

Kognitive Prozesse bei der
indirekten Messung von Einstellungen
mit dem Impliziten Assoziationstest

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Wirtschafts- und Verhaltenswissenschaftlichen Fakultät
der Albert-Ludwigs-Universität
Freiburg i. Br.

vorgelegt von

Jan Mierke
aus Kassel

SS 2004

Erstgutachter:	Prof. Dr. Karl Christoph Klauer
Zweitgutachter:	Prof. Dr. Klaus Fiedler
Drittgutachter:	Prof. Dr. Gerhard Strube
Vorsitzender des Promotionsausschusses der gemeinsamen Kommission der Philologischen, Philosophischen, Wirtschafts- und Verhaltenswissenschaftlichen Fakultät:	Prof. Dr. Hermann Schwengel
Datum der Fachprüfung im Promotionsfach:	22.12.2004

Inhaltsverzeichnis

1	Zusammenfassung.....	7
2	Einleitung.....	9
3	Der Implizite Assoziationstest.....	14
3.1	<i>Befunde zur Validität des IAT.....</i>	<i>19</i>
3.2	<i>Varianten des IAT.....</i>	<i>31</i>
3.3	<i>Allgemeine Probleme impliziter Verfahren.....</i>	<i>33</i>
3.4	<i>Theoretische Modelle.....</i>	<i>35</i>
3.5	<i>Das Aufgabenwechselmodell.....</i>	<i>47</i>
4	Experimente zu lokalen Sequenzeffekten.....	58
4.1	<i>Experiment 1.....</i>	<i>58</i>
4.2	<i>Experiment 2.....</i>	<i>69</i>
4.3	<i>Experiment 3.....</i>	<i>76</i>
5	Experimente zu globalen Effekten.....	82
5.1	<i>Experiment 4.....</i>	<i>82</i>
5.2	<i>Experiment 5.....</i>	<i>91</i>
5.3	<i>Experiment 6.....</i>	<i>97</i>
5.4	<i>Zusammenfassung der Befunde zu globalen Effekten.....</i>	<i>101</i>
6	Experimente zur Rolle exekutiver Kontrollprozesse.....	104
6.1	<i>Experiment 7.....</i>	<i>105</i>
6.2	<i>Experiment 8.....</i>	<i>111</i>
7	Experimente zu methodenspezifischer Varianz.....	120
7.1	<i>Experiment 9.....</i>	<i>124</i>
7.2	<i>Experiment 10.....</i>	<i>127</i>
7.3	<i>Experiment 11.....</i>	<i>130</i>
7.4	<i>Experiment 12.....</i>	<i>136</i>
7.5	<i>Techniken zur Reduktion methodenspezifischer Varianz.....</i>	<i>145</i>

8	Gesamtdiskussion	151
8.1	<i>Bewertung der Befunde in Bezug auf das Aufgabenwechselmodell</i>	152
8.2	<i>Bewertung der Befunde in Bezug auf alternative Prozessmodelle des IAT</i>	159
8.3	<i>Praktische Konsequenzen</i>	163
9	Literatur	170
	Anhang A: Material des Blumen-Insekten Einstellungs-IAT in Experiment 1	185
	Anhang B: Zusätzliches Material des Blumen-Insekten Einstellungs-IAT in den folgenden Experimenten	185
	Anhang C: Material des Selbst-Andere Extraversions-IAT	186

Veröffentlichungsnachweis

Die dieser Arbeit zugrunde liegenden Experimente sind in Teilen bereits in psychologischen Fachzeitschriften erschienen. Die Experimente 1 und 7 sind in Mierke und Klauer (2001) veröffentlicht, die Experimente 9 bis 12 sowie Experiment 2 in Mierke und Klauer (2003).

Danksagung

An dieser Stelle möchte ich mich ganz herzlich bei allen bedanken, die mich bei der Erstellung dieser Arbeit unterstützt haben.

