

# Der Implizite Assoziationstest als Maß automatisch aktivierter Assoziationen: Reichweite und Grenzen

Bertram Gawronski und Frederica R. Conrey

**Zusammenfassung.** Der Implizite Assoziationstest (IAT) ist wahrscheinlich das bekannteste und am weitesten verbreitete Verfahren zur Messung automatisch aktivierter Assoziationen. Trotz der vorliegenden Evidenz für dessen Validität gibt es allerdings nach wie vor Kritik am IAT, die vor allem auf der ungeklärten Frage nach den zugrundeliegenden psychologischen Prozessen beruht. Der folgende Artikel gibt einen Überblick über die vorhandene Forschung zum IAT, um auf diese Weise sowohl die Reichweite als auch die Grenzen des Verfahrens aufzuzeigen. Thematisiert werden dabei Zusammenhänge zu expliziten Fragebogenmaßen, Ergebnisse zur Vorhersagekraft des IAT, Studien zu Kontext- und Materialeffekten, sowie die Rolle systematischer Fehlervarianz. Als Resultat dieses Überblicks stehen eine Reihe von allgemeinen Schlussfolgerungen zur konvergenten, diskriminanten, prädiktiven, inkrementellen und internen Validität des IAT, sowie einige kritische Anmerkungen zu weit verbreiteten Interpretationen in der IAT-Forschung.

Schlüsselwörter: Assoziative Prozesse, Implizite Kognition, Impliziter Assoziationstest, Kognitive Kontrolle, Validität

The Implicit Association Test as a measure of automatically activated associations: Range and limits

**Abstract.** The Implicit Association Test (IAT) is probably the most well known and most widespread measure of automatically activated associations. Even though there is supportive evidence for the validity of the IAT, there are still controversies about its underlying psychological processes. The following article provides an overview of the available research on the IAT, illustrating both the range and the limits of the measure. This review includes the relation to explicit self-report measures, evidence for its predictive ability, studies on context and material effects, and the role of systematic error variance. Based on this review, we draw conclusions about the convergent, discriminant, predictive, incremental, and internal validity of the IAT, and critically discuss some common interpretations in IAT research.

Key words: associative processes, implicit cognition, implicit association test, cognitive control, validity

In der Geschichte der Psychologie hat wahrscheinlich kaum ein anderes Verfahren in so kurzer Zeit eine solche Aufmerksamkeit erlangt wie der Implizite Assoziationstest (IAT; Greenwald, McGhee & Schwartz, 1998). Aufbauend auf das erst vor kurzer Zeit neu entstandene Forschungsprogramm zu Prozessen der impliziten sozialen Kognition (Greenwald & Banaji, 1995) wurde der IAT als Instrument zur Messung individueller Unterschiede in der automatischen Aktivierung von evaluativen oder semantischen Assoziationen entwickelt. Das Interesse an solchen automatischen Aktivierungsprozessen erfuhr mit der Entwicklung des IAT ein fast exponentielles Wachstum, was sich bereits drei Jahre nach der Publikation des IAT in Sonderheften in der *Zeitschrift für Experimentelle Psychologie*

(Plessner & Banse, 2001) und im *Journal of Personality and Social Psychology* (Devine, 2001) widerspiegelte. Ursprünglich aus der Sozialpsychologie stammend ist der IAT mittlerweile in fast alle Bereiche der Psychologie vorgedrungen, angefangen von der Persönlichkeitspsychologie (z. B. Asendorpf, Banse & Mücke, 2002), über die klinische Psychologie (z. B. Teachman, Gregg & Woody, 2001), bis hin zur Konsumentenpsychologie (z. B. Maison, Greenwald & Bruin, 2001), der Gesundheitspsychologie (z. B. Wiers, Van Woerden, Smulders & De Jong, 2002), der Entwicklungspsychologie (z. B. Skowronski & Lawrence, 2001), der Altersforschung (z. B. Hummert, Gartska, O'Brien, Greenwald & Mellott, 2002) und der Neuropsychologie (z. B. Phelps et al., 2000).

Auch wenn es bereits eine Reihe von Studien gibt, die für die Validität des IAT als Instrument zur Messung automatisch aktivierter Assoziationen sprechen, so wird der IAT nach wie vor sehr kontrovers diskutiert. Die dabei am häufigsten genannten Kritikpunkte sind, (a) dass die dem IAT zugrundeliegenden psychologischen Prozesse bis

Diese Arbeit wurde von einem Forschungsstipendium der Deutschen Forschungsgemeinschaft an den Erstautor unterstützt (Ga 747/1-1). Die Autoren danken Rainer Banse und Alexandra Freund für hilfreiche Anmerkungen zu einer früheren Version dieses Artikels.

heute nicht hinreichend geklärt sind, sowie (b) dass unter Umständen nicht-assoziative Prozesse für systematische Varianz in IAT-Scores verantwortlich sind.

Ziel des vorliegenden Artikels ist es, einen Überblick über die bisherige Forschung zum IAT zu geben, um auf diese Weise sowohl die Reichweite als auch die Grenzen des Verfahrens aufzuzeigen (für einen Vergleich verschiedener impliziter Verfahren, siehe Fazio & Olson, 2003). Zu diesem Zweck soll zunächst der IAT als Verfahren kurz vorgestellt und seine Vorteile gegenüber anderen Verfahren zur Messung automatisch aktivierter Assoziationen dargestellt werden. Im Anschluss daran wird ein systematischer Überblick über die bisherige Forschung zum IAT gegeben. Thematisiert werden dabei Zusammenhänge zu expliziten Fragebogenmaßen, Ergebnisse zur Vorhersagekraft des IAT, Studien zu Kontext- und Materialeffekten, sowie die Rolle systematischer Fehlervarianz im IAT. Aufbauend auf diesen Überblick werden abschließend einige allgemeine Schlussfolgerungen über die Validität des IAT als Maß automatisch aktivierter Assoziationen gezogen. Diese Schlussfolgerungen beinhalten auch eine Reihe von kritische Anmerkungen zu gängigen Interpretationen in der IAT-Forschung.

