidence has been more mixed, most researchers acknowledge the confounding influence of cognitive abilities. Accordingly, several, albeit often not fully satisfactory, attempts to control for such influences have been made (e.g., Greenwald, Nosek, & Banaji, 2003). The three manuscripts of this dissertation also address the IAT's contamination by cognitive abilities (see Appendixes A to C).

Salience. In support of their figure-ground asymmetry account, Rothermund and Wentura (2001, 2004) reported experimental data (i.e., manipulations of salience influence IAT effects) and correlational data (i.e., IAT effects are related to measures of salience). These findings corroborate the assumption that salience asymmetries have the potential to contribute to IAT effects as acknowledged by the developers of the IAT (see Greenwald et al., 2005). It is still controversial, however, how pervasive the impact of salience asymmetries is (see Rothermund et al., 2005). Recent studies indicate that only part of the IAT effect can be accounted for by construct-unrelated salience asymmetries, as construct-related compatibilities between the nominal categories have been shown to simultaneously contribute to IAT effects (e.g., Kinoshita & Peek-O'Leary, 2006; see also Houben & Wiers, 2006). Moreover, there is still uncertainty at the conceptual level about how salience should be measured (e.g., Greenwald et al., 2005) and how it is related to other attributes such as familiarity and polarity (e.g., Kinoshita & Peek-O'Leary, 2005, 2006; Proctor & Cho, 2006). However salience should be conceptualized, it appears that the two crucial blocks of the IAT are asymmetrical with regard to certain features (such as salience). These features therefore influence the IAT effect. Accordingly, the IAT's block structure promotes a confounding by features such as salience.

Similarity. The conclusion that the two crucial blocks of the IAT are asymmetrical with regard to certain features prompted De Houwer, Geldof, and De Bruycker (2005) to argue that the IAT might be a general measure of similarity. Their main assumption is that the compatible block of an IAT maps similar categories onto one response key, whereas the incompatible block maps dissimilar categories onto one response key leading to an asymmetry of both blocks with regard to similarity. Because the type of similarity is not specified, all kinds of factors related to similarity can in principle cause variations in IAT effects such as similarity with regard to valence, meaning, salience, or perceptual form (see De Houwer et al., 2005). The type of similarity that is most salient in a given situation should determine which of all possible kinds of similarity finally drive the IAT effect (see Medin, Goldstone, & Gentner, 1993). Although a lot of findings are consistent with the similarity account, it clearly runs the risk of being not falsifiable, given that similarity is a rather unconstrained concept (i.e., everything is similar to everything else in some respect).

Stimuli. The IAT effect has been found to be determined both by the superordinate nominal categories according to which the stimuli have to be categorized (i.e., the *category labels* such as “Black” vs. “White”; De Houwer, 2001, in press; for a similar argument, see Olson & Fazio, 2003) and by the *stimuli* used to represent the categories (e.g., a particular Black or White face; Bluemke & Friese, 2006; Govan & Williams, 2004; Mitchell, Nosek, & Banaji, 2003; Steffens & Plewe, 2001). Influences at the level of the categories (or, category labels) are desired. They assure the experimenter’s control over the nominal categories according to which participants categorize and process the stimuli (see the *relevant* feature account by De Houwer, in press). This allows for determining the construct that the IAT effect should reflect, and also adds to the IAT’s easy applicability to various domains.

Influences at the level of the stimuli, however, are often unintended. For example, several studies have indicated that stimulus selection may force participants to categorize stimuli according to other than the specified category labels (see the *irrelevant* feature account by De Houwer, in press). As Govan and Williams (2004) proposed, participants may re-define the category labels in order to reconcile meaning and/or valence of category labels with meaning and/or valence of stimuli. Biased selections of (target) stimuli can thus have dramatic influences on the magnitude and even on the direction of IAT effects (Bluemke & Friese, 2006; Govan & Williams, 2004; Steffens & Plewe, 2001) which poses a threat to the IAT’s validity. Careful stimulus selection is therefore required in order to exert as much control as possible over the nominal categories according to which participants categorize stimuli. Note that stimulus influences might rely on a *strategic* process of re-defining the category labels as proposed by the strategic recoding account of Rothermund and Wentura (2004) which is presented next.

Strategic effects. Rothermund and Wentura (2004; Rothermund et al., 2005) suggested that participants might strategically recode the double discrimination task in the compatible block of an IAT. They assumed that the consistent mapping of categories onto response keys as provided by the IAT’s block structure allows for a strategic recoding in service of simplifying the task. Recoding the four categories of an IAT into two might rely on any feature that helps to distinguish between the two groups of stimuli that are assigned to different response keys (see also Mierke & Klauer, 2003). Participants might even draw on societal views (so-called “extra-personal” knowledge; Olson & Fazio, 2004) or stereotypes (Devine, 1989) to simplify the tasks (Rothermund & Wentura, 2004) as is discussed below.

Accordingly, the IAT effect might in part reflect those features that participants decided to use for categorization.

Evidence for strategic effects on the IAT came from studies that investigated the fakeability of the IAT. These studies revealed that the IAT outcome can indeed be strategically controlled (a) if participants are told how to fake (Fiedler & Bluemke, 2005; Kim, 2003), (b) if participants are high on self-monitoring and highly motivated to fake (Czellar, 2006), or (c) if participants had experience with at least one prior IAT (Fiedler & Bluemke, 2005; Steffens, 2004). If, however, participants were exposed to an IAT for the very first time (Banse et al., 2001; but see also De Houwer, Beckers, & Moors, 2007) or if they were not advised on how to fake (Asendorpf et al., 2002; Egloff & Schmukle, 2002) there was little evidence for strategic control over the IAT outcome. Accordingly, under certain circumstances, participants might strategically influence the IAT effect. However, it is doubtlessly much easier to exert strategic control over self-reports than over an IAT (Steffens, 2004).

Strategic effects, if they occur, pose all the more a threat to the IAT's validity as individuals might differ in the extent to which and in the success with which they use strategies when completing an IAT. Schnabel et al. (2006b), for instance, showed that only some participants were able to generate strategies that successfully altered their IAT effect. Specifically, only those participants who took the perspective of a non-shy person while working through an IAT revealed lower shyness IAT effects than a control group.

Extra-personal associations. Olson and Fazio (2004) identified another confounding influence on the IAT effect, namely so-called "extra-personal" associations (see also Arkes & Tetlock, 2004; Karpinski & Hilton, 2001). The term extra-personal knowledge refers to culturally shared assumptions (e.g., apples are healthy and thus, are positive) that do not necessarily correspond to personal evaluations (e.g., I don't like apples). To the extent to which behavior is driven rather by personal views than by societal views (or a mixture of both), the IAT's sensitivity to extra-personal associations poses a threat to its validity (see Karpinski & Hilton, 2001; Olson & Fazio, 2004; but see Nosek & Hansen, 2008).

Evidence for the IAT's contamination by extra-personal knowledge is provided by experiments in which the manipulation of extra-personal views led to changes in IAT effects (Han, Olson, & Fazio, 2006). Furthermore, when groups with diverging personal and extra-personal (i.e., societal) views were tested, IAT effects at least sometimes seemed to be in line with societal views. In a racial IAT, for instance, even Black persons showed prejudices against Blacks (Olson & Fazio, 2004). Similarly, independent of their subsequent choice between an apple and a candy bar, participants revealed IAT effects that indicated negative attitudes towards the unhealthy but tasty candy (Spruyt et al., 2007). Also, when the IAT procedure is changed in such a way that it should be less susceptible to the impact of societal views (i.e., by removing error feedback and by using more personalized category labels), evidence for the causal role of societal views becomes weaker (Olson & Fazio, 2004; but see Nosek & Hansen, in press, for criticism of such personalized IAT variants). Last but not least, the assumption that extra-personal associations confound IAT effects is consistent with studies showing the IAT's weakness to predict behavior in health-related domains (e.g., Karpinski & Hilton, 2001; Spruyt et al., 2007), where societal views are prevalent.

Note however, that doubts have been raised about the theoretical significance and validity of the extra-personal account of IAT effects. At the empirical level, recent correlational studies provided little evidence for a link between IAT effects and measures of societal views (Nosek & Hansen, 2008). At the conceptual level, it has been argued that the distinction between personal and extra-personal views actually makes little sense, especially when considering the automatic effects of personal and extra-personal associations (Banaji, 2001; Gawronski & Bodenhausen, 2006; Nosek & Hansen, 2008). Even more important, there is still uncertainty about how extra-personal associations can be conceptualized (for a recent discussion, see Gawronksi, Peters, & LeBel, 2008). Considering the conceptual vagueness, it is not surprising that the processes via which such extra-personal associations contaminate IAT effects are not yet identified. According to Rothermund and Wentura (2004), strategic processes may be at work: Participants might strategically use extra-personal knowledge in order to simplify the complex categorization task, thereby improving their performance.

Compatibility order. Since its introduction in 1998, the IAT is known to be confounded by compatibility order: IAT effects tend to be larger if the compatible block precedes the incompatible block than vice versa (see Nosek et al., 2006). A theoretical account for compatibility-order effects was provided by Klauer and Mierke (2005). Drawing on their task-switching account, Klauer and Mierke suggested that differences in the accessibility of attribute information in the compatible versus incompatible block of the IAT may account for compatibility-order effects.

Compatibility-order effects are difficult to control for, given that compatibility is a function of interindividual differences in the construct of interest and cannot a priori be determined in many applied contexts. For example, for prejudiced individuals who like White persons but dislike Black persons, the compatible block of a racial IAT is the block that maps Black and negative onto one key and White and positive onto the other key. For individuals, however, who like Black persons but dislike White persons the same Black/negative – White/positive block is incompatible because it maps non-associated categories onto one response key. Accordingly, compatibility-order effects constitute an undesirable confounding in the IAT and might influence magnitude and rank order of IAT effects. First attempts to reduce the confounding impact of compatibility order have focused on slight changes of the IAT procedure (see Nosek, Greenwald, & Banaji, 2005).

Relative measure. Last but not least, it has been criticized that the IAT is restricted to the assessment of *relative* association strengths between nominal categories. For instance, a positive score in a racial IAT might indicate that Whites are preferred to Blacks. As a first problem, this limits the applicability of the IAT to constructs that have a natural counterpart. Several constructs of interest, however, do not meet this requirement. For example, if researchers are interested in the participant's fear of spiders, it is difficult to think of a suitable counterpart that could serve as a contrast category. Research has therefore suggested different, albeit not fully satisfying solutions such as contrasting the target category with a neutral category (e.g., Sherman et al., 2003) or employing other implicit measures that allow for the assessment of associations between a single construct category and attribute categories (for the Single Category IAT, see Karpinski & Steinman, 2006; for the GNAT, see Nosek & Banaji, 2001; for the EAST, see De Houwer, 2003a; for the Single Association Test, see Blanton et al., 2006).

A second problem of the IAT as a relative measure has been highlighted by Blanton et al. (2006). The authors argue that the difference model underlying IAT effects makes certain measurement assumptions that have to be questioned based on empirical findings. Blanton et al. further criticize that a given IAT effect is restricted in the conclusions that can be drawn. For instance, a positive IAT effect in a racial IAT – allegedly indicating prejudice – does not necessarily reflect negative attitudes towards Blacks and positive attitudes towards Whites. This effect might as well indicate that a participant holds positive attitudes towards Blacks, but is even more positive towards Whites. In contrast, it might also mean that the participant dislikes both, but evaluates Blacks even more negatively than Whites. Alternatively, the effect might indicate that the participant is neutral towards Blacks, but positive towards Whites, or it could reflect that the participant is neutral towards Whites, but negative towards Blacks. As becomes evident, an IAT effect does neither allow for any conclusions about the participant's evaluation of the single categories, nor does the same IAT effect of different participants necessarily reflect the same attitude.

Summary and Focus of this Dissertation

Despite empirical evidence showing that the IAT captures construct-related variance, numerous factors have been identified that also contribute to IAT effects and cause additional, construct-unrelated variance in the IAT. Most of these confounds are argued to emanate from the IAT's block structure because the consistent mapping of categories onto response keys in the compatible versus incompatible block promotes the confounding influences. Such contaminations pose serious problems: First, size and direction of the IAT effect are not unequivocally interpretable as indices of the relative association strengths between nominal categories (see Greenwald et al., 2005). Although some portions of the IAT effect's variance might be determined by the to-be-measured construct-variance, an unknown amount of variance may be caused by construct-unrelated factors (see De Houwer et al., 2005; Fiedler et al., 2006; Rothermund et al., 2005). Given that it is hitherto not possible to estimate the relative influences of all factors that contribute to the IAT effect, the "valid" (i.e., construct-related) amount of variance in the IAT effect cannot be determined.

Second, some confounding factors such as cognitive abilities or strategic influences differ between individuals. Such interindividually differing factors distort not only the size of IAT effects, but also the rank order which restricts the IAT's predictive power. Confounding factors thus pose a problem for both, the interpretation of absolute IAT effects and the interpretation of the IAT as a measure of interindividual differences.

In view of these considerations, the main question of this dissertation becomes evident: Given that the IAT effect reflects a conglomerate of different factors, some of which are construct-unrelated, how can we obtain less contaminated IAT effects? Two approaches are proposed.

The first approach attempts to tackle the symptoms of the confounding influences in that it leaves the IAT procedure unchanged but suggests ways to extract the construct-related portions of variance in standard IATs. Specifically, the first manuscript (Klauer et al., 2007, Appendix A) proposes a diffusion-model analysis (Ratcliff, 1978) of the IAT that allows for disentangling construct-related variance components and construct-unrelated variance components (for another approach of disentangling different components of the IAT, see the quad model developed by Conrey et al., 2005).

The second approach attempts to tackle the cause of the confounding influences in that it changes a crucial structural feature of the IAT procedure that accounts for many of the identified confounds. Specifically, the second (Teige-Mocigemba et al., in press, Appendix B) and third manuscript (Rothermund et al., in press, Appendix C) propose to get rid of the IAT's block structure. It is argued that this structural change of the standard IAT is sufficient to reduce the impact of several, albeit not necessarily all confounding factors on the IAT effect, leading to less contaminated IAT scores.

The proposed approaches to reduce contaminations of IAT effects are elaborated in the Appendixes A to C. For the last section on possible extensions and limitations of the two approaches, it is assumed that the reader is familiar with the three manuscripts.

Extensions and Limitations of the Proposed Approaches

Decomposing the IAT Effect by Diffusion-Model Analyses

In the recent past, it has become widely accepted that many implicit measures cannot be regarded as process-pure, but rather reflect a conglomerate of different processes (e.g., see Conrey et al., 2005; Klauer, Musch, & Eder, 2005; Klauer, Teige-Mocigemba, & Spruyt, in press; Lane et al., 2007). Accordingly, methods are needed that allow for disentangling such process components. In this context, diffusion-model analyses appear to be particularly suitable because they capitalize on both performance parameters – response times and error rates – to disentangle underlying processes. As the first manuscript (see Appendix A) discusses in detail, our research suggests that diffusion-model analysis represents a useful tool to examine processes of the IAT and to dissociate meaningful components. In particular, IAT effects have been found to be determined by compatibility effects in drift rates (IAT_v), in speed-accuracy settings (IAT_a), and in nondecision components of processing (IAT_t). Importantly, construct-specific variance as measured in terms of correlations with self-reports was selectively mapped on IAT_v, whereas method-specific variance as measured in terms of correlations with control IATs was selectively mapped on IAT_a. In the following, possible extensions and limitations of the approach by diffusion-model analysis are discussed. It is avoided to reiterate the arguments already made by Klauer et al. (2007, Appendix A); instead, some of their considerations are extended and further aspects are added.

Reliability. To evaluate the usefulness of diffusion-model analysis, reliability estimates for the identified process components of the IAT are required. This is especially important if researchers want to use the IAT's process components as meaningful predictors for criterion variables. Reliability of parameters measuring the process components might be hypothesized to be lower than for the standard IAT effect, given that reliability is distributed across three components. First unpublished estimates, however, revealed acceptable internal consistencies for all three components (A. Voss, personal communication, June, 2008).

Even more interesting at the conceptual level are retest-reliability estimates. Such estimates could shed some light on the still unresolved puzzle of the IAT's retest-reliability typically being lower than its internal consistency. As discussed above, these differences have often been linked to the question whether the IAT should be considered as a measure of states or of traits. Retest-reliability estimates of the construct-related process component (i.e., IAT_v) might allow for stronger conclusions regarding this issue. For instance, satisfactory retest-reliability estimates for IAT_v would support the assumption that in principle, the IAT is able to assess stable, trait-like constructs. If at the same time, the other two process components, IAT_a and IAT_t, showed lower retest-reliability estimates, the typically moderate retest-reliability of the IAT might be attributed to the (non-stable) construct-unrelated process components. On the other hand, lower retest-reliability estimates for IAT_v in the range of what is typically found for IAT effects, would support the assumption that the IAT tends to assess context-dependent, state-like constructs.

Further validation of the IAT's process components. Klauer et al. (2007, Appendix A) were the first to apply diffusion-model analysis to IAT data. Accordingly, validation of the process components is in its infancy and further studies are needed in support of the present findings. For example, the speed-accuracy component (i.e., IAT_a) might have been especially pronounced in our studies because participants completed the different IATs in immediate succession. This might have facilitated a direct transfer of speed-accuracy settings from one IAT to the next. Future studies will have to investigate whether method-specific variance still maps on IAT_a, if IATs are not completed in immediate succession.

Further studies are also needed to prove whether the assumption can be upheld that construct-related variance maps on IAT_v. Given the conceptual uncertainty of how to interpret implicit-explicit correlations (see above), such studies should also adopt other approaches both at the group and at the individual level to further investigate the validity of IAT_v. Following the known-group approach, for example, one would expect the groups to differ in the process component IAT_v, but not necessarily in the other process components IAT_a and IAT_t. Similarly, IAT_v should predict meaningful behavior over and above other (e.g., explicit) measures, whereas the other process components should not necessarily relate to such criterion variables. Finally, a comparison of IAT_v and standard IAT scores (i.e., conventional score and *D* measure, respectively) might shed light on the question which measure provides better validity estimates. Fortunately, a huge data base of studies on the IAT's validity is already available, just awaiting re-analyses by diffusion models.

Separating confounding influences on the IAT. When validating the IAT's process components, research should also examine effects of the confounding influences on the IAT (see above). For example, one strength of Klauer et al.'s (2007, Appendix A) findings was that a known contaminant of the IAT effect – method specific variance – could be separated from the construct-related component, thereby controlling for its confounding influence. As argued above, this finding awaits replication with another design in order to investigate the processes underlying method-specific variance (i.e., cognitive abilities and/or speed-accuracy trade offs). Apart from that, research should examine whether other contaminants can also be separated.

For instance, as already discussed by Klauer et al. (2007, Appendix A), faking attempts might be identifiable by diffusion-model analysis. If faking in the IAT relies on strategic delaying of responses in the compatible phase, it should map on the nondecision component IAT_t. If faking relies on strategic settings of speed-accuracy trade-offs, it should map on the speed-accuracy component IAT_a. Accordingly, strategic effects might be partialled out of IAT_v in a diffusion-model analysis, leading to less contaminated IAT effects. As another example, diffusion-model analysis might also contribute to a better understanding of the processes underlying compatibility-order effects in the IAT.

At the same time, diffusion-model analysis might fail to separate confounding influences when it comes to strategic effects in the sense of Rothermund and Wentura's (2004) strategic recoding account. Following this account, participants might simplify the IAT task and decide to categorize stimuli according to categories other than those suggested by the IAT. As a consequence of such a strategic recoding, the IAT might assess associations between these re-defined categories rather than associations between the categories provided by the experimenter (cf. Govan & Williams, 2004). Accordingly, IAT_v as a possibly construct-related component might not primarily reflect the to-be-measured construct, but rather the construct that is assessed as a result of (strategic) recoding.

Direct comparisons with the quad model. Until now, the diffusion model and the quad model (Conrey et al., 2005) are the only models that have been applied to decompose the IAT effect into different components. Both models use different approaches and – as discussed by Klauer et al. (2007, Appendix A) – the components they reveal are not simply related. Nevertheless, it might be interesting to compare the two accounts empirically. For instance, it could be investigated which model better accounts for strategic effects in the IAT such as faking (e.g., see Fiedler & Bluemke, 2005). Similarly, it could be tested on which components the two models map so-called malleability effects in the IAT (see Blair, 2002).

From a theoretical perspective, the quad model has some limitations compared to the diffusion model. Other than the diffusion model, the quad model accounts only for the accuracy data of IATs and does not deal with the latency data. Accordingly, in contrast to the diffusion model, the quad model does not make use of the full information available. As another consequence, a certain proportion of errors is a prerequisite for reliable estimates of the quad-model parameters. Because participants reveal often less than 5% incorrect responses in a standard IAT, some procedural changes of the IAT (e.g., implementation of a response-time window) might be necessary in order to increase error rates to a level sufficient for quad-model analyses. Related to this problem of too little error variance – at least at the individual level – is the problem that some components estimated by the quad model show large confidence intervals (e.g., Conrey et al., 2005). Such findings question the suitability of the quad model for disentangling components at the individual level. It is important to note, however, that not only the quad model, but also the diffusion model requires some, albeit not many errors to achieve acceptable model fit. Last but not least, the diffusion model might be advantageous, because it relies on an established model of decision processes in binary decision tasks (Ratcliff, 1978; Ratcliff & Smith, 2004). Diffusion-model analysis thus provides a theoretically well-grounded means to separate process components.

Reliable estimation. As is also true for the quad model, the diffusion model requires a certain number of trials to obtain stable parameter estimates. The number of trials that is typically used in the standard format of an IAT (comprising about 72 trials) is at the lower bound of what is required for diffusion-model analysis. Future studies will therefore have to clarify in how far diffusion-model analyses of the standard IAT provide sufficiently stable estimates. It might turn out that in order to apply diffusion-model analyses, trial numbers of the IAT have to be increased considerably. This would also dampen the optimism expressed above that reanalyzing the available studies on the IAT's validity by means of the diffusion model might be sufficient to further validate the process components.

In summary, diffusion-model analysis appears to be a promising tool in IAT research that allows for the investigation of several research questions. In the final section, the usefulness of the second proposed approach to reduce the IAT's contamination is considered in more detail.

Eliminating the IAT's Block Structure

Many of the identified confounds of the IAT have been shown to originate from the IAT's block structure. The consistently blocked mapping of categories onto response keys in the compatible versus incompatible block of the IAT appears to promote different processes in the two blocks. Because the IAT effect is based on a comparison of performance in the two separate IAT blocks, such differences in processing directly influence the IAT effect in a confounding manner. A straightforward remedy therefore seems to be the elimination of the IAT's block structure in order to reduce such confounding influences.

Elimination of the block structure was realized in two paradigms called Single Block IAT (SB-IAT) and recoding free IAT (IAT-RF) both of which are introduced in the second (see Appendix B) and the third (see Appendix C) manuscript, respectively. Both paradigms showed reduced susceptibility to known confounding influences of the IAT such as cognitive abilities and stimulus influences, thereby indicating effectiveness and usefulness of such a structural change. As has also been done for the approach by diffusion-model analysis, in the following, possible extensions and limitations of the approach by eliminating the block structure are discussed. Again, it will be avoided to reiterate the arguments already made in the manuscripts (Appendixes B and C).

Further validation of the SB-IAT and the IAT-RF. Both manuscripts strongly focus on providing evidence for the assumption that eliminating the IAT's block structure reduces the impact of the known contaminants on the IAT effect. In addition, some attention has been paid to the validation of the newly developed paradigms. Three validation approaches have been adopted: At the group level, both the SB-IAT and the IAT-RF proved to be valid with regard to universal attitudes showing the to-be-expected preference for flowers over insects. Also, a political attitude SB-IAT distinguished between groups of participants who indicated to vote for different political spectrums. At the individual level, the political attitude SB-IAT correlated with self-reports of political attitude, thus exhibiting implicit-explicit consistency.

Naturally, these findings provide only first evidence for the validity of the SB-IAT and the IAT-RF, and more research is needed to explore their usefulness as measures of implicit attitudes, stereotypes, self-esteem, or personality traits. Given the lively debate on how to interpret implicit-explicit correlations (see validity section), future research should also adopt other validation approaches. For instance, neither the SB-IAT nor the IAT-RF has yet been shown to predict behavior. Validity of both procedures could thus be convincingly demonstrated if the SB-IAT and the IAT-RF were able to predict behavior over and above explicit measures following the additive, the multiplicative, or the double dissociation model as proposed by Perugini (2005).