Mein Dank gilt an erster Stelle meinem Doktorvater Christoph Klauer, der durch seine konsequente Art zu denken sowie durch zahllose methodische und inhaltliche Vorschläge entscheidend zum Gelingen der Arbeit beigetragen hat. Vielen Dank für Ihr Vertrauen, die fruchtbare und gute Zusammenarbeit, die erstklassigen Arbeitsbedingungen und die Zeit, die Sie immer gefunden haben, um alle Probleme und Fragen zu diskutieren.

Dank gebührt auch meinen Kollegen Andreas Eder, Thorsten Meiser, Katja Mierke, Jochen Musch, Christoph Stahl und Zengmei Zhao für die freundschaftliche Arbeitsatmosphäre und viele fruchtbare Diskussionen, insbesondere auch Christoph Stahl, der sich bereit erklärt hat, die Arbeit Korrektur zu lesen.

Danken möchte ich auch Jens Asendorpf, Rainer Banse, Matthias Blümke, Miguel Brendl, Jan De Houwer, Roland Deutsch, Boris Egloff, Klaus Fiedler, Bertram Gawronski, Tony Greenwald, Claude Messner, Roland Neumann, Henning Plessner, Klaus Rothermund, Stefan Schmukle, Konrad Schnabel, Melanie Steffens, Dirk Wentura und zahllosen anderen, deren konstruktive Anmerkungen und kritische Fragen ebenso wertvoll wie inspirierend waren. Ich danke Euch für die positive Stimmung, die den wissenschaftlichen Gedankenaustausch zu einem spannenden und schönen Erlebnis hat werden lassen.

Dank gebührt auch Andrea Epting, Pia Freiermuth, Ralf Grosche, Jörn Güttgemanns, Wera Hatje, Roland Imhoff, Anne Cathrin Kienle, Gabriela Munoz, Gabi Romünder, Anke Simon, Vasilios Siochos und Annette Wille, die tatkräftig und effizient bei der Erhebung der Daten geholfen haben.

Ich danke meinen Eltern, die mir das Studium ermöglicht haben und mich immer mit Rat und Tat unterstützt haben.

Besonderen Dank schulde ich meiner Frau und Freundin Katja Mierke. Deine wertvollen inhaltlichen Kommentare und Anregungen, deine ausdauernde Unterstützung und nicht zuletzt deine Zuneigung haben einen unschätzbaren Anteil am Gelingen dieser Arbeit.

Unschätzbar ist auch der Beitrag meiner Tochter Mona. Ich freue mich jeden Tag aufs Neue darüber, dass du da bist.

1 Zusammenfassung

Der implizite Assoziationstest (IAT; A. G. Greenwald, D. E. McGhee & J. L. K. Schwartz, 1998) ist ein reaktionszeitbasiertes indirektes Verfahren zur Messung der Stärke assoziativer Verknüpfungen. Die theoretische Relevanz des IAT ergibt sich aus der gegenwärtig verbreiteten Strömung, klassische psychologische Phänomene auf der Ebene assoziativer Verknüpfungen zu reformulieren und zu untersuchen. Trotz des weitgehenden Fehlens einer theoretischen Grundlage wird der IAT zunehmend auf konkrete Fragestellungen in vielen experimentellen aber auch differentialpsychologischen und diagnostischen Kontexten angewandt.

Der IAT basiert auf der gemischten Bearbeitung zweier Diskriminationsaufgaben; einer Attribut- und einer Zielkonzeptaufgabe. Jeweils zwei Antwortkategorien dieser Diskriminationsaufgaben werden einer gemeinsamen Antworttaste zugeordnet. Die zentrale abhängige Variable ist der IAT-Effekt, die Reaktionszeitsdifferenz zwischen einer kompatiblen Bedingung, in der assoziierte Konzepte jeweils der gleichen Antworttaste zugeordnet sind, und einer inkompatiblen Bedingung, in der assoziierte Konzepte jeweils verschiedenen Antworttasten zugeordnet sind.