## Der Implizite Assoziationstest

Der IAT besteht aus einer computergestützten Diskriminationsaufgabe, bei der Stimuli zweier dichotomer Dimensionen so schnell wie möglich kategorisiert werden sollen. Ein klassisches Beispiel ist der von Greenwald et al. (1998) verwendete IAT zur impliziten Präferenz von Weißen gegenüber Schwarzen (siehe auch Rudman, Greenwald, Mellott & Schwartz, 1999). In diesem IAT müssen die Versuchspersonen Namen oder Bilder von schwarzen und weißen Personen (Objekt-Stimuli) per Tastendruck so schnell wie möglich zu den Kategorien „schwarz“ oder „weiß“ zuordnen. Alternierend mit dieser Objekt-Diskriminations-Aufgabe wird eine evaluative Entscheidungsaufgabe vorgegeben, in der normativ positive und negative Worte (Attribut-Stimuli) so schnell wie möglich den Kategorien „positiv“ oder „negativ“ zugeordnet werden müssen. Entscheidend sind beim IAT zwei unterschiedliche Kombinationen dieser beiden Diskriminationsaufgaben: eine im Sinne der zu messenden Assoziationen kongruente und eine im Sinne der zu messenden Assoziationen inkongruente. So müssen zum Beispiel weiße Versuchspersonen in der assoziationskongruenten Kombination mit der gleichen Taste auf positive Worte und weiße Personen bzw. auf negative Worte und schwarze Personen reagieren. In der assoziationsinkongruenten Kombination muss dagegen mit der gleichen Taste auf negative Worte und weiße Personen bzw. auf positive Worte und schwarze Personen reagiert werden. Der Unterschied in den mittleren Reaktionszeiten zwischen assoziationskongruenter und assoziationsinkongruenter Kombination wird üblicherweise als Maß für die automatische (assoziative) Präferenz für Weiße gegenüber Schwarzen interpretiert (für eine Diskussion unterschiedlicher Algorithmen zur Berechnung von IAT-Scores, siehe Greenwald, Nosek & Banaji, 2003).

Die weite Verbreitung des IAT hat sicher eine ganze Reihe von Gründen. Zu erwähnen sind dabei auf jeden Fall die im IAT häufig zu findenden hohen Effektstärken im Vergleich zu anderen Verfahren zur Messung automatisch aktivierter Assoziationen (Greenwald et al., 1998). Ein weiterer Grund, der vor allem für differenzielle Forschungszwecke von Bedeutung ist, ist die im Vergleich zu anderen impliziten Verfahren meist sehr hohe interne Konsistenz des IAT. Während eine Reihe von impliziten Verfahren so niedrige interne Konsistenzen aufweisen, dass sie für differenzielle Zwecke nicht geeignet sind (Banse, 1999, 2001; Teige, Schnabel, Banse & Asendorpf, in press), zeigen viele Studien mit dem IAT befriedigende interne Konsistenzen mit Werten zwischen .80 und .90 (Cronbachs  $\alpha$ , z. B. Asendorpf et al., 2002; Egloff & Schmukle, 2002; Gawronski, Ehrenberg, Banse, Zukova, & Klauer, 2003 a; Teige et al., in press). Des Weiteren ist der IAT als Verfahren sehr robust und leicht einsetzbar, so dass er auch für computergestützte Studien außerhalb des Labors ohne Probleme eingesetzt werden kann (z. B. Banse & Fischer, 2002). Schließlich ist der Anwendungsbereich des IAT äußerst flexibel, was sich unter anderem in der bereits angeführten weiten Verbreitung in verschiedenen Teildisziplinen der Psychologie sowie in der Vielzahl der Themen von bisher durchgeführten IAT-Studien zeigt.

## Zusammenhang zu Fragebogenmaßen

Nach Campbell und Fiske (1959) kann die Konvergenz bzw. Divergenz eines neuen Messverfahrens zu bereits etablierten Verfahren als zentraler Validitätshinweis interpretiert werden. Bei der Frage nach dem Zusammenhang zwischen IAT und expliziten Fragebogenmaßen wird allerdings häufig eine gewisse Willkür kritisiert. Zeigt der IAT nur geringe Korrelationen mit expliziten Maßen, so wird dies oft als Evidenz dafür interpretiert, dass der IAT und explizite Maße unterschiedliche Konstrukte messen (diskriminante Validität; z. B. Greenwald et al., 1998). Werden hingegen hohe Korrelationen gefunden, so gilt dies nicht selten als Indikator dafür, dass der IAT und explizite Maße das gleiche Konstrukt messen (konvergente Validität; z. B. Greenwald et al., 2003). Solche arbiträren ad-hoc Annahmen müssen wissenschaftstheoretisch als willkürlich und den wissenschaftlichen Fortschritt hemmend interpretiert werden (Gawronski, 2000). Da die Zusammenhänge zwischen IAT und Fragebogenmaßen zudem sehr stark variieren (für eine Meta-Analyse, siehe Hofmann, Gawronski, Gschwendner, Le & Schmitt, 2003), erscheint es daher sinnvoller, nicht nach absoluten Zusammenhängen zu fragen, sondern systematisch nach möglichen Moderatoren des Zusammenhangs zwischen IAT und expliziten Maßen zu suchen.

Erste Hinweise für solche Moderatoren stammen aus der Vorurteils- und Stereotypforschung. So belegen eine Reihe von Studien, dass Personen mit einer geringen Motivation zur Vorurteilkontrolle höhere Zusammenhänge zwischen implizit und explizit gemessenen Vorurteilen zeigen als Personen mit einer hohen Motivation zur Vor-