In this regard, it might be particularly interesting to investigate domains in which the IAT typically shows weak predictive validity. For example, domains related to health behavior might be a suitable candidate. As discussed above, the IAT's insufficient validity estimates in such domains have been argued to be the result of contaminations by extra-personal knowledge (see De Houwer, Custers, & De Clercq, 2006; Olson & Fazio, 2004;

Spruyt et al., 2007). Following Rothermund and Wentura's (2004) strategic recoding account (see also Teige-Mocigemba et al., 2007, Appendix B), such confounds might be due to the *strategic* use of extra-personal knowledge in order to improve performance on the IAT task. Strategic recoding has been argued to rely on the IAT's block structure. Accordingly, inasmuch as the IAT's contamination by extra-personal knowledge results from strategic recoding, one would expect the SB-IAT and the IAT-RF to be less affected by this confound. Initial evidence from unpublished experiments indeed provided indirect support for the assumption that the SB-IAT is less contaminated by extra-personal knowledge: In contrast to the IAT, a smoking attitude SB-IAT was able to distinguish between smokers and non-smokers (known-group approach) and correlated with the explicit attitude towards smoking as assessed by self-reports (Teige-Mocigemba & Klauer, 2008). Given these preliminary but promising findings, the SB-IAT and the IAT-RF might prove superior to the IAT when it comes to the prediction of health-related behavior.

Preventing confounding factors of the IAT. As is shown in the manuscripts (Appendixes B and C), eliminating the IAT's block structure seems to be sufficient to reduce the confounding influence of task-switching costs, method-specific variance, and biased selection of stimuli. However, these are only some of many identified contaminants of the IAT. Accordingly, future research should focus on the other contaminants not investigated yet. For some such factors, theoretical considerations make empirical testing redundant: For instance, confounding effects of *compatibility order* should be excluded in the SB-IAT and the IAT-RF because this effect clearly relies on the IAT's block structure which both procedures eliminate. In contrast, – as is also true for the IAT – the SB-IAT and the IAT-RF in their present form allow only for assessment of *relative* association strengths between categories, including all problems related to this issue (see above). Note however, that we conducted first experiments in which we successfully employed a Single Category IAT variant that eliminates the block structure (Teige-Mocigemba & Klauer, 2008). This Single Category SB-IAT has the advantage that it combines two desirable modifications of the IAT: First, it eliminates the block structure, thereby reducing the influence of several confounding factors on the IAT and the Single Category IAT (see Karpinski & Steinman, 2006). Second, similar to the Single Category IAT, it allows for the assessment of attitude categories that do not have a natural counterpart.

It still needs to be tested empirically, however, to which degree confounding influences of salience, similarity, strategic effects, and extra-personal associations are reduced by elimination of the IAT's block structure. With regard to *salience* and *similarity*, one might expect the SB-IAT and the IAT-RF to be less affected because these confounds are largely attributed to the IAT's block structure as argued above. Similarly, the SB-IAT and the IAT-RF might be less influenced by *strategic effects*, at least inasmuch as these effects are based on strategic recoding in the sense of Rothermund and Wentura's (2004) account. Note however, that random changes of compatible and incompatible trials within one block – as is the case in the SB-IAT and the IAT-RF – do not necessarily prevent strategic influences. For instance, recent studies provided evidence for strategic effects on affective priming, a procedure which also manipulates compatibility in a trial-wise manner (Klauer & Teige-Mocigemba, 2007; Teige-Mocigemba & Klauer, in press). Such findings indicate that strategic recoding in the sense of Rothermund and Wentura (2004) is not the only process by which participants can strategically influence the outcome of implicit measures.

Indirect evidence for a reduced contamination by *extra-personal associations* in the SB-IAT came from the correlational studies sketched above which indicated superior validity estimates of the SB-IAT as compared to the IAT in the health-related domain of smoking attitudes. Of course, the conclusiveness of these studies is limited because there might be several reasons other than extra-personal knowledge that could account for the findings. Hence, future studies should adopt other approaches to investigate whether the SB-IAT and the IAT-RF are confounded by extra-personal knowledge. Following Han et al. (2006), for instance, extra-personal views might be manipulated experimentally. If the SB-IAT and the IAT-RF proved to be immune against such manipulations, this could be interpreted as evidence for a reduced influence of extra-personal associations.

Last but not least, it should be noted that the SB-IAT and the IAT-RF may have their own shortcomings and confounds to which the IAT is not subjected. For instance, several procedural parameters still require careful testing: Possible confounding influences on the SB-IAT and the IAT-RF might emanate from (a) the presentation time of the fixation cross, (b) the switching between the compatible and incompatible mapping, or (c) the unavoidable confound of word position and compatibility condition in the SB-IAT, to name just a few.

Comparing the SB-IAT and the IAT-RF. Both the SB-IAT and the IAT-RF have been shown to reduce the influence of some confounding factors on the IAT and have been argued to prevent contaminations by other factors, too (see previous section). However, it has not been tested yet whether the results found with one procedure replicate with the other. Nor has it been investigated whether one paradigm is superior to the other. Accordingly, there is an obvious need for research comparing the SB-IAT and the IAT-RF. At the structural level, the clear buildup of the SB-IAT might be an advantage: In contrast to the IAT-RF, the SB-IAT provides participants with a structural feature (i.e., word position) that signals the mapping of categories onto response keys for each trial. In the IAT-RF, such a feature is absent: In each trial, participants are required to read the category labels which indicate the mapping for the respective trial. Accordingly, participants might find it easier to perform the SB-IAT than the IAT-RF. This might be expressed in lower non-systematic error variance of SB-IAT effects which should enhance reliability estimates. On the other hand, the structural feature of the SB-IAT might also be a disadvantage when it possibly serves as a cue misused for the application of strategies that help to simplify the categorization task. Doubtlessly, empirical studies are needed that pit the two paradigms against each other and provide clear recommendations under which circumstances which procedure is to be preferred.

Testable process models. As noted above, for the IAT, it has been criticized that before developing a comprehensive, testable process model, countless studies already applied the IAT to diverse psychological areas without a clear understanding of what the procedure actually measures (cf. Fiedler et al., 2006; Wentura & Rothermund, 2007). Research on the SB-IAT and the IAT-RF should be prevented from running the same risk of becoming empirically driven research without the background of a comprehensive and testable process model that accounts for the relevant factors contributing to the effects. Of course, the present findings corroborate the hypotheses which led to the development of the SB-IAT and the IAT-RF. They do not, however, directly examine the processes underlying the effects in the two paradigms. Instead of inviting researchers to substitute the SB-IAT and/or the IAT-RF for the IAT and to continue applying the new measures blindly to diverse areas, we urge researchers to contribute to the development and testing of comprehensive process models of the SB-IAT and the IAT-RF.

For instance, relevant stimulus-response compatibilities (De Houwer, 2001, 2003b, in press) might account for SB-IAT effects and/or IAT-RF effects (see Rothermund et al., in press, Appendix C). According to this mechanism, compatibility effects may result from an overlap between relevant stimulus features (i.e., their target category membership such as Black vs. White) and response characteristics that are established during the task by assigning a specific attribute category to a response (e.g., right key acquires negative valence, left key acquires positive valence). Effects in the SB-IAT and the IAT-RF might therefore reflect genuine compatibilities between the nominal target and attribute categories of the task. It is beyond question, however, that further research is needed to test whether processes of relevant stimulus-response compatibilities can account for SB-IAT effects and/or IAT-RF effects. Finally, diffusion-model analysis might also contribute to a deeper understanding of the processes underlying the SB-IAT and the IAT-RF.

Conclusion

It has been argued that both the approach by diffusion-model analysis and the approach by eliminating the IAT's block structure provide useful tools to obtain less contaminated IAT scores. Both approaches also allow for the investigation of several unresolved puzzles in IAT research, and they raise further interesting research questions. Future research on the usefulness of the two approaches might help to estimate how far the proposed approaches provide suitable measures of those "unconscious" parts of the self that self-reports cannot reveal. I am confident that both approaches – at least to some extent – contribute to the solution of the problem St. Augustine already posed in the early 5th century when he lamented: "I cannot totally grasp all that I am…" (St. Augustine, trans. 1944).

References

Antonak, R. F., & Livneh, H. (1995). Randomized response technique: A review and proposed extension to disability attitude research. *Genetic, Social, and General Psychology Monographs*, *121*, 97-145.

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or “Would Jesse Jackson ‘fail’ the Implicit Association Test?”. *Psychological Inquiry*, *15*, 257-278.

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.

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

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

Banse, R., & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Skaleneigenschaften und Validierung [The scale motivation to act without prejudice: Psychometric properties and validity]. *Diagnostica*, *49*, 4-13.

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

Baron, A. S. & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6, 10, & adulthood. *Psychological Science*, *17*, 53-58.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242-261.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, *42*, 192-212.

Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects? *Journal of Experimental Social Psychology*, *42*, 163-176.

Bohner, G., & Wänke, M. (2002). *Attitudes and attitude change*. Hove: Psychology Press.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*, 1061-1071.

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.

Bradburn, N., Sudman, S., & Wansink, B. (2004). *Asking questions* (2nd ed.). San Francisco, CA: Jossey Bass.

Brendl, M., Markman, A., & 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.

Chaiken, S., & Trope, Y. (1999). *Dual process theories in social psychology*. New York: Guilford.

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

Conrey, F. R., & Smith E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, *25*, 718-735.

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

Czellar, S. (2006). Self-presentational effects in the Implicit Association Test. *Journal of Consumer Psychology*, *16*, 92-100.

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. (2003a). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77-85.

De Houwer, J. (2003b). 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: Erlbaum.

De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 11-28). Thousand Oaks, CA: Sage.

De Houwer, J. (in press). Comparing measures of attitudes at the procedural and functional level. In R. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures*. New York, NY: Psychology Press.

De Houwer, J., Beckers, T., & Moors, A. (2007). Novel attitudes can be faked on the Implicit Association Test. *Journal of Experimental Social Psychology*, *43*, 972-978.

De Houwer, J., Custers, R., & De Clercq, A. (2006). Do smokers have a negative implicit attitude towards smoking? *Cognition and Emotion*, *20*, 1274-1284.

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

De Houwer, J., & Moors, A. (2007). How to define and examine the implicitness of implicit measures. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 179-194). New York, NY: Guilford Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2008). *Implicit measures: A normative analysis and review*. Manuscript submitted for publication.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.

Dunham, Y., Baron, A., Banaji, M. R. (2007). Children and social groups: A developmental analysis of implicit consistency in Hispanic Americans. *Self and Identity*, *6*, 238-255.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316-326.

Eder, A., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General*, *137*, 262-281.

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. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75-109). San Diego, CA: Academic Press.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013-1027.

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

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229-238.

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

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basis and Applied Social Psychology*, *27*, 307-316.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A" and the "T": Logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*, 74-147.

Freud, S. (1915). Das Unbewußte (The unconscious). In S. Freud (Ed.), *Das Ich und das Es und andere metapsychologische Schriften (1960)* (pp. 73-104). Frankfurt a. M.: Fischer Taschenbuch.

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of implicit and explicit attitude measures in predicting food choice and consumption behavior. *British Journal of Social Psychology*, *47*, 397-419.

Friese, M., Wänke, M., & Plessner, H. (2006). Implicit consumer preferences and their influence on product choice. *Psychology & Marketing*, *23*, 727-740.

Gawronski, B. (Ed.). (2007). What is an attitude? [Special issue]. *Social Cognition*, *25(5)*.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692-731.

Gawronski, B., Peters, K. R., & LeBel, E. P. (2008). What makes mental associations personal or extra-personal? Conceptual issues in the methodological debate about implicit attitude measures. *Social and Personality Psychology Compass*, *2*, 1002-1023.

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

Gray, N. S., MacCulloch, M. J., Smith, J., Morris, M., & Snowden, R. J. (2003). Violence viewed by psychopathic murderers. *Nature*, *423*, 497-498.

Greenwald, A. G., & Banaji, M., R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4-27.

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. (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.

Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the salience asymmetry interpretation of the IAT: Comment on Rothermund and Wentura. *Journal of Experimental Psychology: General*, *134*, 420-425.

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

Gschwendner, T., Hofmann, W., & Schmitt, M. (2008). Differential stability: The effects of acute and chronic construct accessibility on the temporal stability of the Implicit Association Test. *Journal of Individual Differences*, *29*, 70-79.

Häcker, H., Schmidt L. R., Schwenkmezger, P., & Utz, H. E. (1975). *Objektive Testbatterie, OA-TB 75* [*Objective test battery*]. Weinheim: Beltz.

Han, H. A., Olson, M. A., & Fazio, R. H. (2006). The influence of experimentally created extrapersonal associations on the Implicit Association Test. *Journal of Experimental Social Psychology*, *42*, 259-272.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369-1385.

Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit-explicit consistency? *European Review of Social Psychology*, *16*, 335-390.

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology*, *43*, 497-504.

Houben, K., & Wiers, R. W. (2006). A test of the salience asymmetry interpretation of the alcohol-IAT. *Experimental Psychology*, *53*, 292-300.

Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the Implicit Association Test to measure age differences in implicit social cognitions. *Psychology and Aging*, *17*, 482-495.

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*, 349-364.

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

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*, 16-32.

Kawakami, K., & Dovidio, J. F. (2001). Implicit stereotyping: How reliable is it? *Personality and Social Psychology Bulletin*, *27*, 212-225.

Kim, D. Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly*, *66*, 83-96.

Kinoshita, S., & Peek-O'Leary, M. (2005). Does the compatibility effect in the race Implicit Association Test (IAT) reflect familiarity or affect? *Psychonomic Bulletin and Review*, *12*, 442-452.

Kinoshita, S., & Peek-O'Leary, M. (2006). Two bases of the compatibility effect in the Implicit Association Test (IAT). *Quarterly Journal of Experimental Psychology*, *59*, 2102-2120.

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

Klauer, K. C., Musch, J., & Eder, A. (2005). Priming of semantic classifications: Late and response-related, or earlier and more central? *Psychonomic Bulletin and Review*, *12*, 897-903.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluations. *Journal of Experimental Social Psychology*, *43*, 648-655.

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (in press). Contrast effects in spontaneous evaluations: A psychophysical account. *Journal of Personality and Social Psychology.*

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

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility – a model and taxonomy. *Psychological Review*, *97*, 253-270.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far). In B. Wittenbrink &

N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 59-102). New York: Guilford Press.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest*, *1*, 27-66.

Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology*, *14*, 405-415.

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

McConnell, A. R., & Liebold, 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.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254-278.

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.

Murray, H. A. (1943). *Thematic Apperception Test. Manual*. Cambridge, MA: Harvard University Press.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, *106*, 226-254.

Nier, J. A. (2005). How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach. *Group Processes and Intergroup Relations*, *8*, 39-52.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.

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

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

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

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

Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition and Emotion*, *22*, 553-594.

Nosek, B. A., & Hansen, J. J. (in press). Personalizing the Implicit Association Test increases explicit evaluation of target concepts. *European Journal of Psychological Assessment*.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology*, *54*, 14-29.

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

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

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

Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*, 16-31.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277-293.

Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, *44*, 29-45.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811-832.

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.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*, 416-442.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873-922.

Ratcliff, R. & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333-367.

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., & Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*, *6*, 1323-1328.

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, *114*, 363-375.

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

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

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. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*, 139-165.

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Validity of the salience asymmetry account of the IAT: Reply to Greenwald, Nosek, Banaji, and Klauer (2005). *Journal of Experimental Psychology: General*, *134*, 426-430.

Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. P., & Sedikides, C. (in press). Through a glass, less darkly? Reassessing convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment.*

Schmukle, S. C., Back, M. D., & Egloff, B. (in press). Validity of the five-factor model for the implicit self-concept of personality. *European Journal of Psychological Assessment.*

Schmukle, S. C., & Egloff, B. (2004). Does the Implicit Association Test for assessing anxiety measure trait and state variance? *European Journal of Personality*, *18*, 483-494.

Schnabel, K., Asendorpf, J. B., & Greenwald, A. G. (2008). Using Implicit Association Tests for the assessment of implicit personality self-concept. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *Personality theory and assessment (Volume 2): Personality Measurement and Testing* (pp. 508-528). Thousand Oaks, CA: Sage.

Schnabel, K., Banse, R., & Asendorpf, J. B. (2006a). Assessment of implicit personality self-concept using the Implicit Association Test (IAT): Concurrent assessment of anxiousness and angriness. *British Journal of Social Psychology*, *45*, 373-396.

Schnabel, K., Banse, R., & Asendorpf, J. B. (2006b). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: The Implicit Association Procedure (IAP). *Experimental Psychology*, *53*, 69-76.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007). High implicit self-esteem is not necessarily advantageous: Discrepancies between explicit and implicit self-esteem and their relationship with anger expression and psychological health. *European Journal of Personality*, *21*, 319-339.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93-105.

Schwarz, N. (2007a). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638-656.

Schwarz, N. (2007b). Cognitive aspects of survey methodology. *Applied Cognitive Psychology*, *21*, 277-287.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*, 314-335.

Sherman, S. J., Presson, C. C., Chassin, L., Rose, J. S., & Koch, K. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology*, *22*, 13-39.

Smith, E. R., & Conrey, F. R. (2007). Mental representations are “states”, not “things”: Implications for implicit and explicit measurement. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 247-264). New York: Guilford Press.

Smith, E. R., & DeCoster, J. (2000). Dual process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108-131.

Spruyt, A., Hermans, D., De Houwer, J., Vandekerckhove, J., & Eelen, P. (2007). On the predictive validity of indirect attitude measures: Prediction of consumer choice behavior on the basis of affective priming in the picture – picture naming task. *Journal of Experimental Social Psychology*, *43*, 599-610.

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology*, *51*, 165-179.

Steffens, M. C., Kirschbaum, M., & Glados, P. (in press). Avoiding stimulus confounds in Implicit Association Tests by using the concepts as stimuli. *British Journal of Social Psychology*.

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

St. Augustine. (1944). *The Confessions of St. Augustine* (F. J. Sheed, Trans.). London-New York.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220-247.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers. The application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.

Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behaviour consistency for stigmatised behaviour. *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*, *110*, 226-235.

Teachman, B. A., & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology*, *112*, 100-109.

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

Teige-Mocigemba, S., & Klauer, K. C. (2008, July). *Effects of compatibility order on the IAT's validity.* Paper presented at the 19th International Congress of Psychology, Berlin, Germany.

Teige-Mocigemba, S., & Klauer, K. C. (in press). "Automatic" evaluation? Strategic effects on affective priming. *Journal of Experimental Social Psychology*.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (in press). Minimizing method-specific variance in the IAT: A Single Block IAT. *European Journal of Psychological Assessment*.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*, 1206-1220.

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

Wiers R.W., van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, *111*, 648-658.

Wilson, T., Lindsay, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101-126.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262-274.

Wittenbrink, B., & Schwarz, N. (2007). *Implicit measures of attitudes*. New York: Guilford Press.

Ziegler, M., Schmidt-Atzert, L., Bühner, M., & Krumm, S. (2007). Fakeability of different measurement methods for achievement motivation: Questionnaire, semi-projective, and objective. *Psychology Science*, *49*, 291-307.

Footnotes

[1] Note that there is some ambiguity in the use of the terms explicit and implicit. In the literature, the term "explicit measure" is commonly used to refer to direct measures such as questionnaires or other kinds of self-reports. The term "implicit measure" is commonly used to refer to indirect measures such as response-time measures. It has been argued that this terminology makes it difficult to disentangle the empirical measurement level and the construct level. As a result, the outcome of direct and indirect measures might be inadmissibly equated with the underlying explicit and implicit constructs (De Houwer & Moors, 2007; Fazio & Olson, 2003). For these and other reasons, De Houwer (2006) regarded the terms "direct" and "indirect" as more appropriate. Although I acknowledge the problems posed by the wording, I decided to stick with the more commonly used terminology of explicit and implicit measures (see Fazio & Olson, 2003), as also done in the manuscripts (Appendixes A to C). In line with the literature, I will reserve the term implicit for indirect response-time measures (such as the IAT), whereas other indirect measures not based on response times (such as projective tests) will be denoted as indirect measures.

[2] The later introduced affect misattribution procedure (AMP) also shows satisfactory reliability estimates (Payne, Cheng, Govorun, & Stewart, 2005).

[3] In this regard, recent research by Gschwendner, Hofmann, and Schmitt (2008) appears to be enlightening. The authors investigated the closely related issue that the IAT's retest-reliability is typically lower than its satisfactory internal consistency which often has been interpreted as evidence for the assumption that IAT effects might rather reflect states than traits (but see Schmukle & Egloff, 2004). Gschwendner et al. emphasized the impact of construct accessibility on the IAT's temporal stability. In particular, they showed that (a) the IAT's retest-reliability was enhanced in situations in which contextual background features activated specific construct-relevant concepts and that (b) this effect was particularly pronounced for individuals with chronically high accessibility for the relevant concept. These findings suggest that the IAT can be adapted to assess traits that might be better conceptualized as interindividually different, temporally stable patterns of associative activation than as fixed and invariant associative structures (cf. Conrey & Smith, 2007; Smith & Conrey, 2007).

Appendixes

A: Diffusion-Model Analysis of the IAT: Klauer, Voss, Schmitz, and Teige-Mocigemba (2007)

B: Single Block IAT: Teige-Mocigemba, Klauer, and Rothermund (in press)

C: Minimizing Recoding with the IAT-RF: Rothermund, Teige-Mocigemba, Gast, and Wentura (in press)

Appendix A

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

Process Components of the Implicit Association Test: A Diffusion-Model Analysis

Karl Christoph Klauer

Andreas Voss

Florian Schmitz

Sarah Teige-Mocigemba

Albert-Ludwigs-Universität Freiburg

2007, *Journal of Personality and Social Psychology*, *93*, 353-368.

Abstract

The authors present a diffusion-model analysis of the Implicit Association Test (IAT). In Study 1, the IAT effect was decomposed into 3 dissociable components: Relative to the compatible phase, (a) ease and speed of information accumulation are lowered in the incompatible phase, (b) more cautious speed-accuracy settings are adopted, and (c) nondecision components of processing require more time. Studies 2 and 3 assessed the nature of interindividual differences in these components. Construct-specific variance in the IAT relating to the construct to be measured (such as implicit attitudes) was concentrated in the compatibility effect on information accumulation (Studies 2 and 3), whereas systematic method variance in the IAT was mapped on differential speed-accuracy settings (Study 3). Implications of these dissociations for process theories of the IAT and for applications are discussed.

Keywords: implicit measures, implicit social cognition, Implicit Association Test, diffusion model, mathematical model

In recent years, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) has attracted an enormous amount of research interest and debate (for recent reviews, see Fazio & Olson, 2003; Nosek, Greenwald, & Banaji, 2006). The IAT involves two tasks: a concept task, in which exemplars of two target concepts (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 a pair of attribute categories (e.g., as either “positive” or “negative”). In the critical phases of the IAT procedure, both tasks are to be performed in alternation and they are mapped onto the same response keys, which can be done in two 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 compatible mapping, and the other one is called incompatible mapping. The performance difference between the two kinds of mappings is known as the IAT effect.

The IAT rests on the assumption that it is easier to make the same behavioral response to concepts and attributes that are strongly associated than to concepts and attributes that are weakly associated. In this view, direction and size of the IAT effect indicate the relative association strengths between the target concepts and attributes.

We present a diffusion-model analysis of the IAT, applying the diffusion model for two-choice decisions (e.g., Ratcliff, Gomez, & McKoon, 2004; Ratcliff & Rouder, 1998) to the IAT. The model disentangles several components of processing: the rate at which information about the stimulus is accumulated in the decision system; the criteria or thresholds that determine the amount of information that must be accumulated before a decision can be made; nondecision components of processing, such as encoding, task-set retrieval, and response execution; and variabilities in the various components. The diffusion model simultaneously accounts for reaction time distributions of correct and incorrect responses and the error rates. In particular, latency data and accuracy data are mapped onto a common metric in a principled manner (Spaniol, Madden, & Voss, 2006).

By decomposing the IAT effect into separate process components, the diffusion-model analysis contributes to understanding the nature of the processes underlying the IAT effect. In a more applied vein, we address the question of whether different variance components of the IAT map on different process components. The IAT is known to contain method variance, as indicated, for example, by correlations between unrelated IATs (Back, Schmukle, & Egloff, 2005; McFarland & Crouch, 2002; Mierke & Klauer, 2003). IAT measures also contain construct-specific variance relating to the construct to be measured, as indicated, for example, by correlations between IATs and self-report measures of such constructs in many domains (Hofmann, Gawronski, Geschwender, Le, & Schmitt, 2005).

The Diffusion Model

The diffusion model is a model for two-choice decisions. It is assumed that decisions are based on a process of information accumulation over time. Information accumulates on a response-related decision axis on which a starting point lies between two response thresholds, as shown in Figure 1. Each response threshold is associated with one of the two responses. The diffusion process moves from the starting point until one of the thresholds is reached and the response associated with it is initiated. The average rate of accumulation of information is called the drift rate (parameter v). Noise in the accumulation process implies that processes with the same mean drift rate do not always terminate at the same time (producing reaction time distributions) and do not always terminate at the same threshold (causing errors). This variability is called within-trial variability (Ratcliff, Thapar, Gomez, & McKoon, 2004).