In der vorliegenden Arbeit wird ein Modell der kognitiven Prozesse, die bei der Bearbeitung eines IAT ablaufen, vorgeschlagen und geprüft. Dessen zentrale Annahme lautet, dass der IAT-Effekt auf Auswirkungen exekutiver Kontrollprozesse zurückgeht, die beim Wechsel zwischen Attribut- und Zielkonzeptaufgabe ablaufen. Aus allgemeinspsychologischen Befunden und Theorien zu solchen Aufgabenwechselprozessen werden einige spezifische Hypothesen über ihre Auswirkung auf den IAT abgeleitet und experimentell überprüft.

In den Experimenten 1 bis 3 werden Vorhersagen dieses Modells bezüglich lokaler Sequenzeffekte untersucht. Das zentrale Ergebnis dieser Experimente ist, dass Kompatibilitätseffekte im IAT verstärkt auftreten, wenn in einem Durchgang im Verhältnis zum direkten Vorgängerdurchgang zwischen der Attribut- und der Zielkonzeptaufgabe gewechselt werden muss (Experimente 1 bis 3), und dass dieser Effekt geringer wird, je länger das Intervall zwischen den Durchgängen ausfällt (Experiment 1). Diese Effekte werden durch kurzfristige Nachwirkungen vorheriger Aufgabenbearbeitung erklärt, die auf eine Trägheit von Aufgabenwechselprozessen zurückgehen (D. A. Allport, E. A. Styles & S. Hsieh, 1994).

In den Experimenten 4 bis 6 werden längerfristige Nachwirkungen der Bearbeitung von IAT-Blöcken geprüft. Die Bearbeitung von inkompatiblen IAT-Blöcken hat sehr spezifische längerfristige Nachwirkungen auf eine nachfolgende Bewertungsaufgabe, die über einen Wechsel der Antwortmodalität hinweg auftreten und mutmaßlich auf einen erschwerten Abruf von Bewertungsinformation zurückgehen.

In den Experimenten 7 und 8 wird die Beteiligung exekutiver Kontrollprozesse am Zustandekommen von IAT-Effekten untersucht. Die Präsentation so genannter Aufgabenhinweise (z. B. N. Meiran, 1996) verringert in diesen Experimenten den IAT-Effekt, nicht jedoch spezifische Aufgabenwechselkosten. Es wird argumentiert, dass Aufgabenhinweise die Identifikation der zu bearbeitenden Aufgabe erleichtern, nicht jedoch die in den Vorgängerexperimenten aufgetreten kurz- und längerfristigen Nachwirkungseffekte reduzieren können.

In den Experimenten 9 bis 12 wird schließlich geprüft, welche Konsequenzen die Beteiligung von Aufgabenwechselprozessen für die psychometrische Nutzung des IAT hat. Die Ergebnisse zeigen eine reliable Kontaminierung des IAT mit methodenspezifischer Varianz: Es treten selbst dann statistisch bedeutsame Korrelationen zwischen IATs auf, wenn diese keine gemeinsame Inhaltvarianz aufweisen. Dieser Befund kann eine Reihe widersprüchlicher Ergebnisse zur psychometrischen Qualität des IAT erklären.

Die Arbeit schließt mit einer Bewertung alternativer Prozessmodelle des IAT im Lichte der berichteten Befunde und einer Diskussion praktischer Konsequenzen der Ergebnisse für die Nutzung des Verfahrens in Anwendungskontexten. Gleichwohl der IAT als flexibles Werkzeug zur Erfassung assoziativer Verknüpfungen einen fruchtbaren Beitrag leisten kann, bleiben viele Fragen in Bezug auf die interne, konvergente und diskriminante Validität des IAT offen. Eine ausschließlich auf den IAT gestützte Forschungsstrategie scheint in diesem Sinne unbrauchbar. Erst in der Kombination mit andersartigen Verfahren wird ein inhaltlicher Erkenntnisgewinn greifbar.