urteilskontrolle (z. B. Banse & Gawronski, 2003; Banse, Seise & Zerbse, 2001; Gawronski, Geschke & Banse, 2003b; Hofmann, Gschwendner & Schmitt, 2003). Konsistent mit diesen Befunden sind auch Ergebnisse von Florack, Scarabis und Bless (2001a), die zeigen, dass explizit geäußerte Vorurteile einen umso höheren Zusammenhang zu implizit gemessenen Vorurteilen aufwiesen, je stärker sich die Versuchspersonen von der in Frage stehenden Gruppe bedroht fühlten. Dieser Befund deutet darauf hin, dass sich die Motivation zur Kontrolle von negativen Assoziationen verringert, wenn die in Frage stehende Gruppe subjektiv als Bedrohung wahrgenommen wird. Umgekehrt sank dagegen der Zusammenhang zwischen IAT und explizit gemessenen Vorteilen, je höher bei den Versuchspersonen die individuelle Motivation zu einer kognitiv aufwändigen Informationsverarbeitung ausgeprägt war (Florack, Scarabis & Bless, 2001b). Dieses Ergebnis legt nahe, dass allein intensives Nachdenken vorurteilsvolle Assoziationen als Urteilsbasis in Frage stellen kann. In einer Untersuchung zum Einfluss kognitiver Dissonanz konnten Gawronski und Strack (2004) darüber hinaus zeigen, dass implizit und explizit gemessene Einstellungen relativ hoch korrelieren, wenn durch einstellungskonträres Verhalten verursachte kognitive Dissonanz (vgl. Festinger & Carlsmith, 1959) mit Hilfe einer situationalen Erklärung reduziert werden konnte, nicht aber wenn kognitive Dissonanz durch eine (explizite) Einstellungsänderung reduziert wurde. Gawronski und Strack (2004) erklären diese Ergebnisse damit, dass kognitive Dissonanz auf expliziten propositionalen Prozessen beruht, und damit nur explizit, nicht aber implizit gemessene Einstellungen beeinflussen sollte. Darüber hinaus sollten evaluative Assoziationen im Normalfall als Basis für evaluative Urteile dienen, nicht aber wenn kognitive Dissonanz zu einer Zurückweisung von evaluativen Assoziationen führt.

Die zur Zeit wahrscheinlich umfangreichste Studie zur Moderation des Zusammenhang zwischen IAT und Fragebogenmaßen stammt von Nosek (2004). In einer Internetstudie mit über 6000 Versuchspersonen und über 50 verschiedenen IATs konnte Nosek zeigen, dass der Zusammenhang zwischen IAT und Fragebogenmaßen sowohl von personenbezogenen als auch von objektbezogenen Faktoren abhängt. So zeigten IAT und Fragebogenmaße umso höhere Zusammenhänge (a) je geringer bei den Versuchspersonen die Motivation zur Selbstpräsentation ausgeprägt war, (b) je höher die individuelle Stärke der Einstellung war, und (c) je höher bei den Versuchspersonen die wahrgenommene Diskrepanz zwischen der eigenen Einstellung und der anderer Personen war. Darüber hinaus zeigten sich signifikant höhere Zusammenhänge, (d) wenn die Einstellung durch Bipolarität charakterisiert war, d. h., wenn eine positive Einstellung zu einer Kategorie eine negative Einstellung zur anderen Kategorie implizierte.

Auch wenn die Zusammenhänge zwischen IAT und expliziten Fragebogenmaßen systematisch von verschiedenen Moderatoren beeinflusst werden, ist es mindestens ebenso wichtig zu erwähnen, dass die häufig gefundenen

niedrigen Korrelationen zumindest teilweise auf einem hohen Anteil von Fehlervarianz beruhen können. So konnten Cunningham, Preacher und Banaji (2001) zeigen, dass der Zusammenhang zwischen IAT und expliziten Fragebogenmaßen deutlich steigt, wenn die Fehlervarianz beider Maße kontrolliert wird (siehe hierzu auch Hofmann et al., 2003).

Dass der IAT nicht nur konvergente sondern auch diskriminante Validität im Zusammenhang zu expliziten Maßen aufweist, belegt eine Studie von Gawronski (2002). In dieser Studie korrelierte der IAT nur dann mit expliziten Fragebogenmaßen, wenn die erfassten Einstellungen der beiden Maße korrespondierten, nicht aber wenn sich die Einstellungsobjekte unterschieden. So bestand ein signifikanter Zusammenhang zwischen explizit und implizit gemessenen Einstellungen gegenüber Türken sowie zwischen explizit und implizit gemessenen Einstellungen gegenüber Asiaten. Bei nicht korrespondierenden Gruppen zeigten sich dagegen keine signifikanten Zusammenhänge zwischen IAT und expliziten Maßen. Dieses Ergebnis zeigt, dass Rassismus-IATs tatsächlich spezifische Fremdgruppenevaluationen erfassen, und nicht nur eine allgemeine, von der jeweiligen Fremdgruppe unabhängige Tendenz zur Eigengruppenfavorisierung (z. B. Ashburn-Nardo, Voils & Monteith, 2001; Cunningham, Nezlek & Banaji, in press).

## Vorhersagekraft

Aufbauend auf verschiedene Zwei-Prozess-Modelle der sozialen Informationsverarbeitung (z. B. Fazio & Towles-Schwen, 1999; Wilson, Lindsey & Schooler, 2000) wird häufig angenommen, dass implizite Maße wie der IAT vor allem spontanes Verhalten vorhersagen sollten, während explizite Maße primär kontrolliertes Verhalten vorhersagen sollten. Konsistent mit diesen Annahmen konnten Asendorpf et al. (2002) zum Beispiel zeigen, dass schwer zu kontrollierende schüchterne Verhaltensmuster (z. B. Gesichtsadaptoren) substanziell mit einem Schüchternheits-IAT zusammenhängen. Leichter zu kontrollierende schüchterne Verhaltensmuster (z. B. Sprechdauer) zeigten dagegen höhere Zusammenhänge zu expliziten Schüchternheitsmaßen. Ähnliche Dissoziationen in der Vorhersage von spontanem und kontrolliertem Verhalten berichten Egloff und Schmukle (2002) für einen IAT zu Ängstlichkeit, Bosson, Swann und Pennebaker (2000) für einen Selbstwert-IAT, Perugini (in press) für einen IAT zu Esspräferenzen, sowie McConnell und Leibold (2001) für einen IAT zu impliziten Vorurteilen gegenüber Schwarzen.