In fitting the diffusion model to data, the components of processing are also assumed to be variable. Variability in drift rate across trials reflects the fact that the different stimuli are not equivalent sources of response-related information. For example, in a flower-insect IAT, a familiar flower, such as rose, probably supports a faster information-accumulation process toward the flower threshold than does an unfamiliar one, such as hydrangea. This variability is modeled by assuming a normal distribution of drift rates with mean v and standard deviation η across trials. In addition, variability in the starting point (parameter z) and in the nondecision components (parameter t_0) is necessary to account for the shapes of the reaction time distributions of errors and correct responses. For both components, uniform distributions are assumed with ranges s_z and s_t, respectively, following previous work with the model (e.g., Ratcliff, Gomez, & McKoon, 2004).

The model parameters are a, z, v, t_0, η, s_z, and s_t. Parameters a, z, v, and t_0 capture different aspects of the decision-making process.

Parameter a is the separation of the two thresholds and thus quantifies the amount of evidence that must accumulate before a response is initiated. Parameter a thereby reflects participants' speed-accuracy trade-off settings: Large values of a indicate conservative speed-accuracy settings because much information must accumulate before a (slow and accurate) decision is made, whereas small values of a reflect liberal speed-accuracy settings, resulting in faster but less accurate responses.

Parameter z is the starting point of information accumulation. Divided by a, it ranges from 0 to 1 and measures response bias toward one relative to the other of the two responses. For example, a starting point close to the upper threshold implies that comparatively little additional information must accumulate toward the upper threshold before it is crossed and the response associated with it is initiated; conversely, comparatively more information must accumulate toward the lower threshold before it can be crossed. The result is a response bias toward the response associated with the upper threshold and against the response associated with the lower threshold.

Parameter v is the mean drift rate. Drift rate quantifies the direction (toward lower vs. upper threshold) and the speed with which relevant information accumulates. Drift rate thereby determines the decision maker's performance in the decision process itself, high speed of information accumulation implying both fast and accurate decisions. In comparisons between participants, parameter v reflects interindividual differences in the ease of decision making; in comparisons between experimental conditions within participants, it reflects differences in task difficulty. For example, if the task is to decide which of two briefly shown lines is the longer one, the speed with which relevant information accumulates is a function of the individual's sensory ability (people high in visual acuity exhibiting higher drift rates) and of the difference in length between the two lines (stimuli with larger separation affording faster accumulation of response-related information).

The diffusion model elaborates on the decision process in some detail; the contribution of nondecision processes relating to, for example, preparatory encoding of stimuli and motor responses are summarized in nondecision component t_0. Parameter η is the standard deviation in mean drift rate across trials, s_z is the range of the starting point across trials, and s_t is the range of nondecision components t_0. Within-trial variability in drift rate is a scaling factor that was set equal to 1 in the present analyses. For easy reference, the parameters and their meanings are summarized in Table 1.

The diffusion model has been applied to a wide variety of tasks, including lexical decisions (Ratcliff, Gomez, & McKoon, 2004; Ratcliff, Thapar, et al., 2004), memory retrieval (Ratcliff, 1978, 1988), visual signal detection, numerosity judgments, distance judgments (Ratcliff, Thapar, & McKoon, 2003), and animacy categorization (Spaniol et al., 2006), among others. A large body of research (e.g., Ratcliff, 1985; Ratcliff, Van Zandt, & McKoon, 1999; Voss, Rothermund, & Voss, 2004), reviewed by Spaniol et al. (2006), successfully used the model to disentangle the processes captured by the different model parameters.

The diffusion-model analyses allow us to evaluate the roles of the different process components in causing IAT effects. Existing process theories of the IAT can thereby be tested in a more direct manner than was possible before. Another, more applied purpose of the diffusion-model analyses is to investigate how construct-specific variance and method variance in the IAT are reflected in the different process components.

Brendl, Markman, and Messner's (2001) Random-Walk Account

Brendl et al. (2001) hypothesized that information processing in the IAT follows the diffusion model sketched above. More precisely, they assumed that both attribute information and concept information drive the accumulation process. Consider, for example, a flower-insect IAT and a concept stimulus such as rose that is both a flower and positively valenced. In the compatible phase, both sources of evidence, that is the positive valence (attribute information) as well as the membership in the flower category (concept information), move the accumulation process toward the correct response, namely the common response for flowers and positive stimuli. In contrast, in the incompatible phase, the two sources of evidence add to the accumulation process in opposite directions, because flowers and positive stimuli are now mapped on different responses. For this reason, Brendl et al. assumed that drift rate for concept stimuli will be lower in the incompatible block than in the compatible block. Brendl et al. did not, however, expect a compatibility effect on drift rate for the attribute stimuli because there is typically only attribute information but no concept information for such stimuli. For example, the positive attribute stimulus peace is neither insect nor flower.

In this view, attitudes enter the IAT via the concept stimuli. Attitudes associated with the concept stimuli add to the drift rate in compatible blocks and subtract from it in incompatible blocks. This leads to faster and more accurate responses to concept stimuli in the compatible relative to the incompatible block, in proportion to the strength of the attitude associations of the concepts.

According to Brendl et al. (2001), the compatibility effect on drift rate for concept stimuli is assumed to be a valid component of the IAT effect: It correctly reflects the respondents' attitudes or prejudices in an attitude or prejudice IAT, more generally the relative strength of associations between target concepts and attributes. This is not true of a compatibility effect that Brendl et al. postulated for the speed-accuracy settings (i.e., for threshold-separation parameter a). Brendl et al. argued that participants adopt a conservative speed-accuracy setting to the extent to which they perceive a block of trials as difficult. For the IAT, their main hypothesis is that people use different speed-accuracy settings in different phases of the IAT, “which compromises the interpretation of this test as a measure of individual differences in implicit prejudice” (Brendl et al., 2001, p. 769). In this view, the IAT effect is partly caused by the adoption of more conservative speed-accuracy settings in the incompatible phase than in the compatible phase because the former is perceived to be more difficult than the latter.

To summarize, according to Brendl et al. (2001), a compatibility effect is expected for drift rate v in the concept task but not in the attribute task. In addition, a compatibility effect on speed-accuracy parameter a is expected. Alternative process theories of the IAT, such as the account by task-set switching (Klauer & Mierke, 2005; Mierke & Klauer, 2001, 2003) and the quad model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005), are considered in the General Discussion.

Study 1: Process Components of the IAT

In Study 1, a flower-insect IAT was subjected to a diffusion-model analysis to disentangle the process components that might contribute to the IAT effect. A major driving force of IAT research is the fact that IATs often reveal stable interindividual differences in content domains of interest. To be able to capture individual differences, we fitted the diffusion model separately to the data from each participant (see Spaniol et al., 2006) in most of the diffusion-model analyses reported in this article. To obtain stable parameter estimates for concept and attribute task, the flower-insect IAT in Study 1 used more trials than usual. The IATs analyzed in subsequent studies followed standard IAT procedures.

Method

Participants. Twenty University of Freiburg (Freiburg, Germany) students (13 women, 7 men) with different majors participated. Their mean age was 24 years, with a range from 20 to 34 years. All participants were native speakers of the German language and had normal or corrected-to-normal vision. Participants received a gratification of 3.50 euros (U.S. $4.71) for participating.

Materials. The same materials were used as in the flower-insect IATs developed by Mierke and Klauer (2001, 2003). Thus, there were 96 German nouns (24 flower names, 24 insect names, 24 positive words, and 24 negative words) that were matched across categories for word frequency, number of letters, and rated valence. Stimuli were presented in a sans serif font and had a height of 7 mm and a width of 12 mm to 85 mm. Concept stimuli were presented in dark gray and attribute stimuli in black on a light gray background.

Procedure. The present variant of the IAT consisted of five blocks with more trials per block than in standard implementations of the IAT. Participants used the left key A and the right key L on a standard computer keyboard to respond. In Block 1, the concept task was practiced: This block consisted of 96 trials (preceded by eight warm-up trials), across which each insect name and each flower name were presented twice. Participants were to respond with the left key to stimuli belonging to the insect category and with the right key to stimuli from the flower category. In Block 2, the attribute task was practiced. Again, there were 96 trials, across which each attribute stimulus was presented twice, plus 8 warm-up trials. Participants were to respond with the left key to negative stimuli and with the right key to positive stimuli. Block 3 was the so-called compatible block: All 96 stimuli were presented twice, resulting in a total of 192 trials, which were again preceded by eight warm-up trials. Insects and negative words were mapped on the left key and flowers and positive words on the right key. Concept stimuli and attribute stimuli were presented in strictly alternating, but otherwise randomized, order. Block 4 was a practice block for the concept task with reversed response mapping that was otherwise constructed like Block 1. Block 5 was the so-called incompatible block. This block was essentially the same as Block 3. However, the flower category and the negative attribute category were now mapped on one key (i.e., the left key) and the insect and positive attribute categories on the other key (i.e., the right key).

A trial started with the presentation of a fixation cross. After 300 ms, the cross was replaced by the target stimulus, which remained on the screen until the correct key was pressed. In case of a false response, the string XXX was shown in red directly under the stimulus until the correct response was given. Following each block, feedback summarized the mean reaction time and the percentage of correct first responses in that block. We recorded first response and its latency as well as latency to correct response after an error response.

Model analyses. Because the IAT affords only relatively few trials per person, we coded responses as correct versus false, aggregating over the two particular target categories involved (i.e., flowers and insects) for the analyses of the concept task as well as over the two attribute categories (i.e., positive and negative) for the analyses of the attribute task. This implies that possible response biases (e.g., a bias to respond positive) cancel out in the present analyses because across trials they favor the false response equally often as the correct response. As a consequence, parameter z for the mean starting point was set equal to $a/2$ (see also Ratcliff, 2002), that is, equal to the position corresponding to the absence of response bias. Possible response biases do, however, contribute to variability s_z in the starting point z in the present analyses.

The diffusion model was fitted separately to the data from each of the five blocks of the IAT, including the practice blocks, separately for each task (concept task vs. attribute task), yielding seven model analyses per participant based on 96 trials each. Analyzing the data from the practice blocks allowed us to gauge in an exploratory fashion whether there were pronounced changes due to practice or fatigue independently of compatibility and whether the effects of response mapping primarily reflect compatibility gains or incompatibility losses.

Different methods of parameter estimation are in use for diffusion-model analyses (Ratcliff & Tuerlinckx, 2002) that have different strengths and weaknesses. So-called chi-square and weighted least squares methods rely on grouped data, whereas ungrouped latencies and responses are the raw materials of the so-called maximum-likelihood method. Because there were relatively few data points per analysis (namely 96), further aggregation of the data into groups was out of the question and the so-called efficiency of the estimation method became the most important criterion. For this reason, the maximum-likelihood method was chosen. Maximum-likelihood estimation, although making the most efficient use of the data, can, however, be relatively sensitive to outliers and contaminants in the reaction time distributions (Ratcliff & Tuerlinckx, 2002). Trials with response times below 100 ms were therefore discarded, as were trials with latencies that were outliers in the reaction time distribution of the analyzed condition according to Tukey's outlier criterion – that is, latencies that were above the third quartile plus 1.5 times the individual's interquartile range in that condition or below the first quartile minus 1.5 times the interquartile range (Weisstein, 1999).[1] This led to the exclusion of 6.04% of the responses.[2] Four participants made no error in one or more of the analyzed conditions, and their data were also excluded because fitting the diffusion model requires both false and correct responses for stable and nondegenerate parameter estimates.

Results

Figure 2 shows means and 95% confidence intervals for response latencies and error rates in the seven analyzed conditions. As can be seen, there was the usual pronounced IAT effect in the latency domain (M = 167 ms, SD = 117 ms), $t(15) = 5.73$, $p < .01$, and in the error domain (M = 5.24%, SD = 6.00%), $t(15) = 3.50$, $p < .01$.

In Figure 3, means and 95% confidence intervals are shown for nondecision components t_0, speed-accuracy settings a, and the mean drift rates v, as determined through the diffusion-model analysis. In Table 2, results are shown for the variability parameters. It can be seen in Figure 3 that nondecision components t_0 and speed-accuracy parameters a for both tasks are increased in the incompatible block relative to the compatible block, whereas drift rates v are decreased. The differences between compatible block and single-task practice blocks are less pronounced.

Focusing on the critical compatible and incompatible phase, an analysis of variance was computed for each process component with compatibility (compatible vs. incompatible) and task (concept task vs. attribute task) as factors. For the nondecision component t_0, the incompatible phase was associated with significantly larger values of t_0 than the compatible phase, $F(1, 15) = 8.92$, $p < .01$, and this compatibility effect was the only significant effect to emerge: all other Fs$(1, 15) \leq 1.36$, ps $\geq .26$. For the speed-accuracy parameter a, main effects of compatibility, $F(1, 15) = 21.94$, $p < .01$, and task, $F(1, 15) = 6.51$, $p = .02$, were modified by a significant interaction of both factors, $F(1, 15) = 5.13$, $p = .04$. As can be seen in the middle panel of Figure 3, speed-accuracy settings were more conservative in the incompatible phase than in the compatible phase, an effect that was more pronounced for the attribute task than for the concept task. In separate t tests, the effect was in fact significant only for the attribute task, $t(15) = 4.39$, $p < .01$, but not for the concept task, $t(15) = 1.58$, $p = .13$. Finally, for the drift rates v, a main effect of compatibility, $F(1, 15) = 19.92$, $p < .01$, was modified by a significant interaction with task, $F(1, 15) = 6.27$, $p = .02$, indicating a particularly strong decrease of drift rate from compatible to incompatible phase for the concept task. In separate t tests, the decrease was significant for concept task and attribute task, $t(15) = 4.43$, $p < .01$, and $t(15) = 2.85$, $p = .01$, respectively.

The same analyses were conducted for the variability parameters (see Table 2). The variability parameters have comparatively little effect on overall mean latencies and error rates; they are more important for the shapes of the reaction time distributions for correct and false responses. For this reason, they are of secondary interest in accounting for the IAT effect. The effects on the range in nondecision component s_t descriptively mirrored those on the nondecision component t_0 itself, but a particularly strong increase from compatible to incompatible phase for the concept task gave rise to a significant interaction of both factors, $F(1, 15) = 5.50$, $p = .03$, accompanied by main effects of compatibility, $F(1, 15) = 11.90$, $p < .01$, and task, $F(1, 15) = 7.89$, $p < .01$. The range in starting point s_z was larger for the concept task than for the attribute task, $F(1, 15) = 13.53$, $p < .01$, and this main effect of task was the only significant effect to emerge in this analysis (all other Fs < 1), possibly reflecting more polarized response biases for the concept task than for the attribute task. There were no significant effects in the variability η in drift rates: largest $F(1, 15) = 2.98$, smallest $p = .11$.

Model fit was evaluated on three levels. First, we tested the model's ability to reproduce each participant's mean latency and error rates for each of the seven analyzed conditions. Observed mean latencies and error rates correlated almost perfectly with the predicted latencies (r = .99) and error rates (r = .99) across participants and conditions (n = 112). The mean absolute deviations between model predictions and observed values for mean latencies and error rates were 4 ms and 0.6%, respectively. Virtually the same values were obtained when the analyses were restricted to the conditions from the critical compatible and incompatible phase.

Thus, the model reproduced the ingredients of the IAT effect, mean latencies and error rates, satisfactorily for each participant and condition. To assess the degree to which the entire joint distribution of reaction times and (correct vs. false) responses was reproduced, we computed a formal chi-square test of goodness of fit for each participant in each analyzed condition, as explained in the Appendix. This yielded a chi-square distributed goodness-of-fit value for each person and condition with either four or five degrees of freedom, depending on the number of outlier latencies excluded per person and condition. The average chi-square value was 8.48, and the average associated p value was .24. In all, there were 112 goodness-of-fit tests, of which 13 indicated significant deviations between predicted and observed joint distribution of reaction times and responses at the 1% level of significance and 28 at the 5% level.

Finally, for each study, we ran an analysis on data from all participants taken together in which the data from participants without errors and almost all responses were included (only responses with latencies below 100 ms or above 5,000 ms were excluded). Because only few outliers were excluded, aggregate data were fit by minimizing the distance between observed and empirical cumulative distribution functions (i.e., the Kolmogorov-Smirnov distance; see Voss et al., 2004), an algorithm that is less vulnerable to distortions by outliers than the maximum likelihood (Voss & Voss, 2006). For this purpose, responses and reaction times are jointly coded in one variable as follows: For each response, the value of the variable is the reaction time of the response, but for false responses, it is additionally multiplied by –1. Thus, negative values are reaction times of errors and positive values reaction times of correct responses. Figure 4 (top) shows the observed and predicted cumulative distribution function of this variable, separately for compatible and incompatible phase. The portion of the distribution function to the left of the y-axis thereby depicts the distribution of error latencies and the portion to the right of the y-axis the distribution of correct latencies. The empirical and predicted distribution functions intercept the y-axis at the empirical and predicted error rate, respectively.

The figure reveals an excellent fit at the aggregate level, the maximum distance between observed (empirical) and predicted cumulative distribution functions at any point of the horizontal axis being 1.4% in the probability scale (vertical axis). All in all, as in previous applications (e.g., Ratcliff, Thapar, et al., 2004), the model provides a satisfactory first approximation of the data, but there is room for improvement in the model's capability to describe the shape of the individual reaction time distributions.

Discussion

In Study 1, the overall IAT effect of a flower-insect IAT was decomposed into separate process components. Relative to the compatible phase, (a) speed-accuracy settings were more conservative in the incompatible phase, especially for the attribute task; (b) drift rates were reduced, especially for the concept task; and (c) nondecision components required more time. The random-walk account by Brendl et al. (2001) predicts more conservative speed-accuracy settings for both tasks and a reduction of drift rates in the concept task. It cannot account for the reduction in drift rate that was observed for the attribute task, and it is silent with respect to the compatibility effect on the nondecision component. We return to these issues in the General Discussion.

The diffusion-model analyses were broadly consistent with some of Brendl et al.'s (2001) major assumptions. In particular, according to the present analysis, one component of the IAT effect is indeed due to the fact that more conservative speed-accuracy settings tend to be adopted in the incompatible phase than in the compatible phase. Following Brendl et al., this might open the door to unwanted sources of variance in the IAT related to the many and varied factors that influence speed-accuracy trade-offs, such as age, prevention versus promotion focus, instructions, strategies, processing styles, and so forth. However, at this stage, all we can say is that a compatibility effect on speed-accuracy settings contributes to the mean IAT effect. It is not automatically implied that there are substantial and systematic interindividual differences in the size of this compatibility effect that affect the rank order of individuals in the IAT effect itself. Only if there are such interindividual differences do effects on speed-accuracy settings seriously threaten the validity of the IAT. One purpose of the studies reported below was to investigate the nature and extent of interindividual differences in the different process components that contribute to the IAT.

As suggested by the analyses of variance, all three process components, nondecision components, speed-accuracy settings, and drift rates were responsible for sizeable contributions in the IAT effect. From an applied view, the absolute size of the IAT effect is often less important than differences between groups or individuals in the IAT effect. In the following studies, we adopted a correlational approach using external criteria to assess the extent to which the different process components of the IAT effect reflect construct-specific variance (Studies 2 and 3) relating to the construct to be measured and method variance (Study 3).

Study 2: Construct-Specific Variance

In Study 2, we addressed the question of how construct-specific variance in the IAT maps on the process components. For this purpose, a political attitudes IAT was administered along with an external marker of construct-specific variance, explicit ratings of political attitude. The IAT and the attitude ratings contrasted a red political attitude and a black political attitude. In Germany, the red political attitude is associated with the left political spectrum, including issues of social equality, preservation of the environment, and openness to other cultures. The black political attitude is associated with the right political spectrum, including issues of patriotism, authority, and conservative values.

Method

Participants. Sixty University of Freiburg students (26 women, 34 men) with different majors participated. Their mean age was 24 years, with a range from 19 to 40 years. Participants received a gratification of 6.00 euros (U.S. $8.07) for participating.

Explicit measures. Explicit measures were (a) a 10-point Likert scale for the personal political standpoint on a red versus black dimension, (b) separate 10-point thermometer ratings for the red political standpoint and the black political standpoint, and (c) 10-point Likert scales for the valence of each of the concept stimuli used in the political attitudes IAT. The last two sets of ratings were averaged per person, with reverse scoring for ratings pertaining to black political attitudes (and concepts). All three measures were then z-transformed, and the average of the three z scores was the explicit measure of political attitude (Cronbach's α = .90). Participants were also asked to rate their interest in political issues and events and whom they would vote for if elections were held next Sunday.

Political attitudes IAT. We, along with 17 students who did not participate in Study 2, generated a pool of 47 concept stimuli representing either the black or the red political standpoint. The candidate concept stimuli were then rated with respect to typicality for the red versus black political view and with respect to valence by 57 students who did not take part in Study 2. Each of these students also rated the valence of 59 candidate attribute stimuli (29 positive, 30 negative) and provided a rating of his or her political attitude on a red versus black dimension. On the basis of these data, six concept stimuli were selected for the red political attitude (e.g., socialism, multicultural) and six were selected for the black political attitude (e.g., conservative, fatherland), along with six attribute stimuli for the positive attribute category (e.g., joy, love) and six for the negative attribute category (e.g., emergency, poison). Concept stimuli were chosen if they were rated as typical of the respective political attitude and if their valence ratings correlated highly with the overall rating of political attitude. Attribute stimuli were chosen if they were among those rated most positive or most negative. For each concept and attribute category, an additional stimulus was sampled and reserved for use in warm-up trials.

The political attitudes IAT comprised seven blocks of either 24 or 48 trials. In Table 3, the specifics of each block are summarized. Each block was preceded by additional warm-up trials using the stimuli that were reserved for the warm-up trials. There was one warm-up trial for each stimulus category that appeared in the block. Single-task blocks were thus preceded by two warm-up trials; blocks combining both tasks were preceded by four warm-up trials. The same presentation parameters were used as in Study 1, with the exception that the fixation cross preceding each trial was replaced by a blank intertrial interval of 500 ms, in keeping with standard IAT procedures. The order in which the critical combined phases were presented was balanced across participants.

Procedure. Participants first worked through the explicit rating measures, followed by a filler task that lasted about 5 min. Then the IAT was administered, followed by a biographical questionnaire.

Model analyses. Diffusion models were fitted to the data from the critical phases of the political attitudes IAT, separately for each phase and participant. Because of the relatively small number of trials, the analyses did not distinguish between attribute and concept task. Each model analysis was thus based on 72 trials. Using the same criteria as in Study 1, outlier latencies were, however, removed prior to analysis, thereby excluding 5.81% of the responses. In addition, data from 9 participants were excluded because they made no errors in one of the analyzed phases (cf. the Method section of Study 1).

Results and Discussion

The mean IAT effect was 269 ms (SD = 227 ms), $t(50) = 8.45$, $p < .01$, indicating a general preference for the red part of the political spectrum in the sampled population. The mean IAT effect for the error data was 2.96% (SD = 4.27%), $t(50) = 4.95$, $p < .01$. Table 4 shows the mean parameter values of the diffusion model for the two critical phases of the IAT along with the results of t tests for the differences between the two phases. As can be seen, there were significant compatibility effects in nondecision component t_0, speed-accuracy parameter a, and drift rate v. There was little effect on the variabilities in the process components.

The attitude ratings correlated to a moderate degree with the IAT, both in terms of the conventional latency measure (r = .42, p < .01) and in terms of the so-called D_2 measure[3] (Greenwald, Nosek, & Banaji, 2003; r = .64, p < .01). We computed separate compatibility effects IAT_t, IAT_a, and IAT_v for the nondecision component t_0, speed-accuracy parameter a, and mean drift rate v, respectively.[4] The correlations between the attitude ratings and these process component measures IAT_t, IAT_a, and IAT_v were, in order, r = .15 (p = .29), r = .19 (p = .18), and r = .51 (p = .01), indicating that construct-specific variance is captured in mean drift rate v.

Figure 5 shows the results of a regression analysis regressing the attitude ratings on the three process component measures IAT_t, IAT_a, and IAT_v. Whereas IAT_v predicted the explicit measure significantly, neither IAT_t nor IAT_a was responsible for significant contributions to the regression equation.

Note that IAT_v is not a mediator (Judd & Kenny, 1981) of the relationship between political attitudes IAT and attitude ratings. In particular, the variation in the political attitudes IAT effect is not the cause of variation in IAT_a, IAT_v, or IAT_t. Rather, IAT_a, IAT_v, and IAT_t contribute (in nonlinear fashion) to the overall political attitudes IAT effect. For these and other statistical reasons (such as correlated errors between the overall IAT effect and its process components), it is not meaningful to conduct mediational analyses.

Model fit was again evaluated on three levels. First, we tested the model’s ability to reproduce each participant’s mean latency and error rates for the analyzed (compatible and incompatible) blocks. Observed mean latencies and error rates correlated almost perfectly with the predicted latencies (r = 1.00) and error rates (r = .99) across participants and conditions (n = 102). The mean absolute deviations between model predictions and observed values were 4 ms and 0.4% for mean latencies and error rates, respectively.