2 Einleitung

„Nicht wir leben es; Es lebt uns...“, Lars Gustafsson

Die Vorstellung, menschliches Verhalten und Erleben werde nicht von bewussten, intentionalen kognitiven Prozessen gesteuert, sondern durch eine jenseits des Bewusstseins wirkende Mechanik, zieht sich durch die unterschiedlichsten psychologischen Strömungen der Vergangenheit und Gegenwart. Sie findet sich - naturgemäß - sowohl in der Psychoanalyse als auch in radikal-behavioristischen Ansätzen (z.B. Skinner, 1974). Auch in der jüngeren kognitiven Psychologie gibt es ein reges Interesse an unbewussten und spontanen Prozessen (z.B. Bargh, 1994, 1997; Bargh, Chaiken, Raymond & Hymes, 1996; Draine & Greenwald, 1998; Fazio, Sanbonmatsu, Powell & Kardes, 1986; Greenwald, Draine & Abrams, 1996; Klotz & Neumann, 1999; Neumann & Klotz, 1994). Dieses Interesse erklärt sich schon daraus, dass die Formulierung allgemein gültiger wissenschaftlicher Sätze eine Determiniertheit voraussetzt, die mit der Annahme bewussten Handelns und eines freien Willens nur schwer vereinbar ist.

Aus dieser Perspektive scheinen die bewussten Korrelate kognitiver oder affektiver Vorgänge kein zuverlässiger Indikator der ihnen zugrunde liegenden Prozesse zu sein (z. B. Bredenkamp, Klein, von Hayn & Vaterrodt, 1988). Nicht nur deshalb gibt es eine lange Tradition innerhalb der Psychologie, den Blick in die Black-Box mit Verfahren zu versuchen, die von der introspektiven Zugänglichkeit der interessierenden Inhalte unabhängig sind. Der Thematische Apperzeptions-Test (Murray, 1943) und der Rohrschachttest (z.B. Hoek-Gradenwith, 1959; Huth, 1949) sind Beispiele für solche Verfahren, und auch die von Freud eingesetzte Technik der freien Assoziation kann als Versuch gelten, Erleben und Fühlen jenseits der Introspektion zu erfassen.

Trotzdem ist die bei weitem gängigste Methode zur Erlangung von Informationen über mentale Vorgänge die - mehr oder weniger standardisierte - direkte Befragung, nicht zuletzt wohl deshalb, weil sie eine bequeme und nahe liegende Methode ist. Die große Anzahl von Verfahren, die auf eine explizite Befragung der Teilnehmer (im Gegensatz zu einer indirekten oder impliziten Messung) zurückgreift, soll im Folgenden mit dem Oberbegriff explizite Verfahren bezeichnet werden. Seit den einflussreichen Arbeiten von Louis Leon Thurstone (1928) spielt die Nutzung expliziter Verfahren in Form von Selbstberichtsmaßen eine zentrale

Rolle für die sozialpsychologische Erforschung von Einstellungen. Auch in der differentiellen und Persönlichkeitspsychologie stellen standardisierte Selbst- oder Fremdb Berichte, etwa in Form von Fragebögen, gegenwärtig die vorherrschende Methode der Datenerhebung dar.

Probleme expliziter Verfahren

Unglücklicherweise sind die Antworten auf Fragen in klassischen Selbstberichtsmaßen vielfach mit verzerrenden Faktoren konfundiert, die die Gültigkeit der Messungen wesentlich einschränken. Diese Verzerrungen lassen sich in zwei Klassen einteilen: Erstens solche, die auf Begrenzungen der Introspektionsfähigkeit zurückgehen, also einen Mangel an Fähigkeit zur Generierung der gewünschten Information, und zweitens solche, die auf einen Mangel an Bereitschaft zu wahrheitsgemäßen Auskünften zurückgehen (Egloff & Schmukle, 2002; Greenwald et al., 2002), also eher motivationaler Natur sind. In der Literatur wurde eine große Zahl spezifischer Verzerrungen diskutiert, die sich in dieses Raster einordnen lassen. Faktoren, die die *Fähigkeit* zu Introspektion und Selbstbeschreibung verringern können, sind zum Beispiel Selbsttäuschungen (Gur & Sackeim, 1979) und Selbsterweiterungstendenzen (Greenwald, 1980; Taylor & Brown, 1988). Faktoren, die eher die *Motivation* zu korrekter Selbstbeschreibung betreffen, sind allgemeine Selbstdarstellungstendenzen (Tedeschi, Schlenker & Bonoma, 1971), Demand-Characteristics (Orne, 1962), die Angst vor Bewertungen (Rosenberg, 1965), die Tendenz zu sozialer Erwünschtheit, oder die gezielte Verfälschung (Viswesvaran & Ones, 1999). Techniken zur Reduzierung motivational begründeter Verzerrungen, wie der Einsatz von Lügenskalen (Crowne & Marlowe, 1960; Paulhus, 1998) oder die Einführung von Zeitdruck bei der Beantwortung der Items (Holden, Wood & Tomashewski, 2001), sind bei der Lösung dieser Probleme nur begrenzt erfolgreich. Im Speziellen die gezielte Verfälschung von Testergebnissen stellt ein kaum kontrollierbares Problem dar, insbesondere dann, wenn die zu Testenden über Ziel und Funktionsweise des Messinstruments informiert sind.