Über die Vorhersage spontanen Verhaltens hinaus belegen eine Reihe von Studien die prädiktive Validität des IAT bei der Vorhersage von Informations-Verarbeitungs-Prozessen. Aufbauend auf frühere Ergebnisse (z. B. Duncan, 1976; Dunning & Sherman, 1997; Kunda & Sherman-Williams, 1993; Sagar & Schofield, 1980) konnten mehrere Studien zeigen, dass sich verzerrte Interpretationen mehrdeutigen Verhaltens durch den IAT vorhersagen lassen (z. B. Gawronski et al., 2003b; Hugenberg & Boden-

hausen, 2003; Rudman & Glick, 2001). So wurde zum Beispiel in einem Experiment von Gawronski et al. (2003b) mehrdeutiges Verhalten negativer beurteilt, wenn der Akteur türkischer Abstammung war als wenn der Akteur deutscher Abstammung war. Die Gruppenzugehörigkeit des Akteurs spielte jedoch nur bei solchen Versuchspersonen eine Rolle, die in einem IAT stark negative Assoziationen gegenüber Türken zeigten. Bei Personen mit schwachen oder neutralen Assoziationen zeigte sich kein Effekt der Gruppenzugehörigkeit auf die Verhaltensinterpretation. Ähnliche Ergebnisse berichten auch Hugenberg und Bodenhausen (2003) für die Interpretation emotional mehrdeutiger Gesichtsausdrücke bei Afroamerikanern (siehe hierzu auch Hugenberg & Bodenhausen, 2004), sowie Rudman und Glick (2001) für die Interpretation von berufsrelevanten Leistungen bei Männern und Frauen. Explizit gemessene Vorurteile zeigten in keiner der genannten Studien Zusammenhänge mit der Interpretation mehrdeutigen Verhaltens.

Zusammenhänge zwischen IAT-Scores und verzerrten Gedächtnisleistungen lassen sich in einer Studie von Gawronski et al. (2003b) finden. Mit Hilfe eines von Klauer und Wegener (1998) entwickelten multinomialen Modells für das „Who-said-what?“ Paradigma (Taylor, Fiske, Etcoff & Ruderman, 1978) konnten diese Autoren zeigen, dass eine verstärkte implizite Geschlechtsstereotypisierung im IAT mit einem schlechteren Personengedächtnis bei der Zuordnung von individuellen Aussagen zu männlichen und weiblichen Sprechern einherging. Diese verschlechterte Rekognitionsleistung wurde von einer verstärkten geschlechtsstereotypen Ratetendenz begleitet, nach der geschlechtsstereotyp männliche Aussagen verstärkt Männern zugeschrieben wurden, während geschlechtsstereotyp weibliche Aussagen verstärkt Frauen zugeschrieben wurden.

Neben Studien zu nonverbalem Verhalten und zur sozialen Informationsverarbeitung konnte eine Reihe von Studien in der Neuropsychologie Zusammenhänge zwischen dem IAT und der Aktivität unterschiedlicher Gehirnregionen zeigen. So konnten Phelps et al. (2000) mit Hilfe eines IAT zu impliziten Vorurteilen gegenüber Afroamerikanern Aktivierungsprozesse in der Amygdala in Reaktion auf Gesichter von Afroamerikanern vorhergesagen. Da Amygdala-Aktivierung üblicherweise mit spontanem negativen Affekt in Verbindung gebracht wird, spricht dieses Ergebnis für die Annahme, dass der IAT einen direkten Bezug zu automatischen affektiven Reaktionen besitzt. Demgegenüber stehen allerdings Ergebnisse von Chee, Sriram, Soon und Lee (2000), die zeigen konnten, dass die Bearbeitung assoziationsinkongruenter Kombinationen im IAT zu einer verstärkten Aktivierung im dorsolateralen präfrontalen Kortex führt. Eine solche Aktivierung wird üblicherweise mit exekutiven Kontrollprozessen in Zusammenhang gebracht, was im Widerspruch zu einer Interpretation des IAT im Sinne automatischer affektiver Reaktionen zu stehen scheint. Auf die Bedeutung dieses Ergebnisses werden wir im Kontext systematischer Fehlervarianz noch genauer zu sprechen kommen.

## Kontexteffekte

Ursprünglich wurde für den IAT vermutet, dass dieser zeitlich stabile und äußerst robuste Assoziationen misst, die auf langjährigen Sozialisationserfahrungen beruhen (z. B. Wilson et al., 2000). Diese Annahme erscheint mittlerweile aus zwei Gründen fraglich. Auch wenn der IAT gegen bewusste Verfälschungsversuche weniger anfällig zu sein scheint als explizite Fragebogenmaße (Banse et al., 2001; Kim, 2003; Egloff & Schmukle, 2002; Steffens, in press), zeigt der IAT für ein reines Dispositionsmaß nur relativ unbefriedigende Test-Retest-Korrelationen zwischen .50 bis .70 (z. B. Bosson et al., 2000; Schmukle & Egloff, in press; Steffens & Buchner, 2003). Zum anderen belegen eine Vielzahl von Studien, dass der IAT durch relativ simple Kontextvariationen beeinflusst wird (z. B. Blair, Ma & Lenton, 2001; Dasgupta & Greenwald, 2001; Foroni & Mayr, in press; Gawronski & Bodenhausen, 2004; Karpinski & Hilton, 2001; Kühnen et al., 2001; Lowery, Hardin & Sinclair, 2001; Richeson & Ambady, 2001, 2003; Wittenbrink, Judd & Park, 2001). So konnten Dasgupta und Greenwald (2001) zum Beispiel zeigen, dass sich mit dem IAT gemessene Vorurteile gegenüber Afroamerikanern verringerten, wenn den Versuchspersonen vor der Bearbeitung des IAT Bilder von positiv bewerteten Schwarzen (z. B. Michael Jordan) und negativ bewerteten Weißen (z. B. Charles Manson) präsentiert wurden. Ähnliche Ergebnisse berichten Blair et al. (2001), die zeigen konnten, dass sich implizite Geschlechtsstereotype im IAT verringerten, wenn die Versuchspersonen vor der Bearbeitung des IAT gebeten wurden, sich eine dem klassischen Geschlechtsstereotyp widersprechende Frau vorzustellen. Ergebnisse wie diese sind zwar durchaus mit der Annahme konsistent, dass der IAT die Aktivierungsstärke von Assoziationen misst. Sie sprechen jedoch gegen die Annahme, dass es sich bei diesen Assoziationen ausschließlich um chronische Assoziationen handelt. Vielmehr scheinen situativ aktivierte Assoziationen ebenfalls einen starken Einfluss auf den IAT zu haben. Dass der IAT allerdings nicht ausschließlich situativ aktivierte, sondern tatsächlich auch chronische Assoziationen misst, kann aus Studien abgeleitet werden, in denen Verhalten über einen Zeitraum von mehr als einer Woche vorhergesagt werden konnte (z. B. Hugenberg & Bodenhausen, 2004).