Thus, the model again reproduced the ingredients of the IAT effect, mean latencies and error rates, satisfactorily for each participant and compatibility condition. A chi-square distributed goodness-of-fit value was again computed for each person and condition. The average chi-square value was 5.89, and the average associated p value was .26. In all, there were 102 goodness-of-fit tests, of which 7 indicated significant deviations between predicted and observed joint distribution of reaction times and responses at the 1% level of significance and 26 at the 5% level.

Finally, the fit at the aggregate level, including participants without errors and almost all latencies (see Study 1), is shown in Figure 4 (middle). The empirical and predicted cumulative distribution functions of the joint distribution of responses and latencies agreed satisfactorily, the maximum distance being 2.9% on the probability scale. Goodness of fit was thus at a similar level as in Study 1.

The results were straightforward. Construct-specific variance was mapped on the compatibility effect IAT_v in mean drift rate v, whereas neither IAT_a nor IAT_t covaried with explicit ratings of the target construct, that is, with political attitude.

Study 3: Method Variance and Construct-Specific Variance

Method variance in the IAT is indicated by correlations between IATs based on entirely different attributes and concepts for which there is no reason to expect such correlations on a priori grounds. For example, Mierke and Klauer (2003) realized what they called a geometry IAT in which simple geometrical objects (rectangles, triangles, and circles) were used as stimuli. Participants were asked to discriminate red objects from blue objects (concept categories) and small objects from large objects (attribute categories). An association between concepts and attributes was built into the stimuli by means of a contingency so that, for example, red forms were always small and blue forms were always large.

Mierke and Klauer (2003) found that the geometry IAT correlated with a flower-insect IAT and (the absolute size of) the IAT effect in an extraversion IAT, with correlations ranging between .30 and .40. Back et al. (2005) developed an alternative control IAT to operationalize method variance, which they called "task-switch ability" (TSA) IAT. In the TSA IAT, participants have to discriminate letter stimuli (e.g., C, N) from number stimuli (e.g., 4, 7) and words (e.g., shirt, table) from calculations (e.g., 8 – 5 = 3, 2 + 6 = 8). Each concept is associated with one of the attribute categories: Letters are associated with words and numbers with calculations. Using the TSA IAT in combination with an anxiety IAT, Back et al. reported correlations of similar magnitude as those found by Mierke and Klauer between these IATs in three studies. Similarly, McFarland and Crouch (2002) found significant correlations between two control IATs and a flower-insect IAT.

In Study 3, the political attitudes IAT was administered along with markers of construct-specific variance and method variance. Explicit ratings of political attitude were again used as markers of construct-specific variance; two control IATs, geometry IAT and TSA IAT, served as markers of method variance. Because method variance affects the size but not the direction of IAT effects (Mierke & Klauer, 2003), the absolute sizes of the political attitudes IAT effects and their process components entered the analyses involving the control IATs.

Method

Participants. Thirty-two University of Freiburg students (17 women, 14 men; gender information was missing for 1 person) with different majors participated. The mean age was 26.1 years, with a range from 19 to 45 years. Participants received a gratification of 7.00 euros (U.S. $9.41) for participating.

Measures. We used the same explicit measures for political attitudes as in Study 2. In addition, three IATs were administered, the political attitudes IAT already used in Study 2 along with two control IATs, geometry IAT and TSA IAT.

The TSA IAT material consisted of six simple words (shirt, phone, spoon, wall, chair, and pen) and six simple computations (4 + 2 = 6, 1 + 3 = 4, 7 – 4 = 3, 2 + 2 = 4, 4 + 5 = 9, and 8 – 6 = 2) as attribute stimuli and six letters (N, K, B, R, E, and S) and six numbers (7, 3, 8, 5, 2, and 6) as concept stimuli, plus an additional stimulus from each category reserved for the warm-up trials. The geometry IAT presented circles, triangles, and squares in six different sizes (with diameters of approximately 15 mm, 20 mm, and 25 mm for the small size and 35 mm, 40 mm, and 45 mm for the large size) and with outlines colored in one of six colors (red, blue, brown, yellow, green, and purple). The concept task was to discriminate red from blue shapes; the attribute task was to discriminate small from large sizes. Either all red stimuli were small and all blue stimuli were large or all red stimuli were large and all blue stimuli were small, as balanced across participants. The format and presentation parameters were otherwise the same for the control IATs as for the political attitudes IAT.

Procedure. Participants first worked through the political attitudes IAT with combined phase mapping the red political attitude and positive attribute stimuli on the same response key (see Table 3). This was followed by the TSA IAT and the geometry IAT. The order in which these two were administered, compatibility order in the TSA IAT, compatibility order in the geometry IAT, and the nature of the contingency realized in the geometry IAT (red = small vs. red = large) were balanced across participants orthogonally to each other. Finally, the explicit measures were obtained.

Model analyses. Diffusion models were fitted to the data from the critical phases of the political attitudes IAT, separately for each phase and participant as in Study 2. Each model analysis was thus based on 72 trials. Using the same criteria as in the previous studies, outlier latencies were, however, removed prior to analysis, thereby excluding 6.12% of the responses. In addition, data from 5 participants were excluded because they made no errors in one of the analyzed phases.

Results and Discussion

The mean political attitudes IAT effect was 282 ms (SD = 325 ms), $t(26) = 4.51$, $p < .01$, indicating a general preference for the red part of the political spectrum in the sampled population. The IAT effect for the error data was 4.80% (SD = 7.09%), $t(26) = 3.52$, $p < .01$.

Table 4 shows the mean parameter values of the diffusion model for the two critical phases of the political attitudes IAT along with the results of t tests for the differences between the two phases. As can be seen, there were significant compatibility effects in nondecision component t_0, speed-accuracy parameter a, drift rate v, and variability s_t in nondecision component. There was little effect on the variabilities in the other process components.

Construct-specific variance. Attitude ratings again correlated to a moderate degree with the political attitudes IAT, both in terms of the latency measure ($r = .46$, $p = .02$) and in terms of the D_2 measure ($r = .50$, $p < .01$). Figure 6 (top) shows the results of a regression analysis regressing the attitude ratings on the three process components. Whereas IAT_v predicted the explicit measure significantly, neither IAT_t nor IAT_a were responsible for significant contributions to the regression equation. Table 5 shows the correlations between process components and attitude ratings. As can be seen, IAT_v correlated significantly with the explicit measure ($r = .51$, $p < .01$).

Method variance. The correlations between the (absolute size of the) political attitudes IAT on the one hand and the geometry IAT and TSA IAT on the other hand were $r = .32$, $p = .11$, and $r = .38$, $p < .05$, respectively. The correlation between the two control IATs was somewhat higher ($r = .50$, $p < .01$), consistent with the assumption that the political attitudes IAT contains systematic construct-specific variance related to political attitudes over and above method variance. As in Mierke and Klauer (2003), correlations between the control IATs and the political attitudes IAT were eliminated when the D_2 measure was used for the latter (largest $r = .15$, smallest $p = .47$).

How does method variance map onto the process components of the political attitudes IAT? Figure 6 (bottom) shows regression analyses in which the latency measures of the two control IATs are regressed on the process components. As can be seen, component IAT_a uniquely predicted the markers of method variance. In Table 5, the correlations between process components and control IATs are shown. IAT_a correlated significantly with both control IATs, whereas IAT_v and IAT_t played little role in accounting for method variance.

Model fit. Observed mean latencies and error rates for compatible and incompatible blocks correlated almost perfectly with the predicted latencies (r = 1.00) and error rates (r = 1.00) across participants and conditions (n = 54). The mean absolute deviations between model predictions and observed values were 5 ms and 0.4% for mean latencies and error rates, respectively. The average chi-square value for the fit of the individuals' joint distribution of reaction times and responses was 4.77, and the average associated p value was .33. In all, there were 54 goodness-of-fit tests, of which 2 indicated significant deviations between predicted and observed joint distribution at the 1% level of significance and 4 at the 5% level. The fit at the aggregate level, including participants without errors and almost all latencies (see Study 1), is shown in Figure 4 (bottom). The empirical and predicted cumulative distribution functions of the joint distribution of responses and latencies again agreed satisfactorily, the maximum distance being 1.9% on the probability scale.

Discussion. The results are straightforward. As in Study 2, construct-specific variance was focused on the compatibility effect on drift rates, IAT_v. Method variance was primarily mapped on the compatibility effect on speed-accuracy settings, IAT_a. There was little evidence for a pronounced role of IAT_v or IAT_t in accounting for method variance.

General Discussion

We presented a diffusion-model analysis of the IAT. In Study 1, a longer than usual IAT was used to model attribute task and concept task separately. The analysis decomposed the IAT effect into three dissociable compatibility effects, IAT_t, IAT_a, and IAT_v, in nondecision components, speed-accuracy settings, and drift rates, respectively. The IATs used in Studies 2 and 3 followed standard procedures, and the results confirmed that IAT_t, IAT_a, and IAT_v contributed to the mean IAT effects significantly. What do the findings imply for extant process theories of the IAT?

Process Components and Process Theories of the IAT

Random-walk model. As predicted by Brendl et al. (2001), drift rate decreased from compatible block to incompatible block for the concept stimuli (Study 1), whereas speed-accuracy parameter a increased. As also surmised by Brendl et al., construct-specific variance was mapped on the compatibility effect on drift rate in Studies 2 and 3.

As explained in the introduction, Brendl et al. (2001) did not expect a compatibility effect on drift rate for the attribute stimuli, yet drift rate also decreased from compatible block to incompatible block for the attribute stimuli (Study 1). Furthermore, the random-walk account does not explain the compatibility effect that consistently emerged for the nondecision component t_0.

Taken together, the results are broadly consistent with the random-walk account, but they necessitate modifications of that account.

The account by task-set switching. Mierke and Klauer (2001, 2003; see also Klauer & Mierke, 2005) proposed an account of the IAT by task-set switching that shares some elements with Brendl et al.'s (2001) random-walk account. Like the random-walk account, the account by task-set switching is grounded in the idea that participants experience conflict in selecting the appropriate response for concept stimuli in the incompatible phase. The conflict arises because the attitude associations of concept stimuli prime the wrong response in the incompatible phase. As for the random-walk account, this leads one to expect a compatibility effect on drift rate for concept stimuli in proportion to the strength of attitude associations of concept stimuli. The experience of greater conflict in the incompatible phase than in the compatible phase is also consistent with the adoption of more conservative speed-accuracy settings for the same reasons as in the random-walk account.

In the account by task-set switching, an additional idea is that participants capitalize on response synergy in the compatible phase. In the compatible phase, participants can respond fast and accurately, even if they do not switch from the attribute task to the concept task and instead evaluate both kinds of stimuli, attribute stimuli and concept stimuli, on the attribute dimension. For example, in a flower-insect IAT, responding to *tulip* on the basis of its category membership as flower (concept task) or on the basis of its positive evaluation (attribute task) leads to the same response in the compatible phase. This means that participants do not have to perform each and every switch from attribute task to concept task to perform fast and accurately. In contrast, accurate responding in the incompatible phase requires performing each task switch. Task switches are, however, associated with additional performance costs.

Task-switch costs comprise different components, costs of active preparation for upcoming tasks and passive task-switch costs due to so-called task-set inertia. Task switching requires the activation of appropriate task sets and the suppression of competing and interfering task sets. Task-set inertia (Allport, Styles, & Hsieh, 1994) means that task sets, once activated, 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 apply (Mayr & Keele, 2000).

Preparatory costs of task-set switching, such as retrieving from memory the response mapping that is appropriate for the upcoming trial, are likely to add to the nondecision component in the incompatible phase, where task switching is necessary for accurate responding. According to the account by task-set switching, preparatory costs are therefore at least partly responsible for the compatibility effect on nondecision components.

Task-set inertia, in contrast, leads to the expectation that drift rates should also be reduced for attribute stimuli in the incompatible phase. In the incompatible phase, a correct response to a concept stimulus requires suppressing the attribute task set. Inhibiting the attribute task set has aftereffects (Klauer & Mierke, 2005), so that when an attribute stimulus is presented next, it is more difficult than in the compatible phase to gather attribute information. In fact, Klauer and Mierke (2005) found that the inhibition of attribute information that is required in the incompatible phase is strong enough to survive blocks of 24 trials and a change in response modality (from keypresses to vocal responses). On the basis of magnetic resonance imaging data, Chee, Sriram, Soon, and Lee (2000) also suggested that inhibitory processes play a role in IAT responses. Failure to inhibit distracting associations is finally a key element in the quad model of the IAT (Conrey et al., 2005) that is considered below. Because the strength of inhibition depends on the strength of the interfering associations of concept stimuli (Allport et al., 1994), the size of this cost component is again proportional to the strength of the attitude associations of concept stimuli.

Taken together, the present results agree well with the account by task-set switching. The compatibility effects on drift rate for concept stimuli and on speed-accuracy settings are explained in a similar manner as by the random-walk model. In addition, the account by task-set switching offers explanations for the compatibility effects on drift rate for attribute stimuli and on nondecision components.

The quad model. What are the relations to another componential account of the IAT, the quad model by Conrey et al. (2005)? The quad model concurs with the present analysis in assuming that different components contribute to the IAT. On the basis of the error data, it disentangles four components: the automatic activation of an association (AC), the ability to determine a correct response (D), the success at inhibiting automatically activated associations (OB), and the influence of a general response bias (G). There is no simple relationship between these parameters and the diffusion-model parameters, but a few conceptual parallels can be drawn. Mean drift rates, averaged over compatible and incompatible phase, most directly correspond to parameter D; the effects of activated attitude associations (AC) and failure to inhibit them (OB) are reflected in the drop of drift rates from compatible to incompatible phase, that is, in IAT_v. An advantage of the quad model is thus that it disentangles two potentially dissociable aspects, AC and OB, that are conflated in IAT_v. Response bias (G) conceptually corresponds to parameter z of the diffusion model. It is difficult to see how differential speed-accuracy settings might map on the quad model's parameters, given that the quad model is a model of only the accuracy data of IATs and does not deal with the latency data.

The Nature of Interindividual Differences in the IAT Effect

Studies 2 and 3 used external markers to gain more insight into the meaning of the process components of the IAT effect. These studies revealed that IAT_a and IAT_v are dissociable components of the IAT effect that load on different external variables. IAT_v, but not IAT_a, predicted the explicit attitude measure, the marker of construct-specific variance (Studies 2 and 3). Conversely, IAT_a, but not IAT_v, predicted the geometry IAT and the TSA IAT, the markers of method variance (Study 3).

This pattern of relationships invites a number of interesting conclusions. Method variance in Study 3 primarily reflects interindividual differences in speed-accuracy trade-offs. It thus relates to a strategic aspect of processing and the larger field of risky versus cautious processing styles. Given that speed-accuracy trade-offs are influenced by many and varied factors (e.g., age, mood, accuracy rewards vs. speed rewards, instructions, promotion vs. prevention focus, personality traits, and so forth), this component of the IAT effect opens the door for a large variety of situational, strategic, and trait-related influences on the IAT effect that are conceptually independent of the construct-specific variance and thereby systematic contaminants in the effect.

For example, the stability of IAT scores assessed in terms of test-retest reliability is often lower than the internal consistency as assessed by Cronbach's alpha or split-half reliability (Schmukle & Egloff, 2004), indicating that IAT scores capture both stable variance and systematic occasion-specific variance. Given the many and sometimes transient factors that affect speed-accuracy settings, one contribution to occasion-specific variance may be given by the speed-accuracy component of method variance. Conversely, the speed-accuracy component might have been especially pronounced in the present studies because the different IATs were performed in immediate succession, facilitating a direct transfer of speed-accuracy settings from one IAT to the next.

That method variance is taken up in the speed-accuracy parameters meshes well with diffusion-model analyses from the aging literature (e.g., Ratcliff, Thapar, et al., 2004), in which age-related individual differences in speeded reaction time tasks were mapped on differences in speed-accuracy settings. This parallel begins to provide generality across paradigms and domains of investigation.

Note, however, that IAT_a is the difference in speed-accuracy settings between incompatible and compatible phase rather than the mean or overall speed-accuracy setting. Thus, interindividual differences in how individuals choose speed-accuracy settings for the differently difficult IAT phases led to the correlations with the control IATs.

On the basis of the speculations in the literature, including some of our own writing, we initially expected method variance to reflect interindividual differences in cognitive skill. As explained in the introduction, interindividual differences in cognitive skill are captured by the mean drift rates. This led us to assume that method variance would relate to the compatibility effect on drift rates, IAT_v, rather than to IAT_a. There was, however, little evidence in our data for an involvement of IAT_v in accounting for method variance. Nevertheless, we hesitate to abandon the idea of a small residual cognitive skill confound altogether, but the present data do not provide evidence for it.

D_2 Versus IAT_v

From an applied point of view, IAT_v and the diffusion-model analysis offer an alternative to the family of *D* measures recently proposed by Greenwald et al. (2003). IAT_v and D_2 performed roughly similarly in their relationship to external variables:[5] Both were unrelated to the control IATs, the markers of method variance. Conversely, both showed moderately strong correlations with the marker of construct-specific variance, that is, with the attitude ratings in Study 2 (r = .51 and r = .64 for IAT_v and D_2, respectively) and Study 3 (r = .51 and r = .50 for IAT_v and D_2, respectively).

Relative to D_2, we see two advantages of the diffusion-model analysis. First, IAT_v is a principled measure deriving from an established model of decision processes in binary decision tasks. It provides a theory-based means to partial out interindividual differences in threshold setting (IAT_a) and nondecision components (IAT_t), both of which contribute to IAT effects. Second, the diffusion model furthermore provides explicit measures of these additional process components of IAT effects, namely IAT_a and IAT_t. In applications, this allows one to assess the degree to which observed correlations or effects reflect construct-specific variance (IAT_v) or, alternatively, effects on processing styles and strategies (IAT_a) or nondecision components (IAT_t). For example, it would be interesting to assess the degree to which so-called malleability effects in the IAT (e.g., Blair, 2002) reflect changes in implicit attitudes or prejudice (i.e., effects on IAT_v) or, alternatively, changes in risky versus cautious processing style (i.e., effects on IAT_a).

Similarly, it has been suggested that successful faking of IAT effects is driven by strategic delaying of responses in the compatible phase and/or by strategic settings of speed-accuracy trade-offs (Fiedler & Bluemke, 2005). If so, faking might be diagnosed and corrected for through an explicit diffusion-model analysis. Specifically, delaying responses in the compatible phase would lead to elevated parameters t_0 for nondecision components in that phase and strategic settings of speed-accuracy trade-offs would lead to effects on parameter a, but both nondecision components and speed-accuracy settings are automatically partialled out of IAT_v in a diffusion-model analysis. Pronounced misfit of the model could also be diagnostic of faking.

In contrast, there are three technical disadvantages associated with the diffusion-model analysis. First, participants who made no errors in one of the two critical phases of the IAT had to be excluded from the analyses because parameter estimation requires both errors and correct responses for stable and nondegenerate results. Second, the information in the data is spread out among the different parameters of the model rather than - like for the *D* measures - concentrated in one measure. This makes it likely that reliability is lower for the model-based measures than for the *D* measures. For example, the attitude ratings in Study 2 correlated somewhat more strongly with D_2 than with IAT_v, possibly reflecting a larger reliability of D_2 than of IAT_v. Third, but not least, the diffusion-model analysis is costly to compute. Different research groups are, however, developing new algorithms and software, work that is beginning to make diffusion-model analyses much more accessible (Vandekerckhove & Tuerlinckx, 2006; Voss & Voss, in press; Wagenmakers, van der Maas, & Grasman, 2007).

To summarize, IAT effects are contributed to by three dissociable process components. Ease and speed of information accumulation is lower in the incompatible phase than in the compatible phase, more conservative speed-accuracy settings are adopted in the incompatible phase, and nondecision components provide a third contribution to the IAT effect. Method variance was primarily mapped on differential response-threshold settings, suggesting that it reflects differences in processing styles and strategies. In contrast, construct-specific variance was focused on the compatibility effect on ease of information accumulation.

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 & performance XV: Conscious and nonconscious information processing* (pp. 421-452). Hillsdale, NJ: Erlbaum.

Atkinson, A. C. (1970). A method for discrimination between models. *Journal of the Royal Statistical Society*, *32*, 323-345.

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

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242-261.

Bofinger, E. (1973). Goodness-of-fit test using sample quantiles. *Journal of the Royal Statistical Society B*, *35*, 277-284.

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.

Chee, M. W. L., Sriram, N., Soon, C. S., & Lee, K. M. (2000). Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *NeuroReport*, *2*, 135-140.

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

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

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, *27*, 307-316.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

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

Hofmann, W., Gawronski, B., Geschwender, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369-1385.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*, 602-619.

Klauer, K. C. (2002). Model testing and selection, theory of. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 9927-9931). Oxford, England: Pergamon.

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

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

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

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.

Moore, D. S. (1986). Tests of chi-squared type. In R. B. D'Agostino & M. A. Stephens (Eds.), *Goodness-of-fit techniques* (pp. 63-95). New York: Marcel Dekker.

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

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.

Ratcliff, R. (1985). Theoretical interpretations of speed and accuracy of positive and negative responses. *Psychological Review*, *92*, 212-225.

Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*, *95*, 238-255.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, *9*, 278-291.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159-182.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347-356.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, *16*, 323-341.

Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, *65*, 523-535.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*, 438-481.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*, 261-300.

Schmukle, S. C., & Egloff, B. (2004). Does the Implicit Association Test for assessing anxiety measure trait and state variance? *European Journal of Personality*, *18*, 483-494.

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 101-117.

Vandekerckhove, J., & Tuerlinckx, F. (2006). *Fitting the Ratcliff diffusion model to experimental data.* Manuscript submitted for publication.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*, 1206-1220.

Voss, A., & Voss, J. (2006). *A fast numerical algorithm for the estimation of diffusion-model parameters.* Manuscript submitted for publication.

Voss, A., & Voss, J. (in press). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods.*

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An ez-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3-22.

Weisstein, E. W. (1999). Box-and-whisker plot. In *MathWorld - A Wolfram Web resource.* Retrieved, November 10, 2005, from http://mathworld.wolfram.com/Box-and-WhiskerPlot.html

Appendix

An Asymptotically Chi-Squared Distributed Goodness-of-Fit Test

To obtain an asymptotically chi-square distributed goodness-of-fit statistic, we mapped responses and reaction times on a single continuous random variable *X* as follows (Voss et al., 2004): The value of the random variable is the reaction time of the response, but for false responses, it is additionally multiplied by –1. Thus, negative values of *X* are reaction times of errors, and positive values are reaction times of correct responses. Given a one-dimensional real-valued random variable, *X*, standard techniques for assessing goodness of fit can be used (e.g., Klauer, 2002). These rely on partitioning the real line into a finite number of intervals and comparing expected and observed frequencies of observations falling into each interval by means of an appropriate sum of squares. From the many, slightly different quadratic forms that can be computed for such an asymptotically chi-square distributed statistic, we chose one that allowed us to work with a minimum of evaluations of the cumulative distribution function.

The number of intervals, *k,* was chosen, following recommendations derived by Moore (1986), as the largest integer smaller than or equal to $2n^{2/5}$, where *n* is the number of data points in the model analysis. Given *n* continuous real-valued random variables $X_1,\dots X_n$, one modeling each data point, the so-called order statistics $X_1^{'},\dots X_n^{'}$ correspond to the values of the random variables ordered from smallest to largest and thus, $X_1^{'} < \dots X_n^{'}$. Let $0 = \lambda_0 < \lambda_1 < \dots < \lambda_{k-1} < \lambda_k = 1$, and let n_i be the greatest integer less than or equal to $n\,\lambda_i + 1$, $i = 1,\dots k$, and set $n_0 = 0$. The λ_i partition the range of cumulative probabilities into a finite number of intervals. To avoid a subjective element in their choice, intervals of equal size were chosen, so that $\lambda_i = i / k$. Let the parameters of the diffusion model be θ, denote the cumulative distribution function of *X* by $F(\bullet \mid \theta)$, and define the statistic,

$$Y_n^2(\theta) = n\sum_{i=1}^{k}[(F(X_{n_1}^{'} \mid \theta) - F(X_{n_{i-1}}^{'} \mid \theta)) - p_i]^2 / p_i$$

where $p_i = \lambda_i - \lambda_{i-1} = 1 / k$, $i = 1,\dots,k$, $X_0^{'} = -\infty$, and $X_{n+1}^{'} = \infty$. If $\hat{\theta}_n$ is an estimator that minimizes $Y_n^2(\theta)$, then the test statistic $Y_n^2(\hat{\theta}_n)$ is asymptotically distributed as chi-square with $k - q - 1$ degrees of freedom, where *q* is the number of independent parameters (Bofinger, 1973). To compute the goodness-of-fit statistic $Y_n^2(\hat{\theta}_n)$, it was thus necessary to reestimate parameters by minimizing $Y_n^2(\theta)$.