Andere Ergebnisse zeigen wiederum starke Effekte der formalen Gestaltung von Fragebögen, also Effekte von scheinbar belanglosen Eigenschaften der Erhebungsinstrumente. So können Änderungen der Nummerierung von Antwortskalen (Strack, 1994) oder der Reihenfolge, in der Fragen vorgegeben werden, erhebliche Effekte auf die resultierenden Antworten haben (Nebel, Strack & Schwarz, 1989). Insgesamt muss also festgehalten werden, dass die Verwendung von Selbstberichtsmaßen in verschiedener Hinsicht problematisch ist. Erstens kann eine mangelnde Motivation zu wahrheitsgemäßen

Auskünften die Angaben in verschiedener Weise verzerren, zweitens kann die Generierung der Information durch Beschränkungen der Introspektion oder systematischen Selbst- und Fremdbeobachtung verzerrt sein, und drittens kann die Interpretation von vorgegebenen Fragen und Antwortformaten zu systematischen Verzerrungen führen. Eine problematische Eigenschaft solcher systematischer Verzerrungen ist, dass sie häufig wiederholte Messungen in gleicher Weise verzerren und somit einerseits nur schwer von inhaltlich begründeter, „wahrer“ Varianz getrennt werden können und andererseits zu einer Überschätzung der Messgenauigkeit führen.

Ein Teil der angedeuteten Schwierigkeiten liegt darin begründet, dass Verfahren, die auf Selbst- oder Fremdbereichten basieren, die Aufgabe der Informationsgenerierung in die Hand der Betroffenen selbst legen. Wie eine Antwort auf die Frage nach einer Einstellung im Einzelnen zustande kommt, liegt damit weitgehend außerhalb des Einflussbereichs der Forschenden. Im Hinblick auf diese - hier nur kurz umrissenen - Schwierigkeiten wäre es offensichtlich wünschenswert Verfahren zu entwickeln, die die genannten Probleme umgehen.

Implizite Verfahren

Zur Lösung der Probleme expliziter Verfahren wurden viele einfallsreiche neue Prozeduren entwickelt, um Prozesse und Inhalte sozialer Kognitionen zu erfassen. Die meisten dieser Prozeduren beruhen auf der Analyse von Leistungsdaten, meist in Form von Antwortlatenzen (z.B. Bosson, Swann & Pennebaker, 2000; Brauer, Wasel & Niedenthal, 2000; Fazio, 2001; Fazio & Olson, 2003) oder der Gedächtnisleistung (Klauer & Wegener, 1998; Taylor, Fiske, Etoff & Ruderman, 1978). Als Oberbegriff für diese Verfahrensklasse hat sich der Begriff implizite Verfahren durchgesetzt. Obwohl teilweise zwischen impliziten und indirekten Verfahren unterschieden wird, sollen die Begriffe im Kontext dieser Arbeit gleichbedeutend gebraucht werden. Das zentrale Kennzeichen impliziter oder indirekter Verfahren ist, dass die Versuchsteilnehmer nicht explizit gebeten werden Informationen zu einer Fragestellung abzugeben. Vielmehr wird die gewünschte Information durch einen Inferenzprozess aus Leistungsdaten gewonnen. Diese Inferenz setzt häufig eine Reihe von Annahmen voraus, etwa, dass kognitive Repräsentationen in einem assoziativen Netzwerk organisiert sind (z. B. Collins & Loftus, 1975; Collins & Quillian, 1969; Greenwald et al., 2002).