## Materialeffekte

Aufbauend auf eine Studie von De Houwer (2001) wird häufig angenommen, dass der IAT reine Kategorieassoziationen misst, und dass dem spezifischen Stimulusmaterial daher nur eine untergeordnete Rolle zukommt. So konnte De Houwer (2001) zeigen, dass in einem Britisch-Ausländer-IAT die a priori Valenz der gezeigten Objektstimuli (z. B. Adolf Hitler, Mahatma Gandhi, Prinzessin Diana, Maggie Thatcher) keinen Einfluss auf die von britischen Versuchspersonen gezeigte implizite Präferenz für Briten besaß. Dem gegenüber stehen allerdings Ergebnisse von Steffens und Plewe (2001), dass stereotype

Beziehungen evaluativer Attribut-Stimuli zu den in Frage stehenden Objektkategorien IAT-Mittelwerte entscheidend beeinflussen können (z. B. „hysterisch“ vs. „einführend“ als evaluative Attribut-Stimuli in einem IAT zu impliziten Einstellungen zu Männern und Frauen). Govan und Williams (2004) konnten darüber hinaus zeigen, dass sich IAT-Mittelwerte signifikant ändern, wenn die gewählten Objekt-Stimuli im IAT zu einer veränderten Interpretation der Kategorien führt (z. B. „Brennnessel“ vs. „Tulpe“ als Objekt-Stimuli in einem IAT zu impliziten Einstellungen zu Blumen und Insekten). Ähnliche Ergebnisse berichten J. P. Mitchell, Nosek und Banaji (2003), die zeigen konnten, dass implizite Vorurteile gegenüber Afroamerikanern geringer ausfallen, wenn positiv bewertete schwarze Personen (z. B. Michael Jordan) als Objekt-Stimuli verwendet werden. Die wahrscheinlich umfangreichste Studie zum Einfluss von Attribut- und Objekt-Stimuli auf IAT-Scores stammt von Blümke und Friese (2004). Aufbauend auf die bereits berichteten Ergebnisse konnten diese Autoren zeigen, dass Mittelwerte im IAT systematisch in Abhängigkeit stereotyper Beziehungen der Attribut-Stimuli sowie der Valenz der Objekt-Stimuli variieren. Diese Ergebnisse legen nahe, dass IAT-Scores nicht allein von Kategorieassoziationen bestimmt werden, sondern auch vom Stimulusmaterial abhängen. Dies impliziert, dass eine absolute Interpretation von IAT-Scores – trotz der weit verbreiteten Praxis – eigentlich nicht möglich ist (z. B. Werte über Null als Indikator für impliziten Rassismus). Ob die Materialabhängigkeit des IAT auch dessen prädiktive Validität beeinflusst, ist bis jetzt eine offene Frage. Dies ist insofern von zentraler Bedeutung, da Mittelwertsunterschiede auf Stichprobenebene nicht unbedingt eine veränderte Rangordnung der Versuchspersonen auf differenzieller Ebene implizieren. So konnten zum Beispiel in einigen Studien auch mit einem Stichproben-IAT-Score nahe Null differenzielle Unterschiede in der Informationsverarbeitung (z. B. Gawronski et al., 2003 b) oder nonverbales Verhalten (z. B. Asendorpf et al., 2002) valide vorhergesagt werden.

## Systematische Fehlervarianz

Die stärkste Kritik am IAT beruht vor allem auf Studien zu dessen zugrundeliegenden psychologischen Prozessen (interne Validität). So konnten Mierke und Klauer (2001, 2003) zum Beispiel zeigen, dass ein substanzieller Anteil der Varianz im IAT auf unterschiedlichen Aufgabenwechselkosten beruht (Allport, Styles & Hsieh, 1994; Rogers & Monsell, 1995). Solche Unterschiede in Aufgabenwechselkosten implizieren eine zentrale Rolle exekutiver Kontrollprozesse, was wiederum gegen die Annahme spricht, dass der IAT die Stärke automatisch aktivierter Assoziationen misst. Diese Hypothese ist konsistent mit den bereits berichteten Befunden von Chee et al. (2000), dass die Bearbeitung assoziationsinkongruenter Kombinationen im IAT zu einer verstärkten Aktivierung im dorsolateralen präfrontalen Kortex führt, was häufig mit exekutiven Kontrollprozessen in Zusammenhang gebracht wird. Darüber hinaus existieren eine Reihe von differenziellen Faktoren, die mit individuellen Unterschieden in der

Fähigkeit zum Aufgabenwechsel assoziiert sind, so dass diese im IAT mit den eigentlich zu messenden differenziellen Assoziationsstärken konfundiert sein können (McFarland & Crouch, 2002; Mierke & Klauer, 2003). Ein Beispiel hierfür wäre eine altersbedingte Verringerung der Fähigkeit zum Aufgabenwechsel in einem IAT zu Vorurteilen gegenüber alten Menschen (siehe hierzu Kray & Lindenberger, 2000), die nach der klassischen Interpretation des IAT als eine Verstärkung von impliziten Vorurteilen gegenüber alten Menschen mit steigendem Alter missinterpretiert werden würde. Darüber hinaus lassen sich durch perzeptuelle Stimulusmerkmale Aufgabenwechselkosten – und entsprechend IAT-Scores – artifiziell manipulieren, auch wenn a priori angenommen werden kann, dass für das gewählte Stimulusmaterial keine Assoziationen vorliegen (z. B. Mierke & Klauer, 2003).