Author note

Karl Christoph Klauer, Andreas Voss, Florian Schmitz, and Sarah Teige-Mocigemba, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany.

Karl Christoph Klauer and Andreas Voss contributed equally to this article. This research was supported by Deutsche Forschungsgemeinschaft Grants Kl 617/29-2 and Kl 614/31-1 to Karl Christoph Klauer.

Correspondence concerning this article should be addressed to Karl Christoph Klauer, Institut für Psychologie, Methodenlehre und Sozialpsychologie, Albert-Ludwigs-Universität Freiburg, D-79085, Freiburg, Germany. E-mail: klauer@psychologie.uni-freiburg.de.

Footnotes

[1] We used a relatively strict criterion for upper and lower tail treatment of reaction times because of the known susceptibility of the maximum-likelihood method to distorting influences of contaminant reaction times generated by other processes than the diffusion model, such as anticipations or delays caused by distraction. In contrast, excluding data from the upper and lower tails of the reaction time distribution is problematic in that it may by itself lead to distortions in parameter estimates. We reanalyzed the data from all studies with a more lenient criterion, excluding reaction times if they were smaller than 100 ms or above the third quartile plus 3 times the interquartile range. This led to the exclusion of only 2% of the trials, but it also provoked a number of degenerate sets of parameter estimates with unrealistically large drift rate, threshold-separation parameter, and/or variability η in drift rate, presumably reflecting the influence of contaminants that were not excluded. For this reason, we felt that the distortions produced by failing to exclude some of the outliers outweighed possible distortions due to the restrictive upper and lower tail treatment for the present data sets.

Ratcliff and Tuerlinckx (2002) proposed an alternative model-based method to deal with possible contaminants in the reaction time distributions. Their model is basically a probabilistic mixture between the diffusion model and a uniform distribution of reaction times. The mixture coefficient is a new parameter. One way to understand this mixture is to say that the estimated mixture coefficient gauges the strength of the evidence for the uniform distribution model versus the diffusion model in the data (Atkinson, 1970). The relatively small sample size means, however, that the statistical evidence for deciding between the two models is relatively weak for the present data sets. As a consequence, the analysis would not work well here because the mixture coefficient would assume values much too large for the uniform distribution in small data sets and because in small data sets, it is undesirable to distribute the given statistical information over even more parameters.

[2] An anonymous reviewer suggested we exclude all trials with latencies below 300 ms. When these were additionally removed (excluding an additional 0.3% of the data) and the model fits repeated, similar results emerged. In particular, the pattern of significant and nonsignificant effects was the same as that reported in the subsequent *Results* section.

[3] Greenwald et al. (2003) have proposed improved scoring procedures for the IAT, among them the so-called D_2 measure. It differs from the conventional latency measure in several aspects, including modified upper and lower tail treatment of reaction times, the use of response latency to correct response (i.e., after an error is made, the latency to the required

correction is added, implying an error penalty), and a standardization similar to that in Cohen's effect size measure *d*.

[4] The compatibility effects are signed so that larger values correspond to larger IAT effects, indicating greater preference for the red political standpoint.

[5] Conceptually, D_2 should be more sensitive to variations in drift rate and less sensitive to speed-accuracy trade-offs than the latency measure, that is, it should be more similar to IAT_v and less similar to IAT_a than the latency measure. In D_2, both slow latencies and errors are penalized, implying that D_2 is less sensitive to speed-accuracy trade-offs than the latency measure: For example, if latencies go down and errors go up in one of the critical IAT phases, then the net effect on D_2 is comparatively small but the decrease in latencies directly affects the conventional latency IAT measure. At the same time, decreases in drift rate increase both latencies and errors. Because both add to D_2 in the same direction, D_2 is more sensitive to variations in drift rate than is the latency measure. When the empirical correlations between latency IAT effect and D_2 on the one hand and the process component measures IAT_a and IAT_v on the other hand were also computed in Studies 2 and 3, it was indeed found that IAT_v correlates higher with D_2 than with the latency measure, whereas IAT_a correlates higher with the latency measure than with D_2 in both studies. It is difficult, however, to interpret correlations between these different indexes. They are all based on data from the same IAT, causing correlated errors. Correlated errors, in turn, have the potential to distort the observed correlations.

Tables

Table 1

Model Parameters

Parameter	Meaning
t_0	Mean duration of nondecision components of processing
z	Mean starting point of evidence accumulation
a	Threshold separation
v	Mean drift rate
s_t	Range of nondecision components
s_z	Range of starting point
η	Inter-trial variability in drift rate (standard deviation)

Table 2

Mean Variability in the Processing Components in Study 1

			s_t		s_z		η	
Block	Phase	Task	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	Practice	Concept	.18	.06	.61	.33	.04	.17
2	Practice	Attribute	.20	.08	.50	.30	.08	.30
3	Compatible	Concept	.19	.07	.72	.26	.00	.00
		Attribute	.19	.09	.40	.38	.06	.22
4	Practice	Concept	.23	.08	.72	.28	.00	.00
5	Incompatible	Concept	.40	.13	.71	.31	.00	.00
		Attribute	.24	.22	.46	.37	.17	.51

Note. Parameter s_z was divided by the maximum range of z (i.e., by parameter α).

Table 3

Blocks of the Political Attitudes Implicit Association Test in Study 2

			Response key	
Block	Trials	Tasks	Left	Right
1	26[a]	Concept	Black political view	Red political view
2	26[a]	Attribute	Negative	Positive
3	28[b]	Combined Compatible	Black political view or Negative	Red political view or Positive
4	52[b]	Combined Compatible	Black political view or Negative	Red political view or Positive
5	26[a]	Concept	Red political view	Black political view
6	28[b]	Combined Incompatible	Red political view or Negative	Black political view or Positive
7	52[b]	Combined Incompatible	Red political view or Negative	Black political view or Positive

Note. [a]The first two trials were warm-up trials. [b]The first four trials were warm-up trials.

Table 4

Parameter Values in Studies 2 and 3 for the Combined Implicit Association Test Phases

Study and parameter	Phase: Red-positive[a] *M*	*SD*	Black-positive[b] *M*	*SD*	t^c	*p*
Study 2						
t_0	0.47	0.05	0.57	0.12	5.81	< .01
a	1.24	0.44	1.64	0.56	5.47	< .01
v	3.69	1.55	2.24	0.93	–5.89	< .01
s_t	0.17	0.10	0.19	0.19	0.67	.51
s_z	0.43	0.31	0.39	0.38	–0.52	.61
η	0.09	0.27	0.10	0.34	0.12	.90
Study 3						
t_0	0.47	0.05	0.56	0.10	4.63	< .01
a	1.21	0.44	1.57	0.57	3.53	< .01
v	3.19	1.43	1.94	1.14	–3.78	< .01
s_t	0.13	0.09	0.21	0.18	2.30	.03
s_z	0.47	0.33	0.39	0.35	–0.99	.33
η	0.04	0.11	0.06	0.23	0.55	.59

Note. Parameter s_z was divided by the maximum range *a*.

[a] Red and positive were mapped onto the same response key. [b] Black and positive were mapped onto the same response key. [c] In Study 2, *df* = 50, and in Study 3, *df* = 26.

Table 5

Correlations Between IAT_t, IAT_a, and IAT_v and Markers of Method Variance (Control IATs) and the Marker of Construct-Specific Variance (Attitude Ratings) in Study 3

Process Component	Geometry IAT	TSA IAT	Attitude Rating
IAT_t	.09	.07	–.03
IAT_a	.52*	.49*	.38
IAT_v	–.03	–.18	.51*

Note. IAT = Implicit Association Test; TSA = task-switch ability; *p < .01.

Figure Captions

Figure 1. The diffusion model. The decision axis is the vertical axis, and the decision time axis is the horizontal axis. The lower threshold is positioned at zero and the upper threshold at a. Information accumulation begins at z with mean drift rate v.

Figure 2. Mean latencies (in seconds; top) and error rates (bottom) in Study 1 as a function of task and block. Error bars show 95% confidence intervals.

Figure 3. Mean parameter values as a function of task and block for parameter t_0 (in seconds; top), a (middle), and v (bottom) in Study 1. Error bars show 95% confidence intervals for mean parameter values.

Figure 4. Graphical display of model fit for the data from Studies 1-3. The picture shows observed (empirical) and predicted cumulative distribution functions for the joint distribution of responses and latencies for Studies 1-3, separately for compatible phase data and incompatible phase data. Negative values on the horizontal axis are latencies of error responses (multiplied by –1), and positive values are latencies of correct responses. The intercept of the cumulative distribution function indicates the percentage of errors. RTs = response times.

Figure 5. Standardized regression coefficients and p values for the regression of the rating-based measure of political attitudes on the compatibility effects in the processing components t_0, a, and v in Study 2. IAT = Implicit Association Test.

Figure 6. Standardized regression coefficients and p values for the regression of the rating-based measure of political attitudes (top) and the latency IAT effects of control IATs (bottom) on the compatibility effects in the processing components t_0, a, and v of the political attitudes IAT in Study 3. IAT = Implicit Association Test; TSA = task-switch ability.

Figures

Figure 1

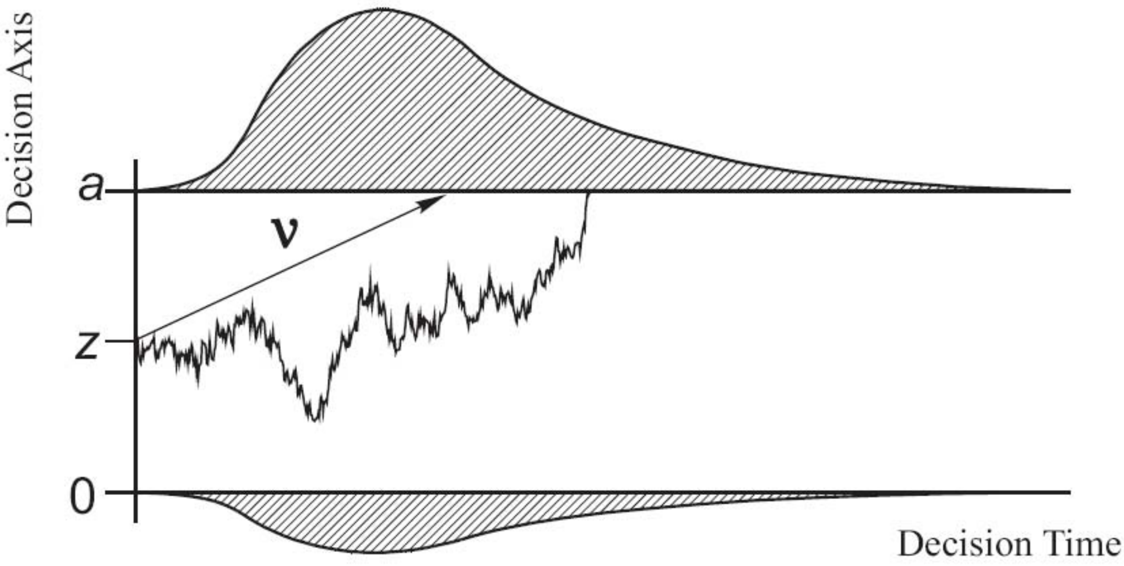
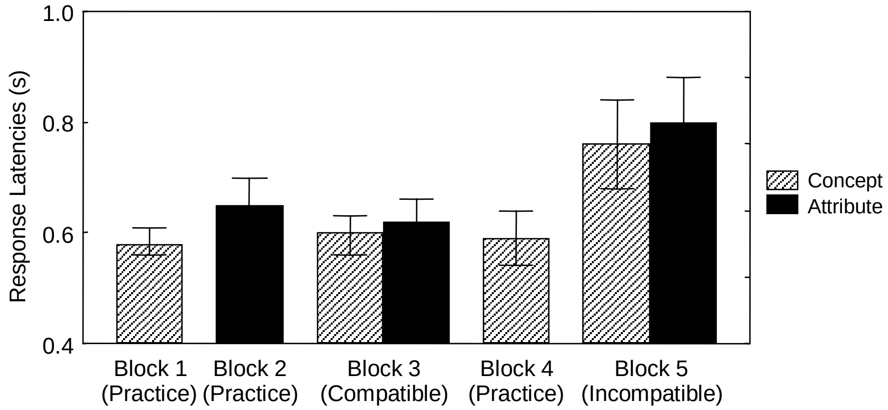
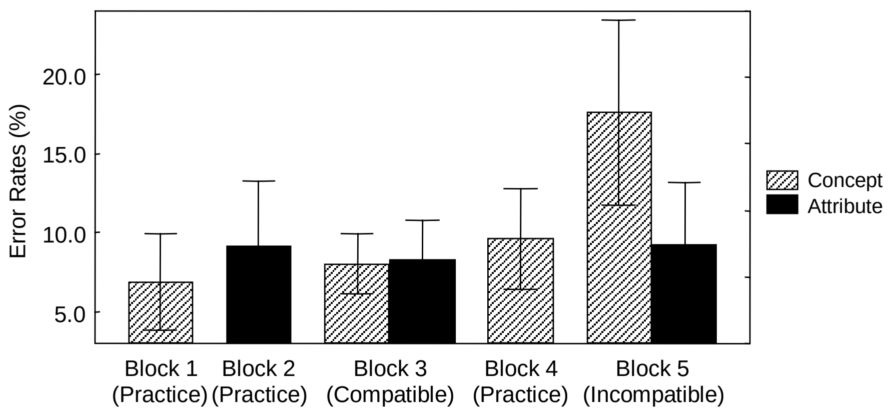


Figure 2



Response Latencies (s)
1.0
0.8
0.6
0.4
Block 1 (Practice)
Block 2 (Practice)
Block 3 (Compatible)
Block 4 (Practice)
Block 5 (Incompatible)
Concept
Attribute



Error Rates (%)
20.0
15.0
10.0
5.0
Block 1 (Practice)
Block 2 (Practice)
Block 3 (Compatible)
Block 4 (Practice)
Block 5 (Incompatible)
Concept
Attribute

Figure 3

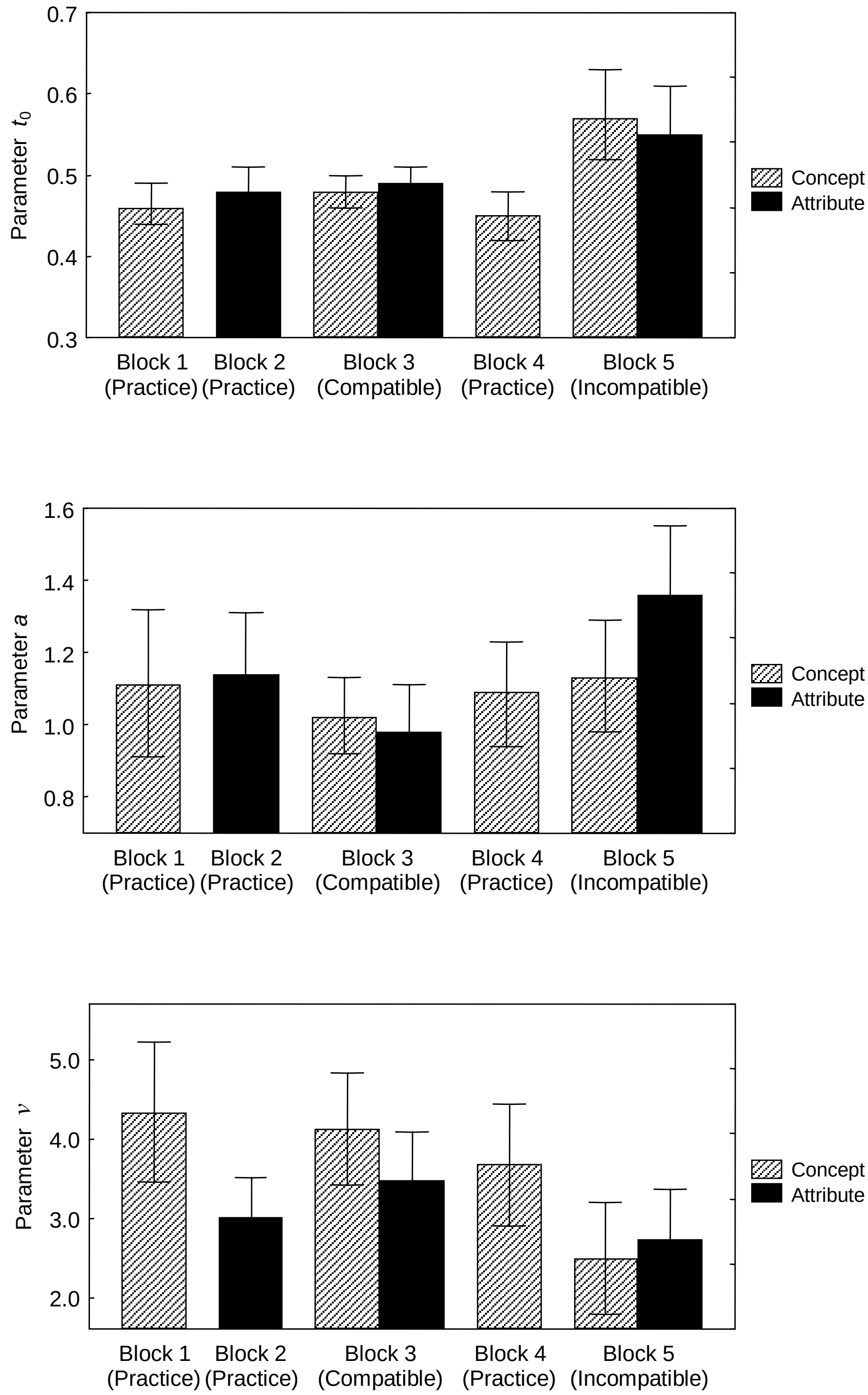

0.7
0.6
0.5
0.4
0.3
Parameter t_0
Concept
Attribute
Block 1 (Practice)
Block 2 (Practice)
Block 3 (Compatible)
Block 4 (Practice)
Block 5 (Incompatible)
1.6
1.4
1.2
1.0
0.8
Parameter a
Concept
Attribute
Block 1 (Practice)
Block 2 (Practice)
Block 3 (Compatible)
Block 4 (Practice)
Block 5 (Incompatible)
5.0
4.0
3.0
2.0
Parameter v
Concept
Attribute
Block 1 (Practice)
Block 2 (Practice)
Block 3 (Compatible)
Block 4 (Practice)
Block 5 (Incompatible)

Figure 4

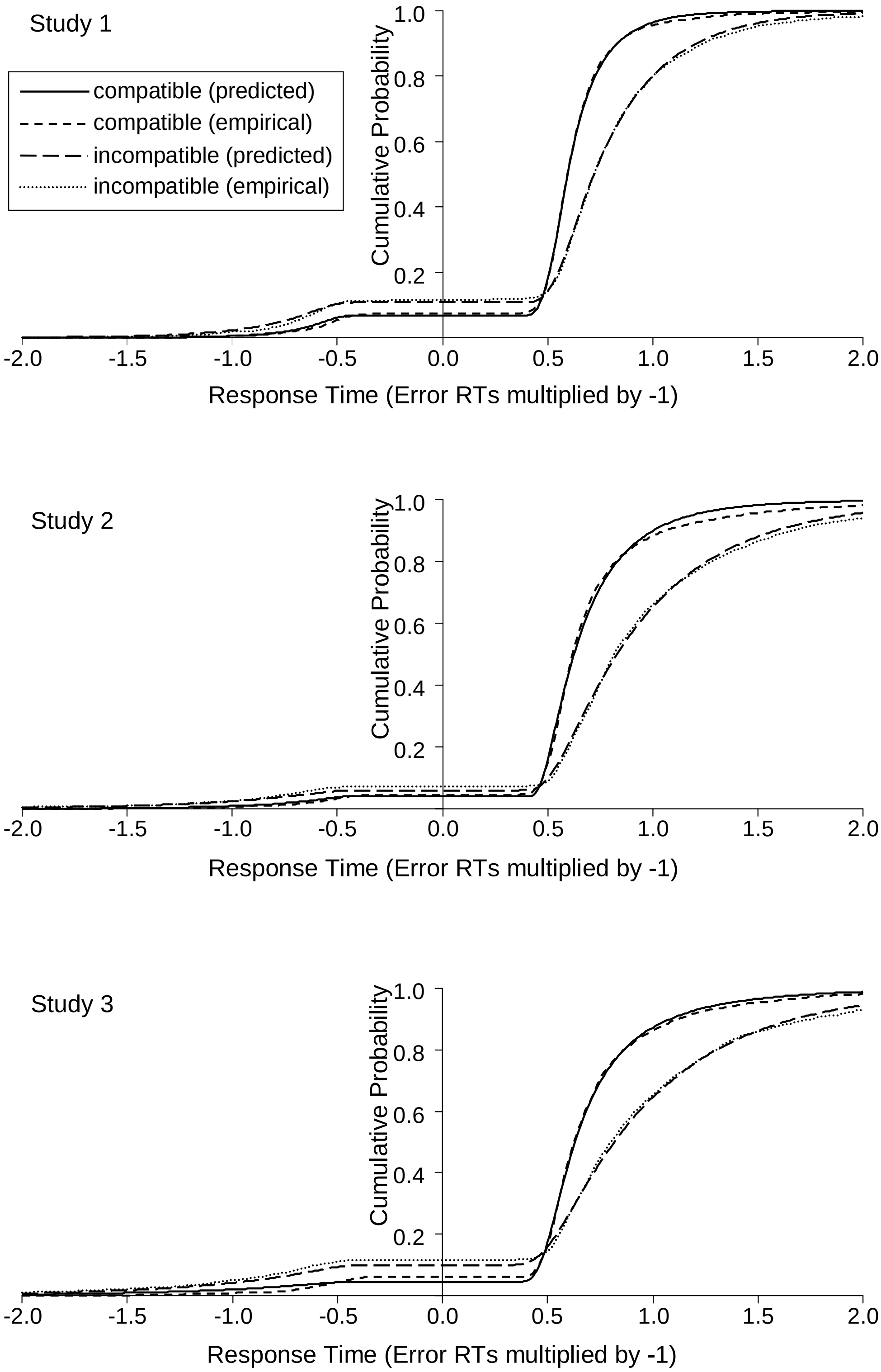

Study 1
compatible (predicted)
compatible (empirical)
incompatible (predicted)
incompatible (empirical)
Cumulative Probability
1.0
0.8
0.6
0.4
0.2
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Response Time (Error RTs multiplied by -1)
Study 2
Cumulative Probability
Response Time (Error RTs multiplied by -1)
Study 3
Cumulative Probability
Response Time (Error RTs multiplied by -1)