Eine zentrale Eigenschaft impliziter Verfahren ist, dass sowohl der Zweck der Messung, als auch der Mechanismus des Inferenzprozesses, der zwischen dem offenen Verhalten der Probanden und einer psychologischen Interpretation dieses Verhaltens vermittelt, verschleiert werden können (Brauer et al., 2000). In dieser Hinsicht stehen aktuelle implizite Verfahren in einer Verwandtschaftsbeziehung mit älteren Verfahren, wie dem TAT (Murray, 1943), die ebenfalls auf einer indirekten Form der Gewinnung psychologischer Daten basieren. Die Tatsache, dass die Kluft zwischen dem offenen Verhalten der Probanden bei der Bearbeitung aktueller impliziter Verfahren und der psychologischen Interpretation der Befunde vergleichsweise groß ist, ist zugleich Vorteil und Nachteil dieser Verfahren. Zum einen erleichtert diese Kluft die Verschleierung des Untersuchungszwecks und verringert damit die Verfälschbarkeit der Verfahren. Zum anderen ist sie jedoch ebenso Ursache einer geringen Augenscheinvalidität und der starken Belastung des Inferenzprozesses mit psychologischen Annahmen.

Bei der Beurteilung impliziter Verfahren werden meist zwei grundlegende Vorteile gegenüber expliziten Maßen genannt. Zum einen seien implizite Maße durch die Verschleierung des Erhebungsmechanismus gegen eine absichtsvolle Verfälschung der Ergebnisse im Sinne der oben genannten Verfälschungstendenzen immun (z.B. Banaji, 2001; Greenwald et al., 2002; Greenwald, McGhee & Schwartz, 1998). Zum anderen wird häufig davon ausgegangen, dass implizite Verfahren nicht - oder in deutlich geringerem Maße - von Einschränkungen der introspektiven Zugänglichkeit des zu messenden Konstrukts abhängen (z.B. Egloff & Schmukle, 2002; Greenwald et al., 2002). Die letztgenannte Position steht auch in Zusammenhang mit der Annahme, implizite und explizite Verfahren erfassen fundamental unterschiedliche Aspekte der zu messenden Konstrukte. So wird etwa angenommen, dass Personen Informationen über sich selbst und ihre Umwelt nicht nur in kognitiv kontrollierter und bewusster Weise verarbeiten, sondern auch implizit, das heißt, automatisch und ohne bewusstes Zutun (Fazio, 1990; Greenwald et al., 2002; Wilson, Lindsey & Schooler, 2000). Entsprechende kognitive beziehungsweise affektive Prozesse und Prozessprodukte könnten außerhalb der subjektiven Aufmerksamkeit liegen und damit introspektiv unzugänglich sein (Nisbett & Wilson, 1977).

Letztlich handelt es sich jedoch um eine empirische Frage, ob und inwieweit die oben angesprochenen Vorteile impliziter Verfahren tatsächlich zutreffen. Die gegenwärtig vorliegenden Ergebnisse lassen in dieser Hinsicht keinen eindeutigen Schluss zu. Die Qualifikation eines Verfahrens als implizit soll daher im Folgenden vor allem auf Eigenschaften des Verfahrens hinweisen. Ob sich die mit impliziten Verfahren erfassten Konstrukte

qualitativ von denjenigen unterscheiden, die durch explizite Verfahren erfasst werden, sei damit offen gelassen.