Ein ähnlicher Einwand gegen den IAT stammt von Brendl, Markman und Messner (2001). Aufbauend auf eine Konzeptualisierung der Reaktionen im IAT als Random-Walk-Prozess (Townsend & Ashby, 1983) argumentieren diese Autoren, dass Reaktionszeitunterschiede zwischen der kompatiblen und inkompatiblen Kombination häufig auf einer bewussten, intentionalen Anpassung von Reaktionsschwellen beruht. So erhöhen Versuchspersonen meist ihre allgemeine Reaktionsschwelle in der inkompatiblen Kombination, während die Reaktionsschwelle in der kompatiblen Kombination oft geringer angesetzt wird. Eine solche Anpassung von Reaktionsschwellen muss jedoch nicht unbedingt auf automatischen Prozessen der Aktivierungsausbreitung in einem assoziativen Netzwerk beruhen. Vielmehr kann eine Anpassung von Reaktionsschwellen intentional oder bewusst erfolgen, was gegen die gängige Annahme spricht, dass der IAT die Stärke automatisch aktivierter Assoziationen misst.

Ein weiterer Einwand gegen die interne Validität des IAT stammt von Rothermund und Wentura (2001, 2004). Diese Autoren konnten zeigen, dass eine entscheidende Varianzquelle im IAT auf Figur-Grund-Kompatibilitäten beruht. Demnach bildet in den meisten IATs jeweils eine Kategorie der beiden vorgegebenen Dimensionen die Figur, während die jeweils andere Kategorie den Grund bildet. Erleichterte Reaktionen kommen nach Rothermund und Wentura vor allem dadurch zustande, dass auf die beiden „Figuren“ mit der gleichen Taste reagiert werden muss, während erschwerte Reaktionen darauf beruhen, dass auf die beiden „Figuren“ mit unterschiedlichen Tasten reagiert werden muss. Problematisch ist dabei, dass solche Figur-Grund-Kompatibilitäten nicht notwendigerweise auf Assoziationen beruhen müssen. Vielmehr lassen sich Figur-Grund-Effekte auch dann im IAT beobachten, wenn gar keine Assoziationen vorliegen, sondern diese allein durch Salienzunterschiede zustande kommen (Rothermund & Wentura, 2001). Zudem tendieren sowohl negativ bewertete als auch wenig vertraute Objekte dazu, die Figur zu bilden (Johnston, Hawley, Plewe, Elliott & De Witt, 1990; Pratto & John, 1991), so dass ein Einfluss der Valenz im IAT häufig mit simplen Vertrautheitseffekten konfundiert ist (Rothermund & Wentura, 2004).

## Fazit

Ziel des vorliegenden Artikels war es, einen systematischen Überblick über die bisherige Forschung zum IAT zu geben, um auf diese Weise sowohl die Reichweite als auch die Grenzen des Verfahrens aufzuzeigen. Aufbauend auf diesen Überblick lassen sich die folgenden Schlussfolgerungen ziehen:

1. Auch wenn die Zusammenhänge zwischen IAT und korrespondierenden Fragebogenmaßen oft nur gering ausfallen und sehr stark variieren, zeigen sich aus einer meta-analytischen Perspektive (Hofmann et al., 2003) eindeutig die zu erwartenden positiven Zusammenhänge (konvergente Validität). Dabei muss allerdings betont werden, dass Korrelationen zwischen IAT und korrespondierenden Fragebogenmaßen sowohl von personenbezogenen als auch von objektbezogenen Faktoren abhängen. Darüber hinaus zeigen sich deutlich höhere Zusammenhänge zwischen IAT und Fragebogenmaßen, wenn der Einfluss von Fehlervarianz kontrolliert wird. Mit nicht korrespondierenden Fragebogenmaßen zeigt der IAT wie zu erwarten keine Korrelationen (diskriminante Validität).

2. Die bisher vorliegende Evidenz zeigt relativ deutlich, dass der IAT systematische Konstruktvarianz zu messen vermag, mit der Verhalten und Informations-Verarbeitungs-Prozesse vorhergesagt werden können (prädiktive Validität). Besonders hervorzuheben ist dabei, dass der IAT vor allem solche Aspekte vorherzusagen vermag, die mit herkömmlichen Fragebogenverfahren nur schwer erfasst werden können (inkrementelle Validität). Beispiele hierfür sind spontane Verhaltensmuster (z. B. nonverbales Verhalten) oder verzerrte Interpretationen mehrdeutiger Informationen.

3. Im Gegensatz zur ursprünglichen Interpretation als reines Dispositionsmaß zeigt der IAT eine hohe Sensitivität gegenüber Kontextmanipulationen. Obwohl eine bewusste Verfälschung von IAT-Scores weitaus schwieriger zu sein scheint als eine Verfälschung von expliziten Selbstberichten, zeigen sich Kontexteinflüsse bereits bei relativ simplen Variationen. Diese Ergebnisse sprechen dafür, dass der IAT nicht nur chronische Assoziationen, sondern auch situativ aktivierte Assoziationen misst. Diese Schlussfolgerung wird auch durch Ergebnisse zur Test-Retest-Reliabilität des IAT gestützt, die für ein reines Dispositionsmaß nur relativ unbefriedigend sind. Dass der IAT nicht ausschließlich situativ aktivierte, sondern tatsächlich auch chronische Assoziationen misst, kann aus Studien abgeleitet werden, in denen Verhalten über einen längeren Zeitraum erfolgreich vorhergesagt werden konnte.

4. Entgegen der Annahme, dass IAT-Scores allein von Kategorieassoziationen abhängen, besitzt das Stimulusmaterial einen entscheidenden Einfluss auf IAT-Mittelwerte. Eine absolute Interpretation von IAT-Scores erscheint aus diesem Grund nicht zulässig. Darüber hinaus legen die Ergebnisse zu Stimuluseinflüssen nahe, dass der

IAT nicht als „Test“ im Sinne eines standardisierten Verfahrens zu betrachten ist, sondern eher als eine Messmethode. Um entsprechende Missverständnisse zu vermeiden, sollte der *Implicit Association Test* daher vielleicht besser in *Implicit Association Task* umbenannt werden. Darüber hinaus impliziert die Materialabhängigkeit des IAT, dass Vergleiche der Mittelwerte thematisch unterschiedlicher IATs (z. B. Rudman, Feinberg & Fairchild, 2002) grundsätzlich nicht zulässig sind. In diesem Zusammenhang ist allerdings zu beachten, dass materialbedingte Mittelwertsunterschiede auf Stichprobenebene nicht unbedingt eine veränderte Rangordnung auf differenzieller Ebene implizieren. Insofern besitzen derartige Mittelwertsunterschiede noch keine Aussagekraft über die Vorhersagekraft der in Frage stehenden IATs. Dies sollte in zukünftigen Studien geklärt werden.