Figure 5

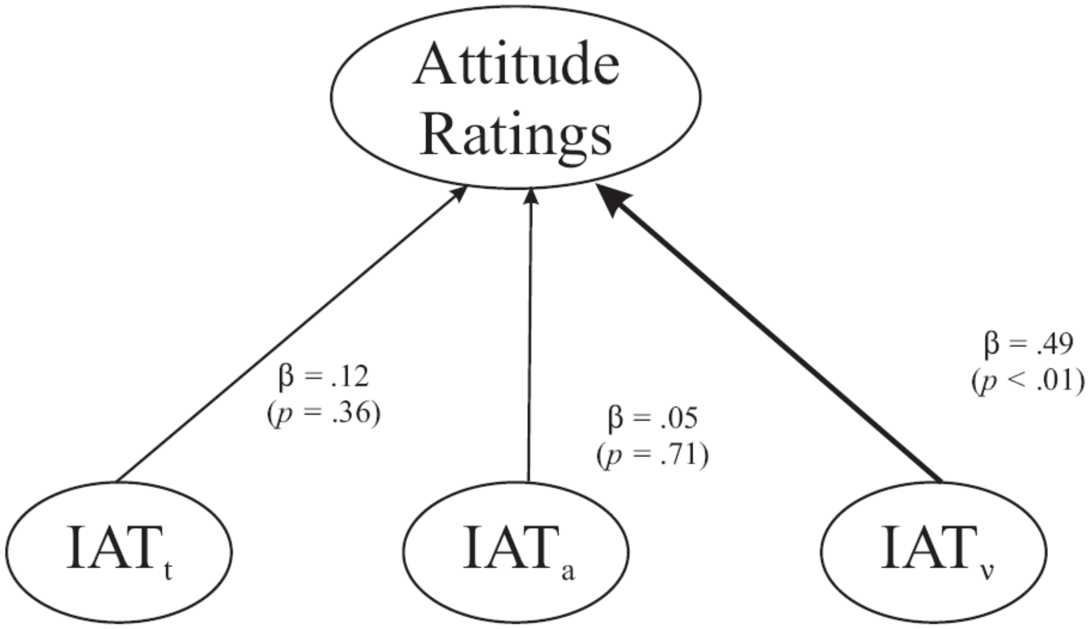
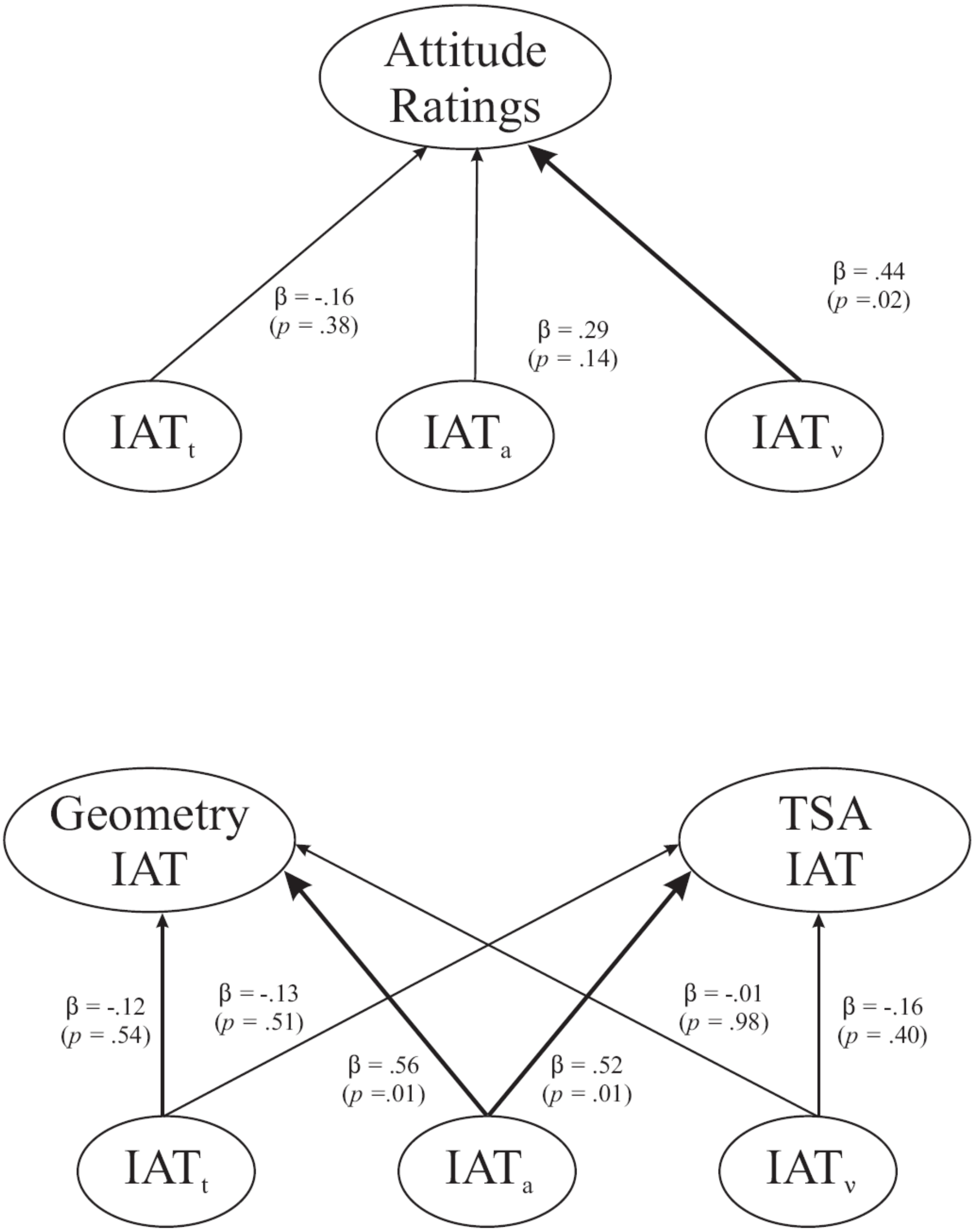


Figure 6


Attitude Ratings
β = -.16 (p = .38)
β = .29 (p = .14)
β = .44 (p =.02)
IAT$_t$
IAT$_a$
IAT$_v$
Geometry IAT
TSA IAT
β = -.12 (p = .54)
β = -.13 (p = .51)
β = .56 (p =.01)
β = .52 (p = .01)
β = -.01 (p = .98)
β = -.16 (p = .40)
IAT$_t$
IAT$_a$
IAT$_v$

Appendix B

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (in press). Minimizing method-specific variance in the IAT: A Single Block IAT. *European Journal of Psychological Assessment*.

Minimizing Method-Specific Variance in the IAT: A Single Block IAT

Sarah Teige-Mocigemba
Albert-Ludwigs-Universität Freiburg

Karl Christoph Klauer
Albert-Ludwigs-Universität Freiburg

Klaus Rothermund
Friedrich-Schiller-Universität Jena

in press, *European Journal of Psychological Assessment*

Abstract

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

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

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

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

Implicit Association Test

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

Method-Specific Variance in the IAT

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

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

Cognitive Skills as Reliable Contamination of the IAT Effect

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

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

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

How to Deal with Method-Specific Variance

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

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

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

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

The main difference between the SB-IAT and the standard IAT is that the SB-IAT compares performance on compatible versus incompatible trials *within the same* (i.e., a single) *block*, whereas the standard IAT compares performance on compatible versus incompatible trials *between two different* (i.e., compatible vs. incompatible) *blocks*. Thus, in the SB-IAT, the response mapping (compatible vs. incompatible) may randomly change from trial to trial and is not consistently blocked anymore. This should impede any kind of recoding strategies, because recoding processes rely on a consistent mapping of categories onto response keys (Strayer & Kramer, 1994).

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

Study 1

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

Method

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

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

SB-IATs. All SB-IATs consisted of eight blocks of either 26 or 52 trials. In Table 1, specifics of each block are summarized for the flower-insect SB-IAT. Note that the TSA SB-IAT and the geometry SB-IAT followed an analogous format. Participants started out practicing the concept and the attribute tasks. First, they were to categorize target stimuli in the upper half of the screen (e.g., left key for insect stimuli and right key for flower stimuli). Then, they were to categorize the same target stimuli in the lower screen half (e.g., left key for flower stimuli and right key for insect stimuli). The tasks of the first two blocks were combined in the third block, in which participants had to correctly assign target stimuli depending on word position. In the fourth block, participants were to categorize attribute stimuli (e.g., left key for negative stimuli and right key for positive stimuli). In four ensuing test blocks, the tasks of the third and the fourth block were combined. For example, in the flower-insect SB-IAT, participants were to discriminate insect and negative from flower and positive (upper screen) or flower and negative from insect and positive (lower screen) depending on word position. All blocks were preceded by additional warm-up trials using stimuli that were reserved for the warm-up trials, one trial and one stimulus per category that appeared in the block. Single-task blocks were thus preceded by two warm-up trials; blocks combining both tasks were preceded by four warm-up trials. Participants used the left key “A” and the right key “5” on a standard computer keyboard to respond.

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

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

Results

Following the conventional scoring procedure (Greenwald et al., 1998), analyses were based on log-transformed response latencies of correct responses, and latencies smaller than 300 ms or greater than 3,000 ms were recoded to 300 ms or 3,000 ms, respectively. Mean response latencies and error rates across the three SB-IATs (M = 764 ms, SD = 136 ms, and M = 8%, SD = 4%, respectively) were comparable to those known from prior IAT research, thereby indicating the feasibility of the SB-IAT task. Note that one participant’s flower-insect SB-IAT data were excluded from all analyses because her mean latency in the flower-insect SB-IAT was an extreme outlier in the distribution of the total sample according to Tukey (1977; mean latency was above the third quartile plus three times the interquartile range).

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

Discussion

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

Importantly, lower internal consistencies of the present SB-IATs are not surprising given that reliability estimates depend on the amount of interindividual variability. The IAT's satisfactory reliability is thought to stem from two systematic, but conflated sources of variance (i.e., construct- and method-specific variance; Mierke & Klauer, 2003), whereas the

SB-IAT's reliability should just stem from one systematic source of variance (i.e., construct-specific variance). Thus, if variability in the construct of interest is rather low, as should be the case for the associations assessed in Study 1, reducing method-specific variance in the SB-IAT should be accompanied by reduced reliability estimates. If, however, participants' variability in the construct of interest is high, reliability estimates for the SB-IAT should also be higher. Consequently, Study 1 might have underestimated the SB-IAT's reliability. Therefore, and because we were also interested in the SB-IAT's validity (here, with regard to implicit-explicit consistency), Study 2 examined the SB-IAT's psychometric properties.

Study 2

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

Method

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

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

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

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

Results

Response latencies were preprocessed as in Study 1. Participants needed a little more time to complete the political attitude SB-IAT (M = 998 ms, SD = 270 ms) as compared to the mean response latencies for the SB-IATs in Study 1, whereas error rates were in the same range (M = 8%, SD = 7%). As expected, the political attitude SB-IAT showed acceptable internal consistency computed as in Study 1, Cronbach's α = .74. The SB-IAT proved to be valid both at the group level and at the correlational level. It discriminated between participants who indicated an intention to vote for the red political spectrum and participants who indicated an intention to vote for the black political spectrum[3], $t(37) = 4.77$, $p < .001$, and correlated to a moderate degree with the attitude rating, $r = .43$, $p < .01$, as has also been found for the highly comparable IAT of Klauer et al. ($r = .42$).[4] Importantly, the SB-IAT's prediction of voting intention (red vs. black political spectrum) was mediated by the attitude rating: If the SB-IAT and the attitude rating entered a logistic regression separately, they both predicted voting intention, $B = 2.10$, $SE = .73$, $p < .01$, and $B = 7.29$, $SE = 3.11$, $p = .02$, respectively. Also, the SB-IAT predicted the attitude rating, $\beta = .43$, $p < .01$. If however, the SB-IAT and the attitude rating entered the logistic regression simultaneously, only the attitude rating ($B = 7.20$, $SE = 3.43$, $p = .04$), but not the SB-IAT ($B = .76$, $SE = .77$, $p = .33$) predicted voting intention.

Discussion

The SB-IAT reliably assessed interindividual differences in political attitudes and proved to be valid, both in terms of discriminating between red- vs. black-voters and in terms of implicit-explicit correlations. However, it did not show incremental validity in predicting voting intention over and above the attitude rating, as has also been shown for the IAT: In socially insensitive domains (e.g., political attitudes), explicit measures outperformed the IAT with regard to predictive validity (see Greenwald et al., in press). Interestingly, the SB-IAT's impact on voting behavior was mediated by its impact on the attitude rating, a finding consistent with recent models that suggest a (default) bottom-up influence of associative processes on propositional/reflective processes of evaluation (e.g., see Gawronski &

Bodenhausen, 2006). One might have expected that reducing the amount of method-specific variance should be accompanied by an increase of construct-specific variance and thus, higher implicit-explicit correlations for the SB-IAT as compared to the standard IAT. The SB-IAT's correlation with the attitude rating, however, was of the same magnitude as the IAT's correlation with the attitude rating in Klauer et al.'s (2007) Study 2. Importantly, this finding corresponds to recent research showing that even partialing out method-specific variance by means of a control IAT only slightly increases implicit-explicit correlations (see Back et al., 2005; Mierke & Klauer, 2003) as can be easily calculated using a formula provided by Mierke and Klauer (2003, p. 1188).

General Discussion

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

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

Secondly, Olson and Fazio (2004) showed that IAT effects are confounded by "extra-personal associations", that is, culturally shared assumptions (e.g., apples are healthy and thus, are positive) that do not necessarily correspond to personal evaluations (e.g., I don't like apples). Although not explicitly stated by the authors, one might hypothesize that this confounding is based on differences in the extent to which participants *strategically* use extra-personal knowledge "when solving the mapping problem posed by the IAT" (Olson & Fazio, 2004, p. 661). Importantly, strategy use requires the consistent mapping of categories onto response keys across many consecutive trials, such as in the critical IAT blocks (cf. Fazio & Olson, 2003; Strayer & Kramer, 1994). Inasmuch as the IAT's contamination by extra-personal associations is based on its block structure, we would expect the SB-IAT to be less affected by this confound, although of course, this assumption needs empirical testing.

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

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

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

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

References

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

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

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

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

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

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

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

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692-731.

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

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

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

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

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

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

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

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

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

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

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

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

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

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2006). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (pp. 265-292). New York, NY: Psychology Press.

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

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

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

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

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

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

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

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

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

Author note

Sarah Teige-Mocigemba and Karl Christoph Klauer, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany; Klaus Rothermund, Institut für Psychologie, Friedrich-Schiller-Universität Jena, Jena, Germany.

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

Correspondence concerning this article should be addressed to Sarah Teige-Mocigemba, Institut für Psychologie, Sozialpsychologie und Methodenlehre, Albert-Ludwigs-Universität Freiburg, 79085 Freiburg, Germany. E-mail: teige@psychologie.uni-freiburg.de. phone: +49-761-2032474, fax: +49-761-2032417.

Footnotes

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

[2] Note that adapting the D_2-score (Greenwald et al., 2003) also revealed non-significant correlations: The flower-insect SB-IAT did not correlate with the TSA SB-IAT, $r = .17$, $p = .37$, or with the geometry SB-IAT, $r = .29$, $p = .12$, nor were the correlations between the TSA SB-IAT and the geometry SB-IAT significant, $r = .24$, $p = .21$.

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

[4] Note that using the D_2-score (Greenwald et al., 2003) led to the same results. The SB-IAT showed satisfactory reliability ($\alpha = .81$), discriminated between red- vs. black-voting participants, $t(37) = 3.47$, $p = .001$, and correlated with the attitude rating, $r = .44$, $p < .01$.

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

Tables

Table 1

Single-Block Implicit Association Test (SB-IAT): Task Sequence

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

Note. [a] = 2 warm-up-trials + 24 practice trials;

[b] = 2 warm-up-trials + 24 practice trials (2 [position] x 2 [category] x 6 [stimuli]);

[c] = 4 warm-up-trials + 24 compatible test trials (6 stimuli per category) + 24 incompatible test trials (6 stimuli per category);

Note that with the exception of the geometry SB-IAT in Study 1 all SB-IATs presented each stimulus once above and once below the dashed line in each test block, respectively.

In the TSA SB-IAT, the category labels ‘number’ and ‘letter’ substituted the labels ‘insect’ and ‘flower’, and ‘calculation’ and ‘word’ substituted ‘negative’ and ‘positive’. In the geometry SB-IAT, ‘red’ and ‘blue’ substituted ‘insect’ and ‘flower’, and ‘small’ and ‘large’ substituted ‘negative’ and ‘positive’. In the political attitude SB-IAT, only target category labels changed: ‘black’ and ‘red’ substituted the labels ‘insect’ and ‘flower’.

Appendix C

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

Minimizing the Influence of Recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF)

Klaus Rothermund
Friedrich-Schiller-Universität Jena

Sarah Teige-Mocigemba
Albert-Ludwigs-Universität Freiburg

Anne Gast
Friedrich-Schiller-Universität Jena

Dirk Wentura
Saarland Universität Saarbrücken

in press, *Quarterly Journal of Experimental Psychology*

Abstract

Recoding processes can influence the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) in a way that impedes an unequivocal interpretation of the resulting compatibility effects. We present a modified version of the IAT that aims to eliminate recoding, the IAT-RF (short for "IAT-recoding free"). In the IAT-RF, compatible and incompatible assignments of categories to responses switch randomly between trials within a single experimental block. Abandoning an extended sequence of consistent category-response mappings undermines recoding processes in the IAT-RF. Two experiments reveal that the IAT-RF is capable of assessing compatibility effects between the nominally defined categories of the task and effectively prevents recoding. By enforcing a processing of the stimuli in terms of their task-relevant category membership, the IAT-RF eliminates the confounding of compatibility effects with task-switch costs and becomes immune against biased selections of stimuli.

Keywords: Implicit Association Test, recoding processes, task-switch costs, consistent mapping, stimulus effects.

Since its publication in 1998, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) soon became one of the most popular response-time measures for the implicit assessment of cognition (for a recent review of IAT research, see Nosek, Greenwald, & Banaji, 2006). Despite its popularity among researchers who use the IAT as a convenient tool for an implicit assessment of compatibilities (or “associations”) between categories, there is still a lively debate regarding the underlying processes that produce IAT effects and, relatedly, how IAT effects can or should be interpreted (e.g., Arkes & Tetlock, 2004; Blanton, Jaccard, Gonzales, & Christie, 2006; Brendl, Markman, & Messner, 2001; De Houwer, 2001, 2003b; De Houwer, Geldof, & De Bruycker, 2005; Fiedler, Messner, & Bluemke, 2006; Greenwald, Nosek, Banaji, & Klauer, 2005; Kinoshita & Peek-O’Leary, 2005, 2006; Klauer & Mierke, 2005; Mierke & Klauer, 2001, 2003; Mitchell, 2004; Rothermund & Wentura, 2001, 2004; Rothermund, Wentura, & De Houwer, 2005; Steffens et al., 2004; Wentura & Rothermund, 2007).

In this article, we focus on a specific question regarding the underlying processes and possible interpretation of IAT effects – namely, we investigate the influence of recoding processes in the IAT. Recoding means that the presented stimuli are not categorized in terms of the so-called “nominal categories” of the IAT (Greenwald et al., 2005). Instead, the stimuli can be categorized on the basis of some other feature that discriminates both between the two target categories and between the two attribute categories, which typically helps to simplify the dual-categorization task in the compatible block of an IAT. We argue that recoding has a strong influence on the IAT and prevents an unequivocal interpretation of IAT effects. We then present a modified version of the IAT that aims to eliminate recoding, the IAT-RF (short for “IAT–recoding free”). Results of two experiments reveal that the IAT-RF is able to minimize recoding and is capable of assessing compatibility effects of the nominal categories.

Recoding in the IAT

In a standard IAT (Greenwald et al., 1998), participants have to categorize stimuli of two target categories (A, B) and two attribute categories (a, b) by the help of two response keys (right vs. left). In the compatible block of the task, compatible target and attribute categories are assigned to the same response keys (a/A-right vs. b/B-left), whereas in the incompatible block of the task, incompatible target and attribute categories are mapped onto the same response keys (a/B-right vs. b/A-left). An IAT effect is computed as the difference between the average response latencies in the incompatible and compatible blocks.

Compatibility effects in the IAT can be confounded by recoding processes (De Houwer et al., 2005; Rothermund et al., 2005; Wentura & Rothermund, 2007). Recoding means that the target and/or attribute stimuli are categorized on the basis of features that do not match the nominal definitions of the respective target and/or attribute categories. For example, if all (or most) of the target exemplars can be mapped consistently onto the two attribute categories (i.e., exemplars of the target category A are characterized by the attribute a, whereas exemplars of the target category B are characterized by the attribute b), participants may tend to categorize the target stimuli on the basis of their attribute qualities (instead of their target category membership) in the compatible block of the task (a/A vs. b/B). Such a recoding strategy reduces the complex 4-to-2 categorization task of the IAT to a single binary decision (target and attribute stimuli are both categorized according to the attribute dichotomy a vs. b). Since such a recoding strategy is not possible in the incompatible block of the IAT (a/B vs. b/A), response latencies are longer in the incompatible block, which produces an IAT effect.

To illustrate the possibility of recoding, we refer to a recent study by Gray, MacCulloch, Smith, Morris, and Snowden (2003) who used a version of the IAT with the attribute categories “pleasant” versus “unpleasant” (with words like UGLY written in upper case) and target categories “peaceful” versus “violent” (with words like kill written in lower case). The average IAT effect (denoting that violence is negatively evaluated) amounted to about 500 ms. Given the arguments above, it can be assumed that a recoding in terms of pleasantness suggests itself for the congruent block: Presumably, the words of the “violence” category would receive unpleasantness ratings as high as the words of the “unpleasant” category if they were interspersed in a list of stimuli of a norm study for obtaining pleasantness norms. And presumably, the same applies for the “peaceful” category in terms of pleasantness values. Therefore, such a version of the IAT can be criticized for revealing something trivial: In the compatible block, participants will press one key for anything pleasant and one key for anything unpleasant while ignoring the case information (and target category) of the words; in the incompatible block, however, they have to meticulously focus on case to decide which key has to be pressed for a pleasant or unpleasant word in the given trial.[1] It was clearly not intended by the creators of the IAT that IAT effects reflect this kind of process (see, e.g., Greenwald & Nosek, 2001; Nosek & Hansen, in press).

However, recoding is not restricted to a reframing of the target categories in terms of the attributes. In principle, any feature that allows for a simple mapping of one target and one attribute category onto one response and of the remaining two categories onto the other response can be used for recoding in the compatible block. Even structural or perceptual qualities that are completely unrelated to the meaning or definition of the categories can be used for such a recoding (De Houwer et al., 2005; Mitchell, 2004; Rothermund & Wentura, 2001, 2004). For example, Rothermund and Wentura (2001, 2004) showed that salience asymmetries within the target and attribute dimensions are used for a recoding of categories that are conceptually unrelated (see also Kinoshita & Peek-O'Leary, 2005, 2006).

Evidence for the influence of recoding in the IAT has been reported by Mierke and Klauer (2001, 2003; see also Rothermund & Wentura, 2001). They found that effects of switching from the target categorization task to the attribute categorization task (or vice versa) typically produced much smaller switch costs (Rogers & Monsell, 1995) in the compatible block than in the incompatible block of an IAT. This reduction of task-switch costs indicates that only one categorization task is applied in the compatible block—that is, the target and attribute dimensions are recoded in terms of a single feature dichotomy.

Additional evidence for the influence of recoding in the IAT comes from research on stimulus effects in the IAT. Several studies have demonstrated that selecting different sets of target exemplars that are confounded with the attributes of the task can have a dramatic influence on the magnitude and even on the direction of compatibility effects in the IAT (Bluemke & Friese, 2006; Govan & Williams, 2004; Mitchell, Nosek, & Banaji, 2003; Steffens & Plewe, 2001). The cognitive basis of these stimulus effects is a recoding process: Exemplars are processed not just with respect to their task-relevant category membership but also with respect to other features (e.g., target items are processed with regard to the attributes of the task).

Minimizing Recoding Processes: The IAT-RF

To eliminate recoding in the standard IAT, we propose a modified version of the IAT, termed "IAT-RF". Basically, the IAT-RF undermines recoding by abandoning the block structure of the standard IAT.[2] In the standard IAT, trials with compatible and incompatible response assignments are presented in separate blocks of the task. This structure creates an extended consistent mapping of categories onto responses, which invites recoding processes in the compatible block. In the IAT-RF, however, compatible and incompatible assignments of categories to responses vary randomly between trials within a single experimental block. Response assignments are indicated at the beginning of each trial by presenting the category labels in the respective corners of the screen that correspond to the response keys to which the categories are assigned in the upcoming trial.[3] This random switching of response assignments across trials should prevent any type of recoding, because efficient recoding requires a consistent mapping of categories onto responses (Roßnagel, 2001; Shiffrin & Schneider, 1977; Strayer & Kramer, 1994): Any attempt to simplify the task by a recoding of the categories would lead to an overall error rate of approximately 25% (i.e., 50% of the incompatible trials).

Due to the random switching between compatible and incompatible response assignments in the IAT-RF, participants are forced to categorize all presented stimuli on the basis of their task-relevant category membership. This entails that response-time differences between the compatible and incompatible trials can be attributed to intrinsic compatibilities (and/or incompatibilities) between the nominal categories of the task. Thus, markers of recoding processes like task-switch costs and stimulus influences should be clearly reduced in the IAT-RF compared to the standard IAT.

Experiment 1: Task-Switch Costs in the IAT and IAT-RF

The first experiment had two aims: We wanted to provide evidence for the IAT-RF's capability of detecting genuine compatibility effects, and we wanted to show that the IAT-RF is less susceptible to influences of recoding than is the standard IAT. We conducted a standard IAT and an IAT-RF with the target categories flowers/insects and the attribute categories good/bad. We chose these categories because there is reason to assume that the compatibility effect that was found in the standard flower/insect IAT (e.g., Greenwald et al., 1998) reflects some kind of genuine compatibility between the nominal categories of the task. First, compatibilities between the target categories (flowers vs. insects) and valence have been demonstrated previously with other implicit measures that are not affected by recoding processes (De Houwer, 2003a). Second, Kinoshita and Peek-O'Leary (2006) recently demonstrated that although recoding processes based on salience asymmetries have an influence on compatibility effects in the Flower/Insect x Good/Bad IAT, a considerable portion of the effect is determined by other processes that may reflect genuine compatibilities between the categories. On the basis of these findings we expected to find a significant compatibility effect between the categories flowers/insects and good/bad in the IAT-RF - that is, even when recoding processes are prevented. We also propose that task-switch costs are reduced in the compatible block of the standard IAT due to recoding, but should be of equivalent magnitude for sequences of compatible and incompatible trials in the IAT-RF.[4]

Method

Participants. A total of 55 students with different majors from the Universities of Jena and Saarbrücken took part in the experiment in exchange for a small gift (a piece of fruit or a chocolate bar). A total of 16 participants took part in the standard IAT, while 39 completed the IAT-RF. Data of 3 other participants were discarded because of very slow responses (outliers in the overall distribution of mean response times) or a large amount of missing data (more than 25% invalid responses).