Affektives Priming

Ein Teil der gegenwärtig häufiger verwandten impliziten Verfahren basiert auf dem Paradigma des semantischen Primings (Meyer & Schvaneveldt, 1971; Neely, 1991), etwa das affektive Priming (Bargh, Chaiken, Govender & Pratto, 1992; Bargh et al., 1996; Draine & Greenwald, 1998; Fazio et al., 1986; Klauer, 1998; Klauer & Musch, 2001). In diesem experimentellen Paradigma soll erfasst werden, in welchem Ausmaß Einstellungen durch die Präsentation von Einstellungsobjekten automatisch aktiviert werden.

In dem Experiment von Fazio et al. (1986) bestand die Aufgabe der Versuchsteilnehmer darin, die Valenz eines dargebotenen Wortes, des so genannten Targets, durch Druck einer Antworttaste zu bewerten. Der zentrale Befund ist, dass die Klassifikation dann schneller abläuft, wenn dem zu klassifizierenden Wort (z.B. „warm“) ein assoziiertes Primewort vorangestellt ist (z.B. „Sonne“). Diese schnellere Verarbeitung der Targetwörter, denen ein evaluativ kongruentes Primewort vorangestellt ist, bezeichnet man als affektiven Primingeffekt. Das Auftreten eines Primingeffekts gilt als Beleg dafür, dass eine automatische Evaluation der Primewörter erfolgt ist. Im Verlauf der letzten Jahre ist eine Vielzahl von Varianten dieses Paradigmas vorgeschlagen und angewandt worden. Da eine Übersicht über diese Verfahren den Rahmen der vorliegenden Arbeit sprengen würde, seien hier nur einige Beispiele genannt. Werden zum Beispiel Primewörter verwandt, die eine soziale Gruppe repräsentieren, so kann mit diesem Paradigma die Stärke der Einstellung der Versuchsteilnehmer bezüglich dieser Gruppe erfasst werden (Fazio et al., 1986). Einige Studien zeigen in diesem Zusammenhang, dass durch die Stärke des Primingeffekts rassistisches Verhalten vorhergesagt werden kann (Dovidio, Kawakami, Johnson, Johnson & Howard, 1997; Fazio, Jackson, Dunton & Williams, 1995).

Obwohl die Klasse primingbasierter Verfahren ein durchaus brauchbares Paradigma zur Untersuchung von Mittelwerten und Gruppenunterschieden darstellen, sind die Effektgrößen ebenso wie die Reliabilitäten dieser Verfahren meist eher unbefriedigend (Banse, 2001; Bosson et al., 2000; Cunningham, Preacher & Banaji, 2001). Die niedrigere Reliabilität ist insgesamt ein entscheidender Nachteil impliziter im Vergleich zu expliziten Verfahren. Für explizite Selbstbeschreibungen und Persönlichkeitsfragebögen finden sich meist Reliabilitäten zwischen .70 und .90, Werte die deutlich über den Reliabilitäten primingbasierter Verfahren

liegen. Eine Interpretation der Effekte auf individueller Ebene, wie sie in vielen Anwendungskontexten wünschenswert erscheint, wäre also kaum zu verantworten (Greenwald & Banaji, 1995).

Der Affektive Simoneffekt

Ein anderes implizites Verfahren macht sich den so genannten Simoneffekt (Fitts & Seeger, 1953; Kornblum & Lee, 1995; Simon, 1990) zu nutze. Als Simoneffekt bezeichnet man den Befund, dass Antwortlatenzen bezüglich symbolischer Reize durch die Relation einer irrelevanten Reizposition mit der Position einer instruierten Antwort beeinflusst werden. Die Antwortlatenzen sind üblicherweise geringer, wenn auf einen links dargebotenen Reiz mit einer links positionierten Antworttaste reagiert werden soll, und relativ dazu verlangsamt, wenn auf einen links dargebotenen Reiz mit einer rechts positionierten Antworttaste reagiert werden soll.