5. Auch wenn der IAT nach den vorliegenden Ergebnissen tatsächlich systematische Konstruktvarianz zu messen vermag, scheint ein substanzieller Varianzanteil im IAT auf systematischer Fehlervarianz zu beruhen (interne Validität). Dies hat zum einen die Implikation, dass diagnostische Umkehrschlüsse von IAT-Scores auf zugrundeliegende Dispositionen (z. B. Gray, McCulloch, Smith, Morris & Snowden, 2003) auf jeden Fall vermieden werden sollten. Zum anderen sollten potenzielle Alternativklärungen diskutiert und (wenn möglich) ausgeschlossen werden. Dies trifft für den IAT sowohl als abhängige als auch als unabhängige Variable zu. Eine weitere Implikation ist, dass Korrelationen zwischen inhaltlich verschiedenen IATs nicht notwendigerweise auf geteilter Konstruktvarianz beruhen, sondern auch durch einen hohen Anteil geteilter Fehlervarianz zustande kommen können.

## Ausblick

Zusammenfassend kann festgehalten werden, dass die bisherige Forschung zum IAT positive Evidenz für dessen konvergente, diskriminante, prädiktive und inkrementelle Validität zeigt. Die interne Validität des IAT erscheint allerdings nach wie vor kritisch. Auch wenn es mittlerweile eine Reihe von Ansätzen gibt, die sowohl systematische Konstruktvarianz als auch systematische Fehlervarianz integrieren (z. B. De Houwer, 2003; Rothermund & Wentura, 2004), sind die dem IAT zugrundeliegenden kognitiven Prozesse bis heute nur unzureichend geklärt. Entsprechend sollten diese Prozesse in zukünftigen Studien noch genauer erforscht werden. Nach unserer Auffassung ist dabei von zentraler Bedeutung, die bereits identifizierten im IAT konfundierten Prozesse voneinander zu trennen und (wenn möglich) zu isolieren. Interessante Möglichkeiten bieten hier Prozess-Dissoziations-Verfahren der kognitiven Psychologie, mit denen der Beitrag qualitativ unterschiedlicher Prozesse innerhalb einer gegebenen Aufgabe mathematisch abgeschätzt werden können (z. B. Lindsay & Jacoby, 1994). So haben zum Beispiel Conrey, Sherman, Gawronski, Hugenberg und Groom (2004) kürzlich ein multinomiales Modell für Reaktions-Kompatibilitäts-Aufgaben entwickelt, das sich auch

auf den IAT anwenden lässt. Mit diesem Modell können vier qualitativ unterschiedliche, im IAT konfundierte Prozesse voneinander getrennt, und deren relativer Beitrag mathematisch abgeschätzt werden. Auch wenn sich diese Forschung noch im Anfangsstadium befindet, ist die Anwendung von Prozess-Dissoziations-Verfahren auf implizite Messmethoden nach unserer Auffassung eine vielversprechende Perspektive, mehr über die im IAT beteiligten Prozesse zu lernen, um auf diese Weise das Verfahren fruchtbarer – und gleichzeitig kritischer – einsetzen zu können.

## Literatur

- Allport, D. A., Styles, E. A. & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.
- Asendorpf, J. B., Banse, R. & Mücke, D. (2002). Double dissociation between explicit and implicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83*, 380–393.
- Ashburn-Nardo, L., Voils, C. I. & Monteith, M. J. (2001). Implicit associations as seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81*, 789–799.
- Banse, R. (1999). Automatic evaluation of self and significant others: Affective priming in close relationships. *Journal of Social and Personal Relationships, 16*, 803–821.
- Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition and Emotion, 15*, 501–520.
- Banse, R. & Fischer, I. (2002, Juli). *Implicit and explicit aggressiveness and the prediction of aggressive behavior*. Poster bei der 11<sup>th</sup> European Conference on Personality by the European Society for Personality Psychology, Jena.
- Banse, R. & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Skaleneigenschaften und Validierung. *Diagnostica, 49*, 4–13.
- Banse, R., Seise, J. & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48*, 145–160.
- Blair, I. V., Ma, J. & Lenton, A. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828–841.
- Blümke, M. & Friese, M. (2004). *Do irrelevant features of stimuli influence IAT effects?* Manuscript submitted for publication.
- Bosson, J. K., Swann, W. B., Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*, 631–643.
- Brendl, C. M., Markman, A. B. & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 760–773.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.
- Chee, M. L., Sriram, N., Soon, C. S. & Lee, K. M. (2000). Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *Neuroreport, 11*, 135–140.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K. & Groom, C. (2004). *Beyond automaticity and control: The Quad-Model of behavioral response*. Manuscript submitted for publication.
- Cunningham, W. A., Preacher, K. J. & Banaji, M. R. (2001). Implicit attitude measurement: Consistency, stability, and convergent validity. *Psychological Science, 12*, 163–170.
- Cunningham, W. A., Nezlek, J. B. & Banaji, M. R. (in press). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*.
- Dasgupta, N. & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*, 800–814.
- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443–451.
- De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Erlbaum.
- Devine, P. G. (2001). Implicit prejudice and stereotyping: How automatic are they? *Journal of Personality and Social Psychology, 81*, 757–759.
- Duncan, B. L. (1976). Differential perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology, 34*, 590–598.
- Dunning, D. & Sherman, D. A. (1997). Stereotypes and tacit inference. *Journal of Personality and Social Psychology, 73*, 459–471.
- 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. & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297–327.
- 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.
- Festinger, L. & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*, 203–210.
- Florack, A., Scarabis, M. & Bless (2001a). Der Einfluss wahrgenommener Bedrohung auf die Nutzung automatischer Assoziationen bei der Personenbeurteilung. *Zeitschrift für Sozialpsychologie, 32*, 249–260.
- Florack, A., Scarabis, M. & Bless, H. (2001b). When do associations matter? The use of automatic associations towards ethnic groups in person judgments. *Journal of Experimental Social Psychology, 37*, 518–524.
- Foroni, F. & Mayr, U. (in press). The power of a story: New, automatic associations from a single reading of a short scenario. *Psychonomic Bulletin and Review*.
- Gawronski, B. (2000). Falsifikationismus und Holismus in der experimentellen Psychologie: Logische Grundlagen und methodologische Konsequenzen. *Zeitschrift für Sozialpsychologie, 31*, 3–17.
- Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice related IATs. *Experimental Psychology, 49*, 171–180.
- Gawronski, B. & Bodenhausen, G. V. (2004). *Accessibility effects on implicit social cognition: The role of accessible content versus accessibility experiences*. Manuscript submitted for publication.
- Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J. & Klauer, K. C. (2003a). It's in the mind of the beholder: The impact of stereotypic associations on category-based and individuating impression formation. *Journal of Experimental Social Psychology, 39*, 16–30.