Stimuli and materials. Four words were selected for each of the four categories (flowers: rose, tulip, lily, and syringa; insects: maggot, spider, wasp, and moth; good: peace, healthy, humour, and holidays; bad: miserable, war, pain, and cruel). All words were presented in black color in the centre of a white screen. Category labels indicating the assignment of the categories to the left and right keys were shown in black color in the upper left and upper right part of the screen. Experiments were programmed in E-Prime (Psychology Software Tools, 2002) and were executed on Pentium computers. Participants responded by pressing the key “D” (left) or the key “L” (right) of the computer keyboard.

Procedure. Participants took part in groups of up to 4 persons and were seated in individual cubicles at a distance of approximately 50 cm from the screen. Written instructions for the tasks were given on the computer screen. Completing the experiment took less than 15 minutes.

The standard IAT consisted of five phases. First, flower and insect words had to be categorized according to their category membership (target dimension, 16 trials). Second, evaluatively positive and negative words had to be categorized on the basis of their valence (attribute dimension, 16 trials). In a third block, stimuli of all four categories were presented in a random sequence, in which flower and insect words had to be categorized on the basis of their category membership, whereas all other stimuli had to be categorized on the basis of their valence. The first 16 trials of the combined block served as practice trials, followed by 64 experimental trials (consisting of 50% target words and 50% attribute words). In a fourth block, the valence categorization (attribute dimension) was practiced again, this time with a reversed response assignment (16 trials). In a final block, the combined categorization task was conducted again with the reversed response assignment for the valence categories (16 practice trials followed by 64 experimental trials). Assignment of categories to responses and the sequence of compatible and incompatible blocks were counterbalanced across participants. Each trial consisted of the following sequence of events: A word was presented in the middle of the screen and remained on the screen until a response was registered. Following a correct response, the word was removed from the screen, and after an intertrial interval of 200 ms, the word of the next trial was presented. In case of an erroneous response, an error message was shown below the stimulus (“Error. Press the correct key to continue...”), and the stimulus was removed from the screen after the correct response key had been pressed. The category labels always remained on the screen.

The IAT-RF consisted of three phases. Simple valence discriminations (attribute dimension) were practiced in the first block (16 trials). In a second practice block, the flower/insect categorization (target dimension) was practiced; category-response assignments switched randomly between trials (16 trials). In a third block, both categorization tasks were combined: Flower and insect words had to be categorized on the basis of their membership in one of the two target categories; good/bad words had to be categorized on the basis of their valence (32 practice trials, 128 experimental trials). Trials with compatible and incompatible response assignments (as well as target and attribute stimuli) were randomly intermixed within the combined block (50% compatible, 50% incompatible). Response assignments for the attribute dimension (good vs. bad) remained constant throughout the entire experiment for each participant but response assignments for the target categories (flowers vs. insects) switched randomly between trials. Flower, insect, good, and bad words were presented in a random sequence throughout the combined block. Each trial started with the presentation of the four differently paired category labels in the corresponding upper left and right parts of the screen. This ensured that participants could grasp the key assignments for each trial. A period of 1,000 ms later, a fixation cross appeared in the middle of the screen to reorient participants' attention from the labels to the spatial position where the next stimulus word would appear. The fixation cross was replaced after another 500 ms by a stimulus word that remained on the screen until a response was registered. Following a correct response, the stimulus was removed from the screen, whereas after an erroneous response, an error message was shown below the stimulus ("Error. Press the correct key to continue..."), and the stimulus was removed from the screen after the correct response key had been pressed. The category labels were also removed from the screen simultaneously with the stimulus. After an intertrial interval of 250 ms during which the screen was blank the next trial started with the presentation of the category labels for the upcoming trial.

Results

Experimental trials of the combined blocks were used for the analyses. Response latencies of erroneous responses and outlier values[5] (IAT: 2.49%; IAT-RF: 1.82%) were excluded from the analyses. The two tasks did not differ with regard to average response latencies – IAT: M = 791 ms, SD = 92; IAT-RF: M = 817 ms, SD = 97; $t < 1$, $d = 0.27$ – or error frequencies – IAT: 6.69%; IAT-RF: 4.91%; $t(53) = 1.64$, *ns*, $d = 0.49$. Split-half reliabilities of compatibility effects were $r_{tt} = .80$ for the standard IAT and $r_{tt} = .63$ for the IAT-RF.

Compatibility effects. A large compatibility effect was found for the standard IAT, $t(15) = 5.67$, $p < .01$, $d = 1.42$, indicating that responses were faster in the compatible block ($M = 712$ ms, $SD = 98$) than in the incompatible block ($M = 869$ ms, $SD = 116$; see Figure 1).[6] For the IAT-RF, a significant compatibility effect of medium size emerged, $t(38) = 3.39$, $p < .01$, $d = 0.54$, indicating that responses were faster for trials with a compatible response assignment ($M = 801$ ms, $SD = 94$) than for incompatible trials ($M = 832$ ms, $SD = 107$; see Figure 1). A combined analysis revealed that compatibility effects were significantly larger for the standard IAT than for the IAT-RF, $F(1, 53) = 30.50$, $p < .01$.

Task-switch costs. Switch costs were computed as the average response-time difference between trials in which the categorization task changed compared to the preceding trial (target-attribute, attribute-target) and trials in which the categorization task was repeated (target-target, attribute-attribute). For the standard IAT, switch costs were significantly larger in the incompatible block ($M = 157$ ms, $SD = 67$) than in the compatible block ($M = 33$ ms, $SD = 63$), $t(15) = 5.51$, $p < .01$, $d = 1.38$ (see Figure 2).[7] To compare switch costs for compatible and incompatible trials in the IAT-RF, we analyzed only those trials in which the response assignments of the previous trial were repeated in the current trial (compatible-compatible, incompatible-incompatible). Only those trials allow for a categorization of switch costs as referring to sequences of either compatible or incompatible trials[8] and are comparable with the switch-cost analyses for the standard IAT. In the IAT-RF, switch costs did not differ significantly for sequences of compatible ($M = 50$ ms, $SD = 77$) and incompatible trials ($M = 72$ ms, $SD = 87$), $t(38) = 1.26$, *ns* (see Figure 2). A combined analysis yielded the predicted three-way interaction between switch costs, compatibility, and IAT type, $F(1, 53) = 10.96$, $p < .01$, indicating a significant difference between switch costs in the compatible and incompatible blocks of the standard IAT, but no such difference in the switch costs for sequences of compatible and incompatible trials of the IAT-RF.

Discussion

Experiment 1 yielded a replication of the large compatibility effect that is typically observed in the Flower/Insect x Good/Bad standard IAT (e.g., Greenwald et al., 1998; Kinoshita & Peek-O'Leary, 2006). Although we assumed that some portion of this effect reflects compatibilities between the nominal target categories and attribute categories, it is nevertheless likely that the effect is a mixture of genuine compatibilities and recoding processes. An influence of recoding on the standard IAT is to be suspected because switch costs were markedly reduced in the compatible block of this task.

A significant compatibility effect of medium size was also found in the Flower/Insect x Good/Bad IAT-RF. Unlike in the standard IAT, switch costs in the compatible and incompatible conditions did not differ significantly, indicating that recoding was prevented in the IAT-RF. The compatibility effect thus attests to the existence of an overlap between the cognitive representations of the nominal categories flowers/insects and good/bad, and it also attests to the IAT-RF's capability of detecting existing compatibilities between nominally defined categories.

Experiment 2: Stimulus Confounds in the IAT and IAT-RF

By avoiding extended phases of a consistent mapping of categories onto responses, the IAT-RF undermines recoding and enforces a processing of the exemplar stimuli in terms of their task-relevant category membership. As was shown in Experiment 1, randomly switching between compatible and incompatible response assignments in the IAT-RF eliminates the reduction of task-switch costs for sequences of compatible trials that is typically observed in the compatible block of a standard IAT.

Another implication of inducing a strict processing focus on the task-relevant features of the exemplar stimuli is that compatibility effects in the IAT-RF should be immune against biased selections of target stimuli. Several studies have demonstrated strong effects of biased selections of target stimuli on compatibility effects in the standard IAT (Bluemke & Friese, 2006; Govan & Williams, 2004; Mitchell et al., 2003; Steffens & Plewe, 2001). These findings reveal that the target exemplars are also processed in terms of their attribute category membership, indicating a recoding of the target categorization task.

In the IAT-RF, however, target items should be processed first of all with respect to their task-relevant category membership, because this is the only information that consistently allows for an accurate performance on both compatible and incompatible trials. Processing of other features would lead to conflict in half of the target trials. Enforcing a focus on the task-relevant information, therefore, should neutralize effects of biased selections of target stimuli in the IAT-RF.

Two standard IATs and two IAT-RFs using old/young and good/bad as target and attribute categories, respectively, were conducted to test this assumption. Trait adjectives that were stereotypically "old" or "young" were selected as exemplars for the two target categories. In one variant of both IATs, the target items of the category "old" consisted entirely of negatively valent adjectives, whereas only positive adjectives were chosen to represent the category "young" (valence-consistent version). For the other variant of both tasks, the opposite bias was implemented in the selection of the target stimuli (valence-inconsistent version). We predicted that confounding the selection of target items with valence should have a marked influence on compatibility effects in the standard IAT but should leave compatibility effects in the IAT-RF unaffected.

Method

Participants. A total of 69 students with different majors from the University of Jena took part in the experiment in exchange for a small gift (a piece of fruit or a chocolate bar). Data of 9 other participants were discarded because they were not native German speakers ($n = 4$), because of very slow responses (outliers in the overall distribution of mean response times), or due to a large amount of missing data (more than 25% invalid responses). The 5 participants that were discarded due to a large number of slow or erroneous responses were evenly distributed across the four types of IAT (2 participants were discarded who had conducted the IAT-RF with positive old and negative young words; for each of the remaining three versions of the IAT, 1 participant was discarded). Each participant accomplished one of four possible Old/Young x Good/Bad IATs ($16 \leq n \leq 19$ for each IAT variant): The four IATs resulted from combining IAT type (standard IAT vs. IAT-RF) with exemplar valence for the target categories (old-negative/young-positive vs. old-positive/young-negative).

Stimulus materials and procedure. Four words were selected for each category (good: peace, humour, gain, and holidays; bad: murder, anger, hatred, and terror; old-negative: frail, senile, confused, cranky; old-positive: experienced, kind, dignified, considerate; young-negative: immature, naïve, spoilt, careless; young-positive: healthy, spontaneous, modern, easy-going). Words for the categories old and young were selected from a pilot study ($N = 15$) to ensure that all sets of stimuli were equally stereotypical for the respective categories (scale ranging from –4, highly typical for old people, to +4, highly typical for young people) and had a clear positive or negative valence (scale ranging from –4, highly negative, to +4, highly positive). The selected word sets satisfied these criteria: old-negative: $M_{\text{stereo}} = -2.4$, $SD = 0.56$, range = (–3.1, –1.8); $M_{\text{valence}} = -2.3$, $SD = 0.69$, range = (–2.9, –1.5); old-positive: $M_{\text{stereo}} = -2.3$, $SD = 0.42$, range = (–2.9, –1.8); $M_{\text{valence}} = +2.5$, $SD = 0.47$, range = (+1.8, +3.1); young-negative: $M_{\text{stereo}} = +2.3$, $SD = 0.84$, range = (+1.9, +2.7); $M_{\text{valence}} = -1.8$, $SD = 0.55$, range = (–2.5, –1.3); young-positive: $M_{\text{stereo}} = +2.5$, $SD = 0.69$, range = (+2.2, +3.0); $M_{\text{valence}} = +2.4$, $SD = 0.56$, range = (+1.7, +3.5). Stimuli for the attribute categories (good vs. bad) were identical in all IATs. Stimulus sets for the categories old and young consisted either of negative adjectives for old and positive adjectives for young (old-negative/young-positive) or of positive adjectives for old and negative adjectives for young (old-positive/young-negative). Category labels were identical in all four IATs (old vs. young, good vs. bad). Procedural details of the IAT and IAT-RF were identical to those in Experiment 1.

Results

Experimental trials of the combined blocks were used for the analyses. Response latencies of erroneous responses and outlier values ($IAT_{valence\text{-}consistent}$: 3.77%; $IAT_{valence\text{-}inconsistent}$: 2.98%; $IAT\text{-}RF_{valence\text{-}consistent}$: 3.41%; $IAT\text{-}RF_{valence\text{-}inconsistent}$: 2.34%; see Footnote 6) were excluded from the analyses. Comparing the four tasks with regard to average response latencies revealed a significant main effect of IAT type, $t(67) = 2.90$, $p < .01$, $d = 0.70$, indicating faster responses for the IAT than for the IAT-RF ($IAT_{valence\text{-}consistent}$: $M = 768$ ms, $SD = 106$; $IAT_{valence\text{-}inconsistent}$: $M = 762$ ms, $SD = 67$; $IAT\text{-}RF_{valence\text{-}consistent}$: $M = 824$ ms, $SD = 117$; $IAT\text{-}RF_{valence\text{-}inconsistent}$: $M = 856$ ms, $SD = 129$). No significant differences between tasks emerged with regard to error frequencies, $t < 1$ ($IAT_{valence\text{-}consistent}$: 5.47%; $IAT_{valence\text{-}inconsistent}$: 5.71%; $IAT\text{-}RF_{valence\text{-}consistent}$: 4.98%; $IAT\text{-}RF_{valence\text{-}inconsistent}$: 5.47%). Split-half reliabilities of compatibility effects were $r_{tt} = .81$ for the $IAT_{valence\text{-}consistent}$, $r_{tt} = .34$ for the $IAT_{valence\text{-}inconsistent}$, $r_{tt} = .60$ for the $IAT\text{-}RF_{valence\text{-}consistent}$, and $r_{tt} = .57$ for the $IAT\text{-}RF_{valence\text{-}inconsistent}$.

Compatibility effects. A large compatibility effect was found for the valence-consistent variant (old-negative/young-positive targets) of the standard IAT, $t(16) = 6.66$, $p < .01$, $d = 1.62$, indicating that responses were faster in the block in which old and negative were assigned to the same response ($M = 667$ ms, $SD = 87$) than when young and negative shared the same response ($M = 876$ ms, $SD = 150$; see Figure 3). This compatibility effect was reversed, however, for the valence-inconsistent variant (old-positive/young-negative targets) of the standard IAT, $t(15) = -1.87$, $p < .05$ (one-tailed), $d = 0.47$, indicating that responses were now faster if old and positive were assigned to the same response ($M = 750$ ms, $SD = 78$) than in the block in which young and positive shared the same response ($M = 785$ ms, $SD = 78$). A combined analysis of the two standard IATs revealed a main effect of compatibility, $F(1, 31) = 22.04$, $p < .01$, indicating that, overall, responses were faster if old and negative were assigned to the same response. This main effect was qualified by a compatibility x target valence consistency interaction, $F(1, 31) = 43.22$, $p < .01$, indicating that the compatibility effect that was found in the old-negative/young-positive version of the IAT was reversed in the old-positive/young-negative version of the task.[9]

For the IAT-RF, compatibility effects were similar for the two variants of the task: For the valence-consistent (old-negative/young-positive targets) version of the task, responses were faster if old and negative were assigned to the same response ($M = 808$ ms, $SD = 117$)

than when young and negative shared the same response (*M* = 841 ms, *SD* = 122), $t(18) = 2.90$, $p < .01$, $d = 0.66$ (see Figure 3). A similar effect was found for the valence-inconsistent (old-positive/young-negative targets) version of the IAT-RF, $t(16) = 3.07$, $p < .01$, $d = 0.74$, again indicating faster responses for the old-negative response assignment (*M* = 838 ms, *SD* = 128) than for the young-negative response mapping (*M* = 874 ms, *SD* = 134). Accordingly, only the main effect of compatibility reached significance in a combined analysis of the two IAT-RFs, $F(1, 34) = 17.75$, $p < .01$, whereas the compatibility x target valence consistency interaction was not significant, $F < 1$.

To test whether the pattern of findings differed significantly between IAT and IAT-RF, a combined analysis of all four IATs was conducted, revealing the predicted three-way interaction of compatibility, IAT type, and target valence consistency, $F(1, 65) = 39.44$, $p < .01$.

Discussion

Conceptually replicating previous findings, confounding the selection of target stimuli in terms of the attributes had a marked influence on the resulting compatibility effect of a standard Old/Young x Good/Bad IAT. Selecting exemplars that are valence consistent with the prototypical negative old age stereotype and with a positive stereotype of the category "young" produced corresponding compatibility effects that were reversed, however, for a selection of target stimuli with an opposite valence bias.[10]

By contrast, the same manipulation of the valence of the target items did not have an influence on the direction or magnitude of the compatibility effects in the Old/Young x Good/Bad IAT-RF. A significant compatibility effect of the same direction and size was found for both variants of the IAT-RF, indicating that responding is facilitated if the categories "old" and "bad" are assigned to one response, and "young" and "good" are assigned to the other response. Apparently, the IAT-RF was successful in establishing a processing focus on the task-relevant category information, which prohibited a processing of the target stimuli in terms of their (task-irrelevant) valence. By implication, the compatibility effect in the IAT-RF indicates more or less pure categorization processes with respect to the nominal categories of the task. The resulting compatibility effect reflects an overlap between the target and attribute categories that can be taken as an inherent quality of the mental representations of the concepts "old" and "young".

General Discussion

We introduced a variant of the IAT that aims to eliminate recoding processes, which is why we called it the "recoding-free" IAT (IAT-RF). The IAT-RF abandons the separation of compatible and incompatible blocks of trials that is typical for the standard IAT. Trials with compatible and incompatible response assignments are randomly intermixed within a single experimental block. This random sequence of response assignments precludes a consistent mapping of categories and responses over an extended period of time, which is a prerequisite for an efficient recoding of the categorization task.

Compatibility effects of the standard IAT and IAT-RF were compared in two experiments. In Experiment 1, the IAT-RF revealed genuine compatibility effects between the categories flowers/insects and good/bad. In addition, we found that whereas task-switch costs were strongly reduced in the compatible blocks of the standard IAT, no significant difference was found between the switch costs for sequences of compatible and incompatible trials in the IAT-RF. Experiment 2 revealed that compatibility effects in an Old/Young x Good/Bad IAT-RF were not affected by the choice of exemplars. Instead, a highly similar compatibility effect was obtained for a stimulus selection that confounded old with bad (and young with good) and for a selection containing the opposite confound. In line with previous findings, compatibility effects in the standard version of the Old/Young x Good/Bad IAT were reversed by such a valence manipulation (Bluemke & Friese, 2006; Govan & Williams, 2004; Mitchell et al., 2003; Steffens & Plewe, 2001). Taken together, these findings support our hypothesis that standard IAT effects are contaminated by recoding processes whereas the IAT-RF effectively prevents recoding and is capable of detecting connotations of the nominal target categories of the task.

Process Models of the IAT

The results of this study have important implications for process models of the IAT and IAT-RF. The finding that task-switch costs were much smaller in the compatible block than in the incompatible block of the standard IAT implies that a recoding of target stimuli in terms of the attributes plays an important role in this task. Similarly, the dependence of standard IAT effects on the valence of the target exemplars indicates that the attribute categorization task is also applied to the target stimuli. These findings yield further support for the task-switching account of the standard IAT (Klauer & Mierke, 2005; Mierke & Klauer, 2001, 2003).

On the other hand, we still found significant medium-sized compatibility effects in the IAT-RF although switch costs did not differ between sequences of compatible and incompatible trials and although the valence of the target stimuli had no effect on the resulting compatibility effects for this task. Compatibility effects in the IAT-RF thus have to be attributed to other causes. The relevant feature account proposed by De Houwer (2001, 2003b) provides a plausible explanation of recoding-free compatibility effects in the IAT-RF. According to this account, compatibility effects can result from an overlap between relevant stimulus features (i.e., their target category membership) and response characteristics that are established during the task by assigning a specific attribute category to a response. We therefore assume that IAT-RF effects reflect genuine compatibilities between the nominal target and attribute categories of the task.

Recoding is a Multifaceted Phenomenon

The evidence that we provided in this study was mainly concerned with recoding in its most blatant form—that is, a recoding of the target categories in terms of the attributes. This kind of recoding is most likely to occur with regard to attitude IATs, in which positive and negative valence represent the attribute categories. As was already stated in the introduction, however, recoding processes are not restricted to a recoding of the target categories in terms of the attributes. Other features of the stimuli that help to discriminate between the target and attribute categories in a consistent fashion can also be used for recoding (e.g., familiarity, salience, size, color; cf. Mierke & Klauer, 2003; Rothermund & Wentura, 2001, 2004).

Importantly, recoding is not restricted to strategic processes but can also occur automatically - that is, without a conscious plan or strategy (i.e., recoding can result from an implicit learning of covariations between features and responses; Lewicki, Hill, & Czyzewska, 1992). Due to the subtlety and opaqueness of the recoding process, it can be difficult or even impossible to identify the feature (or combination of features) that is responsible for a given IAT effect.

Implications for the Use of the IAT and of the IAT-RF

As a consequence of the indeterminacy of the exact nature of the features that might have been used for recoding, IAT effects cannot be interpreted with confidence as reflecting compatibilities between the nominal categories of the task (De Houwer et al., 2005; Rothermund et al., 2005). For example, it certainly makes a difference for the interpretation of an Old/Young x Bad/Good IAT whether “old” and “bad” (“young” and “good”) are compatible in the sense that “bad” (“good”) is an intrinsic feature of the mental representation of the category “old” (“young”) that is automatically activated whenever a stimulus is categorized as “old” (“young”), or whether familiarity, salience, or any other feature was used to discriminate between the pairs of categories.

Recoding also poses a problem for the interpretation of interindividual differences in IAT effects because it introduces an additional source of variance into IAT effects that reflects differences regarding the use of recoding (e.g., in order to fake the IAT; Fiedler &

Bluemke, 2005; Kim, 2003; Steffens, 2004) or with respect to the efficiency of recoding processes (e.g., method variance due to a “cognitive skill confound”; Back, Schmukle, & Egloff, 2005; McFarland & Crouch, 2002; Mierke & Klauer, 2003). Inasmuch as interindividual differences in recoding are unrelated to compatibilities between the nominal categories of an IAT, they may lead to biases in the strength of IAT effects.

The IAT-RF was developed to solve the specific problem that the IAT is not a pure measure of compatibilities between the nominal target and attribute categories. Researchers who are interested in an unbiased assessment of genuine compatibilities between nominal categories might prefer the IAT-RF because it is capable of enforcing a processing of all stimuli in terms of their task-relevant category memberships and because it represents a pure measure of category effects that is immune against stimulus confounds.

Due to its immunity against influences of recoding in terms of explicit preferences, the IAT-RF also is a promising measure for the prediction of behavior, particularly with regard to behavior that is underdetermined by explicit preferences. Recent studies attest to the predictive validity of the IAT-RF for political attitudes (Teige-Mocigemba et al., in press) and in the domain of ambivalent behaviors like smoking (Teige-Mocigemba, Klauer, & Rothermund, 2007) or alcohol consumption (Houben, Rothermund, &Wiers, 2007). Based on the promising findings in these studies, we expect that the IAT-RF will turn out to be a sensitive indicator of category evaluations related to attitudes.

We want to state explicitly, however, that an assessment of “pure” conceptual compatibilities is not the only legitimate research objective and that the standard IAT might outperform the IAT-RF in other regards. In particular, the standard IAT might be better suited to predict explicit attitudes (except perhaps in a situation where participants try to hide their preferences or have a motive to fake the IAT). The reason for this superiority of the standard IAT is that recoding can be based on explicit attitudes, which would increase correlations between IAT effects and explicit measures. Inasmuch as the maximization of such a predictive relation is the main research goal, we would recommend the standard IAT in most cases. A boost of implicit-explicit correlations that is mediated by an influence of explicit attitudes on recoding processes, however, does not necessarily attest to the validity of the IAT as an implicit measure; nor would the lack of a correlation between an explicit measure and the IAT-RF necessarily betray the validity of the latter measure.

To conclude, we would like to recommend the IAT-RF as a process-pure implicit measure for assessing compatibilities between nominal categories. Although effect sizes will certainly be smaller than what is typically found with the standard IAT, IAT-RF effects have the advantage that they can be more easily interpreted in terms of genuine compatibilities between mental representations of the nominal categories of the task.

References

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?". *Psychological Inquiry*, *15*, 257-278.

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

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, *42*, 192-212.

Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects? *Journal of Experimental Social Psychology*, *42*, 163-176.

Brendl, M., Markman, A., & 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.

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. (2003a). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77-85.

De Houwer, J. (2003b). 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 Associates, Inc.

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

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

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology*, *27*, 307-316.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A" and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*, 74-147.

Gast, A., & Rothermund, K. (2007). *No matter if senile or experienced? Effects of stimulus-selections on four different implicit measurement methods*. Unpublished manuscript, Friedrich-Schiller-University Jena, Jena.