- Gawronski, B., Geschke, D. & Banse, R. (2003b). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology, 33*, 573–589.
- Gawronski, B. & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40*, 535–542.
- Govan, C. L. & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by redefining the category labels. *Journal of Experimental Social Psychology, 40*, 357–365.
- Gray, N. S., McCulloch, M. J., Smith, J., Morris, M. & Snowden, R. J. (2003). Violence viewed by psychopathic murderers: Adapting a revealing test may expose those psychopaths who are most likely to kill. *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., McGhee, D. E. & Schwartz, J. K. L. (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., Gschwendner, T., Le, H. & Schmitt, M. (2003). *A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures*. Manuscript submitted for publication.
- Hofmann, W., Gschwendner, T. & Schmitt, M. (2003). *On the moderation of the implicit-explicit consistency: The role of introspection and adjustment*. Manuscript submitted for publication.
- Hugenberg, K. & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*, 640–643.
- Hugenberg, K. & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in racial categorization. *Psychological Science, 15*, 342–345.
- Hummert, M. L., Gartska, 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 & Aging, 17*, 482–495.
- Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119*, 397–411.
- Karpinski, A. & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 774–788.
- Kim, D. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66*, 83–96.
- Klauer, K. C. & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology, 75*, 1155–1178.
- Kray, J. & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology & Aging, 15*, 126–147.
- Kühnen, U., Schießl, M., Bauer, N., Paulig, N., Pöhlmann, C. & Schmidhals, K. (2001). How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. *Zeitschrift für Experimentelle Psychologie, 48*, 135–144.
- Kunda, Z. & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin, 19*, 90–99.
- Lindsay, D. S. & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception & Performance, 20*, 219–234.
- Lowery, B. S., Hardin, C. D. & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*, 842–855.
- Maison, D., Greenwald, A. G. & Bruin, R. (2001). The Implicit Association Test as a measure of implicit consumer attitudes. *Polish Psychological Bulletin, 32*, 1–9.
- McConnell, A. R. & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435–442.
- McFarland, S. G. & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20*, 483–510.
- Mierke, J. & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift für Experimentelle Psychologie, 48*, 107–122.
- Mierke, J. & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85*, 1180–1192.
- Mitchell, J. P., Nosek, B. A. & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455–469.
- Nosek, B. A. (2004). *Moderators of the relationship between explicit and implicit attitudes*. Manuscript submitted for publication.
- Perugini, M. (in press). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C. & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729–738.
- Plessner, H. & Banse, R. (2001). Attitude measurement using the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 82–84.
- 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.
- Richeson, J. A. & Ambady, N. (2001). Who's in charge? Effects of situational roles on automatic gender bias. *Sex Roles, 44*, 493–512.
- Richeson, J. A. & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology, 39*, 177–183.
- Rogers, R. D. & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.
- Rothermund, K. & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 94–106.
- Rothermund, K. & Wentura, D. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General, 133*, 139–165.
- Rudman, L. A., Feinberg, J. & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition, 20*, 294–320.
- Rudman, L. A. & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 743–762.
- Rudman, L. A., Greenwald, A. G., Mellott, D. S. & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437–465.
- Sagar, H. A. & Schofield, J. W. (1980). Racial and behavioral cues in black and white children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology, 39*, 590–598.

- Schmukle, S. C. & Egloff, B. (in press). Does the Implicit Association Test for assessing anxiety measure state and trait variance? *European Journal of Personality*.
- Skowronski, J. J. & Lawrence, M. A. (2001). A comparative study of the implicit and explicit gender attitudes of children and college students. *Psychology of Women Quarterly*, 25, 155–165.
- Steffens, M. C. (in press). Is the Implicit Association Test immune to faking? *Experimental Psychology*.
- Steffens, M. C. & Buchner, A. (2003). Implicit Association Test: Separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology*, 50, 33–48.
- 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.
- Taylor, S. E., Fiske, S. T., Etcoff, N. L. & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, 36, 778–793.
- 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.
- Teige, S., Schnabel, K., Banse, R. & Asendorpf, J. B. (in press). Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task (EAST). *European Journal of Personality*.
- Townsend, J. T. & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. New York. Cambridge University 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. D., Lindsey, S. & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101–126.
- Wittenbrink, B. Judd, C. M. & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, 81, 815–827.

---

Dr. Bertram Gawronski

Department of Psychology  
Northwestern University  
2029 Sheridan Road  
Evanston, IL 60208-2710  
USA  
E-Mail: [gawronski@northwestern.edu](mailto:gawronski@northwestern.edu)

---

Dr. Frederica R. Conrey

Department of Psychology  
Northwestern University  
2029 Sheridan Road  
Evanston, IL 60208-2710  
USA  
E-Mail: [f-conrey@northwestern.edu](mailto:f-conrey@northwestern.edu)