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

Gray, N. S., MacCulloch, M. J., Smith, J., Morris, M., & Snowden, R. J. (2003). Violence viewed by psychopathic murderers. *Nature*, *423*, 497-498.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie*, *48*, 85-93.

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

Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the salience asymmetry interpretation of the Implicit Association Test: Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General*, *134*, 420-425.

Houben, K., Rothermund, K., & Wiers, R. W. (2007). *Eliminating recoding in the Alcohol-IAT: An application of the IAT-RF*. Manuscript submitted for publication.

Kim, D.-Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly*, *66*, 83-96.

Kinoshita, S., & Peek-O'Leary, M. (2005). Does the compatibility effect in the race Implicit Association Test (IAT) reflect familiarity or affect? *Psychonomic Bulletin and Review*, *12*, 442-452.

Kinoshita, S., & Peek-O'Leary, M. (2006). Two bases of the compatibility effect in the Implicit Association Test (IAT). *Quarterly Journal of Experimental Psychology*, *59*, 2102-2120.

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

Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist*, *47*, 796-801.

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, C. J. (2004). Mere acceptance produces apparent attitude in the Implicit Association Test (IAT). *Journal of Experimental Social Psychology*, *40*, 366-373.

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

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

Nosek, B. A., & Hansen, J. J. (in press). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition and Emotion*.

Psychology Software Tools (2002). *E-Prime 1.1*. Pittsburgh, PA: Psychology Software Tools, Inc.

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

Roßnagel, C. S. (2001). Revealing hidden covariation detection: Evidence for implicit abstraction at study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1276-1288.

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. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*, 139-165.

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Validity of the salience asymmetry account of the Implicit Association Test: Reply to Greenwald, Nosek, Banaji, and Klauer (2005). *Journal of Experimental Psychology: General*, *134*, 426-430.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-190.

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology*, *51*, 165-179.

Steffens, M. C., Lichau, J., Still, Y., Jelenec, P., Anheuser, J., Goergens, N. K., & Hülsebusch, T. (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 of reaction time differences in the Implicit Association Test (IAT)]. *Zeitschrift für Psychologie*, *212*, 57-65.

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

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

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2007, September). *Beyond the IAT's block structure: A Single Block IAT*. Paper presented at the 11th Meeting of the Section of Social Psychology of the German Association of Psychology, Freiburg, Germany.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (in press). Minimizing method-specific variance in the IAT: A single block IAT. *European Journal of Psychological Assessment*.

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

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

Author note

Correspondence should be addressed to Klaus Rothermund, Friedrich-Schiller-Universität Jena, Institut für Psychologie, Am Steiger 3, Haus 1, D-07743 Jena, Germany. E-mail: klaus.rothermund@uni-jena.de.

We want to thank Johannes Kretzschmer for developing the programs for the experiments and Catharina Casper, Carina Giesen, Franziska Meißner, Bettina Seifert, and Lisa Schubert for collecting the data.

Footnotes

[1] The main point of the article by Gray et al. (2003) was that the IAT effect was significantly reduced for psychopathic murderers (to about 350 ms). It is unclear, however, whether the difference in IAT effects between the two groups of participants is due to a difference in the strength of associations, to a difference in recoding, or both.

[2] A similar variant of a blockless IAT has recently been proposed by Eichstaedt (2007), however, with a different theoretical purpose.

[3] Another variant of the IAT-RF that is currently being tested in our laboratories (Teige-Mocigemba, Klauer, & Rothermund, in press) uses an additional feature of the stimulus (like word position) to indicate the category-response assignment for the respective trial (e.g., if the stimulus appears in the upper/lower half of the screen, responses have to be given on the basis of the compatible/ incompatible mapping).

[4] It should be noted that the IAT-RF introduces a new type of switch costs that refer to a switch between compatible and incompatible response mappings (compared to those trials in which the response mapping of the previous trial is repeated). These “mapping-switch costs”, however, are unrelated to recoding processes because they do not refer to a switching between the target and attribute categorization tasks and thus are not indicative of whether the two tasks are reduced to a single task.

[5] Values that were below 250 ms or that were more than three interquartile ranges above the median of the overall response-time distribution were treated as outliers (Tukey, 1977).

[6] We also computed the *D* measure for the compatibility effect of the standard IAT (Greenwald, Nosek, & Banaji, 2003) and found a large compatibility effect ($M = 0.45$, $SD = 0.39$), $t(15) = 4.62$, $p < .01$.

[7] A similar task switch x compatibility interaction also emerged if the *D* measure (Greenwald et al., 2003) was used as dependent variable instead of simple response-time indicators, $t(15) = 3.00$, $p < .01$. *D* was more than twice as large for the task-switch trials ($M = 0.64$, $SD = 0.51$) as it was for the task-repetition trials ($M = 0.31$, $SD = 0.42$), indicating that the *D* measure does not eliminate effects of recoding in the standard IAT.

[8] Trials of the IAT-RF in which response assignments were switched between the previous and the current trial (“mapping-switch trials”, compatible-incompatible, incompatible-compatible) cannot be classified as being either compatible or incompatible with regard to task-switch costs because the first trial of the task-switch or task-repetition sequence is always opposite in compatibility to the second. A categorization of the respective

task-switch or task-repetition RT as referring to a sequence of either compatible or incompatible trials is logically impossible in this case.

[9] A similar reversal of compatibility effects for the standard IAT was also found for the *D* measure (Greenwald et al., 2003). The difference between the old-negative/young-positive version (M = +0.65, SD = 0.35) and the old-positive/young-negative version (M = –0.01, SD = 0.39) was highly significant, $t(31) = 5.11$, $p < .01$. Apparently, using the *D* algorithm does not eliminate recoding effects of biased stimulus sets in the standard IAT.

[10] It has been suggested that biasing influences of stimuli in the standard IAT should be avoided by selecting sets of target stimuli that are balanced with regard to valence (Steffens et al., 2004). In our view, however, this strategy does not eliminate the problem of recoding completely. In another study, we found that recoding still had an influence on compatibility effects in the standard IAT, even if the target stimuli were balanced with regard to valence (Gast & Rothermund, 2007). Specifically, compatibility effects differed significantly for stimuli of opposite valence within a target category, indicating that the valence of the target stimuli had a substantial influence on response times. Apparently, balancing the target stimuli with regard to valence does not suffice to eliminate recoding, as long as the compatible and incompatible trials are presented in separate blocks. Furthermore, a recoding of the targets in terms of the attributes is just one possibility of how recoding can operate. Any feature that can help to reduce the complexity of the categorization task can be used for recoding (familiarity, salience, etc.). Balancing the targets with respect to valence thus does not rule out the possibility that other features are used for recoding.

Figure Captions

Figure 1. Average response latencies (and standard errors) for compatible and incompatible response assignments in the standard IAT and in the IAT-RF for the Flower/Insect x Good/Bad IAT (Experiment 1).

Figure 2. Average task-switch costs (and standard errors) in the compatible and incompatible blocks of a standard IAT and for sequences of compatible and incompatible trials in the IAT-RF for the Flower/Insect x Good/Bad IAT (Experiment 1).

Figure 3. Average response latencies (and standard errors) for compatible (old/negative vs. young/positive) and incompatible (old/positive vs. young/negative) response assignments in the standard IAT and in the IAT-RF depending on the valence bias of target items (Experiment 2).

Figures

Figure 1

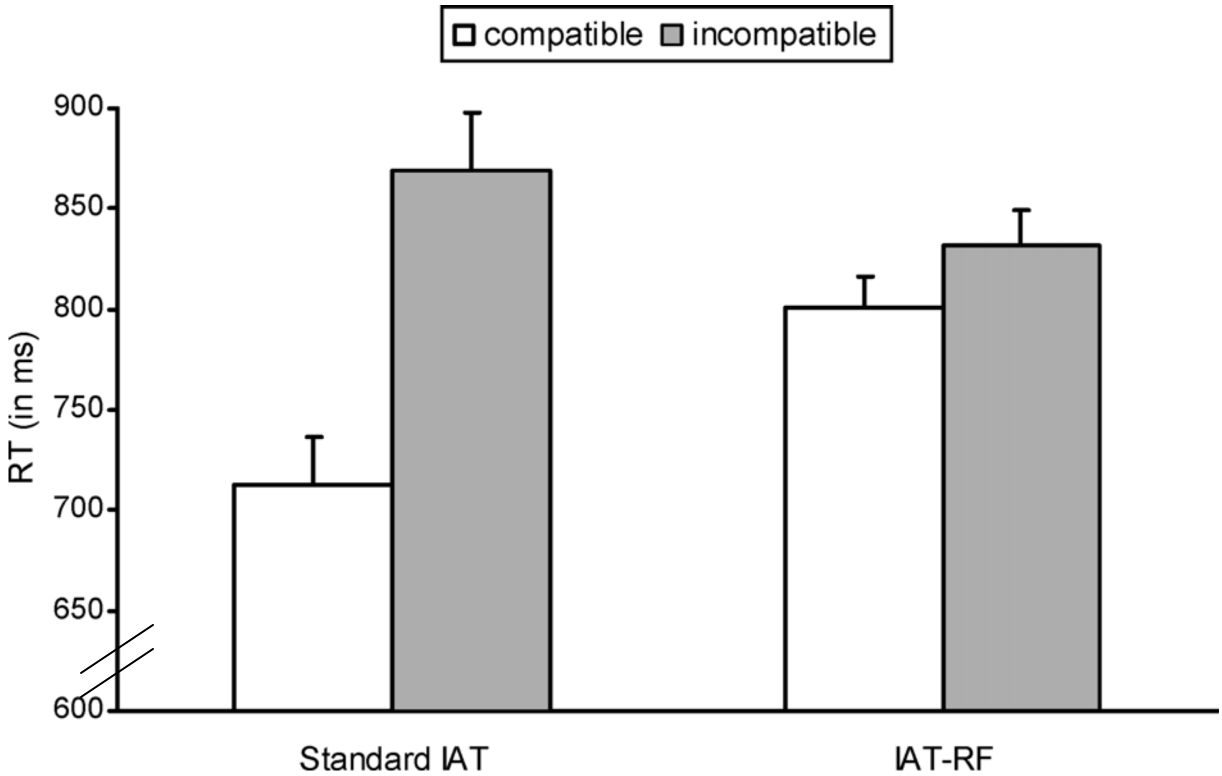


Figure 2

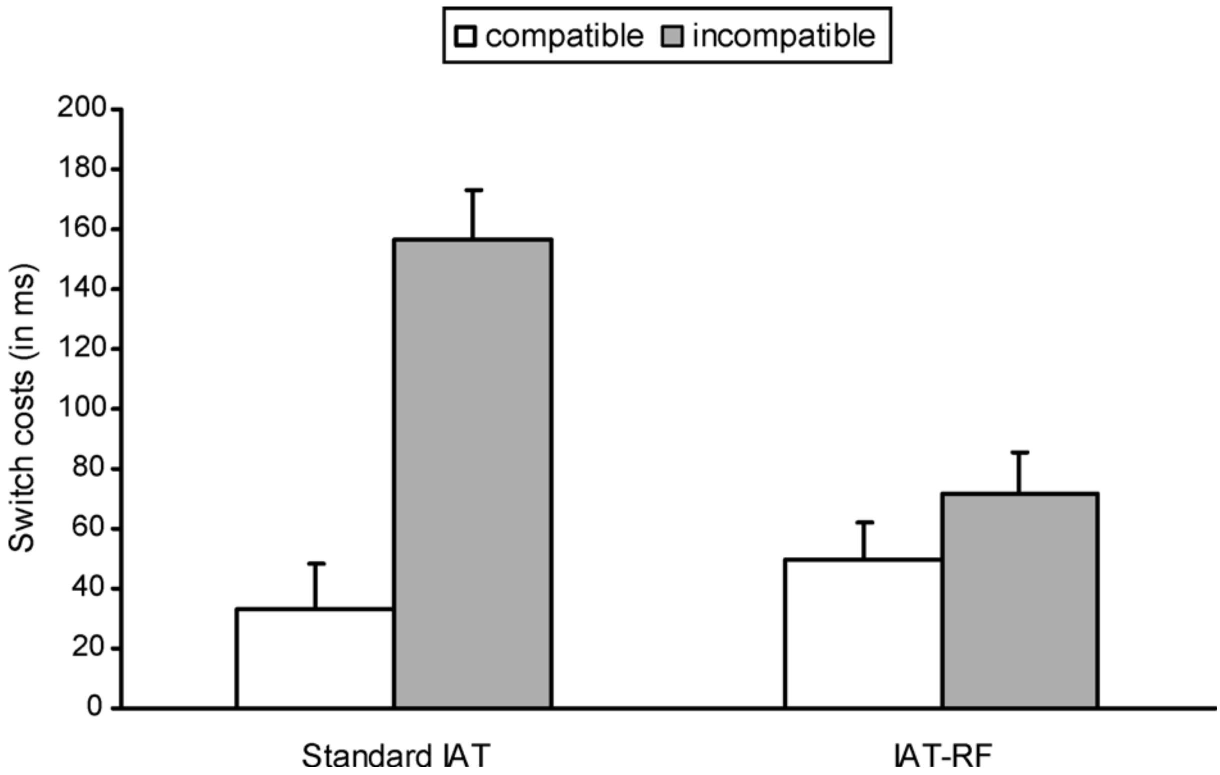
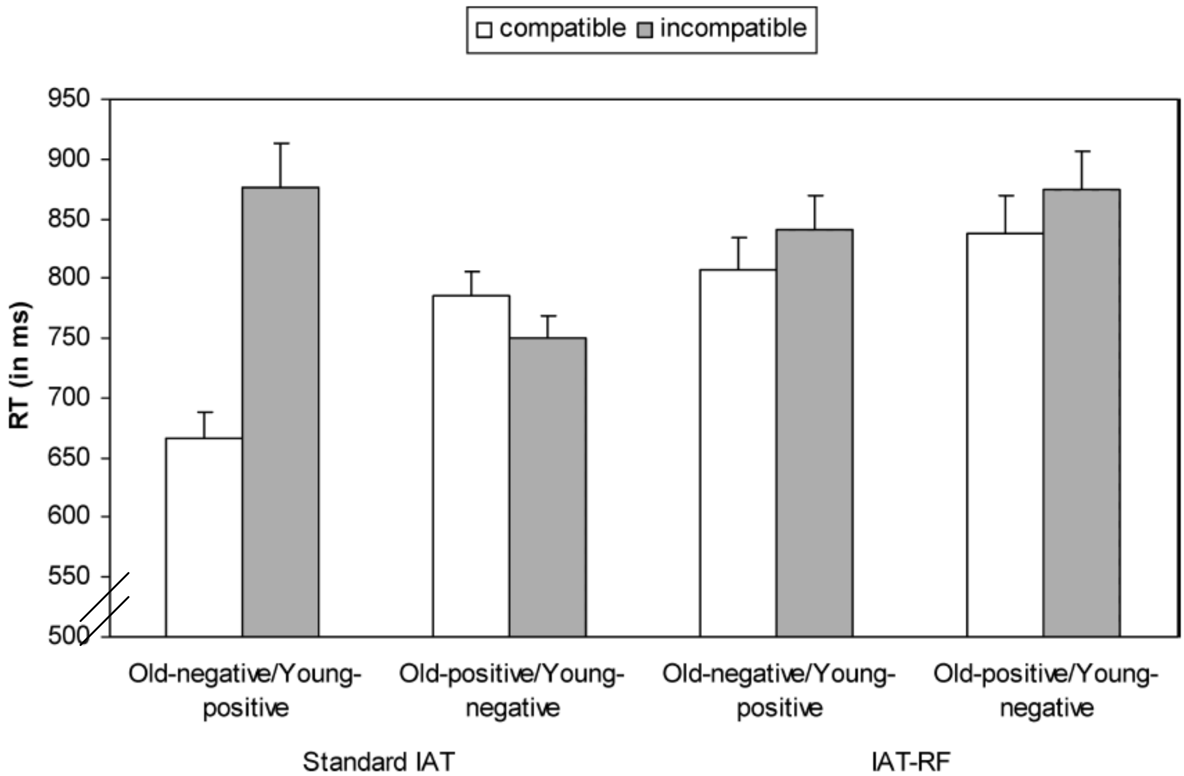

compatible
incompatible
200
180
160
140
120
100
80
60
40
20
0
Switch costs (in ms)
Standard IAT
IAT-RF

Figure 3


compatible
incompatible
950
900
850
800
750
700
650
600
550
500
RT (in ms)
Old-negative/Young-positive
Old-positive/Young-negative
Old-negative/Young-positive
Old-positive/Young-negative
Standard IAT
IAT-RF

Curriculum Vitae

Dipl. - Psych.

Sarah Teige-Mocigemba

Personal Data

Date of Birth	01.03.1979
Place of Birth	Berlin-Wilmersdorf (Germany)
Family Status	Married

Contact Data

Office	University of Freiburg Institute of Psychology Social Psychology and Methodology Engelbergerstraße 41 79106 Freiburg ++49-761-2032474 sarah.teige@psychologie.uni-freiburg.de
Private	Hauptstraße 13b 79104 Freiburg ++49-761-2089548

Professional Positions

since 11/2004	**Research and Teaching Assistant** at the Department of Social Psychology and Methodology at University of Freiburg
since 07/2001	**Freelancer** (Human Resources) at Deutsche Bahn AG
09/2001-09/2003	**Student Research and Teaching Assistant** at the Department of Personality Psychology at Humboldt-University, Berlin
02/2002-04/2002	**Internship in the Medical Clinic Charité**
04/2001-07/2001	**Freelance Interviewer** in the research project "Application of New Media in Civil Engineering" at the Department of Media Consulting, University of Technology, Berlin
11/2000-09/2001	**Student Research Assistant** in the research project "Implicit Personality Assessment" at Humboldt-University, Berlin
02/1999-02/2000	**Interviewer** at EMNID (market and opinion research)

Education

07/2007-08/2007	**EAPP Summer School** on "Implicit Measures of Personality", Bertinoro (Italy)
05/2006	**Workshop for PhD students** (So-Doc), University of Jena
since 11/2004	Several **Workshops for PhD students** (Oberrhein-group), Universities of Freiburg, Basel, Heidelberg, and Mannheim
04/2003-04/2005	**Sinology** at Humboldt-University, Berlin
10/1998-12/2003	**Psychology** at Humboldt-University, Berlin Grade of diploma: 1.0; grade of intermediate diploma: 1.1
08/1991-06/1998	Georg-Herwegh-Gymnasium Berlin (grade of Abitur: 1.4)
02/1987-06/1991	Victor-Gollancz-Grundschule, Berlin
08/1985-01/1987	Grundschule Buisdorf-Fahrdorf, Schleswig-Holstein

Awards

2003	**'Best Diploma Thesis 2003'** at the Department of Psychology, Humboldt-University, Berlin

Teaching

SS 08	Seminar: Social Cognition
WS 07/08	Seminar: Intimate Relationships
WS 06/07 – SS 07	Research Course: 'Strategical recoding and extra-personal associations in IAT variants'
WS 05/06 – SS 06	Research Course: 'Single Block IAT'
WS 05/06	Seminar: Psychological Research Methods
WS 04/05 – SS 05	Research Course: 'The influence of stimulus prototypicality on the EAST and IAT'
WS 02/03 – SS 03	Research Course: 'Assessment of multiple implicit traits using the EAST'
WS 01/02 – SS 03	Seminar: Methods of Observational Research

Ad-hoc Reviewing

Experimental Psychology

Cognition and Emotion

Social Cognition

Journal of Applied Social Psychology

Memberships

Deutsche Gesellschaft für Psychologie (DGPs; associated member)

Society for Personality and Social Psychology (SPSP)

Publications

Note. References printed in bold are part of this dissertation.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (in press). Minimizing method-specific variance in the IAT: A Single Block IAT. *European Journal of Psychological Assessment*.

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

Teige-Mocigemba, S., & Klauer, K. C. (in press). 'Automatic' evaluation? Strategic effects on affective priming. *Journal of Experimental Social Psychology.*

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (in press). Contrast effects in spontaneous evaluations: A psychophysical account. *Journal of Personality and Social Psychology.*

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

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43,* 648-655.

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

Conference Contributions

Teige-Mocigemba, S., & Klauer, K. C. (2008). *Effects of compatibility order on the IAT's validity.* Paper presented at the 19th International Congress of Psychology, Berlin, Germany, July 20-25, 2008.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2008). *Beyond the IAT's block structure: A Single Block IAT.* Invited talk at the Department of Experimental Clinical and Health Psychology (Jan De Houwer), Ghent, Belgium, April 24-25, 2008.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2007). *Beyond the IAT's block structure: A Single Block IAT.* Paper presented at the 11th Meeting of the Section of Social Psychology of the German Association of Psychology, Freiburg, Germany, September 17-19, 2007.

Rothermund, K., Casper, C., Gast, A., Giesen, C., Meißner, F., Schubert, L., Seifert, B., Teige-Mocigemba, S., & Wentura, D. (2007). A*usschaltung von Rekodierungseffekten im IAT: Der IAT-RF.* Paper presented at the 11th Meeting of the Section of Social Psychology of the German Association of Psychology, Freiburg, Germany, September 17-19, 2007.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2007). *Minimizing the influence of recoding processes on the IAT-effect: The Single Block IAT (SB-IAT).* Paper presented at the 9th Transfer of Knowledge Meeting of the European Social Cognition Network, Brno, Czech Republic, September 05-09, 2007.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2007). *On the problem of the IAT's block structure.* Paper presented at the EAPP Summer School on Implicit Measures of Personality, Bertinoro, Italy, July 29-August 03, 2007.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2007). *Minimizing the Influence of Recoding Processes on the IAT: A Single Block IAT.* Paper presented at the 12th Biennial Meeting of the International Society for Studying Individual Differences (ISSID), Gießen, Germany, July 22-27, 2007.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2007). *Minimizing Recoding Processes in the IAT: A Single Block IAT.* Paper presented at the 9th European Conference on Psychological Assessment (EAPA), Thessaloniki, Greece, May 3-6, 2007.

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (2007). *Contrastive Effects in Evaluative Priming.* Paper presented at the 49th meeting of Experimental Psychology (TeaP), Trier, Germany, March 26-28, 2007.

Teige-Mocigemba, S., & Klauer, K. C. (2007). *'Automatische' Evaluation? Strategische Effekte im affektiven Priming.* Paper presented at the 49th meeting of Experimental Psychology (TeaP), Trier, Germany, March 26-28, 2007.

Teige-Mocigemba, S., Rothermund, K., & Klauer, K. C. (2006). *Reduzierung methodenspezifischer Varianz im IAT: Der blocklose IAT.* Paper presented at the 45th meeting of the German Association of Psychology (DGPS), Nürnberg, Germany, September 17-19, 2006.

Schmitz, F., Teige-Mocigemba, S., Voss, A., & Klauer, K. C. (2006). *Validität eines IAT zur politischen Einstellung: Die moderierende Rolle von interindividuellen Unterschieden im Arbeitsgedächtnis.* Paper presented at the 45th meeting of the German Association of Psychology, Nürnberg, Germany, September 17-19, 2006.

Voss, A., & Teige-Mocigemba, S. (2006). *Der "Stimulus- Comparison- Test" (SCoT): Ein neues Verfahren zur indirekten Erfassung von Assoziationen.* Paper presented at the 48th meeting of Experimental Psychology (TeaP), Mainz, Germany, March 26-29, 2006.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2006). *Prozesskomponenten des IAT.* Paper presented at the 48th meeting of Experimental Psychology (TeaP), Mainz, Germany, March 26-29, 2006.

Teige-Mocigemba, S., Rothermund, K., & Klauer, K. C. (2006). *Minimierung von Rekodierungsprozessen im IAT-Effekt: Der blocklose IAT.* Paper presented at the 48th meeting of Experimental Psychology (TeaP), Mainz, Germany, March 26-29, 2006.

Stahl, C., Teige, S., Klauer, K. C., & Degner, J. (2005). *Effekte der Stimulusprototypikalität auf IAT und EAST: Eine Anwendung des Quad-Modells.* Poster presented at the 47th meeting of Experimental Psychology (TeaP), Regensburg, Germany, April 04-07, 2005.

Schmitz, F., Teige, S., Voss, A., & Klauer, K. C. (2005). *Working memory load in the IAT.* Paper presented at the 5th Workshop on Implicit Representations and Personality, Berlin, Germany, June 24-25, 2005.

Voss, A., Schmitz, F., Teige, S., & Klauer, K. C. (2005). *Cognitive Capacity, Task Shifts, and Measurement Issues in the EAST Procedure.* Paper presented at the EAST Workshop, Oppurg, Germany, April 08-10, 2005.

Eidesstattliche Erklärung (declaration of originality)

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- beziehungsweise Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

(I hereby declare that this dissertation is my own work and that all the sources that I have used or quoted have been acknowledged by means of complete references. This work has not been submitted previously for a degree at any university or other academic institution.)

Freiburg, den 06.08.2008