Während der klassische Simoneffekt auf einer räumlichen Kongruenz von irrelevanter Reizposition und relevanter Antworttaste basiert, liegt dem verwandten affektiven Simoneffekt (De Houwer & Eelen, 1998; De Houwer, Hermans & Eelen, 1998) eine affektive Kongruenz zwischen der irrelevanten Evaluation eines Reizes und der evaluativen Konnotation der instruierten Antwort zugrunde. Antworten sind in diesem Paradigma dann schneller, wenn auf einen positiven Reiz mit einer positiv konnotierten Antwort, etwa dem Aussprechen des Wortes „positiv“ selbst, reagiert werden soll. Die Differenz zwischen affektiv kongruenten und inkongruenten Reiz-Reaktions-Paaren kann als Indikator der Evaluation der Reize interpretiert werden.

3 Der Implizite Assoziationstest

Eines der gegenwärtig populärsten impliziten Verfahren und Gegenstand der vorliegenden Arbeit ist der Implizite Assoziationstest (IAT, Greenwald et al., 1998). Der IAT basiert auf der Idee, dass es leichter fällt, mit der gleichen Antworttaste auf subjektiv assoziierte Konzepte zu reagieren, als mit einer entgegengesetzten Antworttaste. Werden die Versuchsteilnehmer zum Beispiel gebeten, in Abhängigkeit davon, ob ein präsentierter Reiz eine Blume oder ein Insekt beziehungsweise ein positives oder negatives Adjektiv darstellt, eine von zwei Antworttasten zu drücken, so sind die Antwortlatenzen üblicherweise geringer, wenn Blumen und positive Adjektive einer Taste zugeordnet sind (kompatible Bedingung) und langsamer, wenn Blumen und negative Adjektive einer Taste zugeordnet sind

(inkompatible Bedingung, vgl. Greenwald et al., 1998; Mierke & Klauer, 2001). Der IAT-Effekt ist definiert als Differenz der Reaktionszeiten in kompatibler und inkompatibler Bedingung und wird häufig als Schätzung der Stärke einer assoziativen Verknüpfung interpretiert (Banaji, 2001; Greenwald et al., 2002; Greenwald et al., 1998).

In den vergangenen fünf Jahren wurde der IAT in einer kaum überschaubaren Zahl unterschiedlicher Bereiche eingesetzt. In der Sozialpsychologie wurden IATs zur Messung von Vorurteilen (Cunningham et al., 2001; Greenwald et al., 1998), der Einstellung von Rauchern und Nichtrauchern (Sherman, Rose, Koch, Presson & Chassin, 2003; Swanson, Rudman & Greenwald, 2001), oder der Einstellung zur Benutzung von Kondomen (Marsh, Johnson & Scott Sheldon, 2001) eingesetzt. In der differentiellen Psychologie wurden IATs zur Messung der Schüchternheit (Asendorpf, Banse & Mücke, 2002) oder Ängstlichkeit (Egloff & Schmukle, 2002) verwendet, und in der klinischen Psychologie wurden IATs zur Erfassung von Phobien (De Jong, van den Hout, Rietbroek & Huijding, 2003; Teachman, Gregg & Woody, 2001) und Essstörungen (Roefs & Jansen, 2002) erprobt, um nur einige Beispiele zu nennen.

Die in dieser Vielfalt von Anwendungen sichtbare Popularität des IAT geht nicht zuletzt darauf zurück, dass Greenwald et al. (1998) explizit darauf hinweisen, dass der IAT auch in Bereichen angewandt werden kann, in denen Stärke und Richtung der assoziativen Verknüpfungen nicht a priori bekannt sind. Insbesondere die Überlegung, dass in diesem Fall die Richtung des Kompatibilitätseffekts als abhängige Variable Auskunft über *individuelle* Einstellungen bzw. Assoziationen geben könnte, hat zu einem regelrechten Boom in vielen Bereichen der Sozial- und differentiellen Psychologie geführt. Der Wandel der in einigen älteren Vorträgen von Greenwald und Kollegen auftauchenden Aufschlüsselung des Kürzels IAT als „Implicit Attitude *Task*“ zur mittlerweile geläufigen Aufschlüsselung „Implicit Association *Test*“ ist ein Ausdruck dieser Auffassung.

Prozedur des IAT

Die von Greenwald et al. (1998) beschriebene Standardprozedur des IAT basiert auf der alternierenden Bearbeitung zweier einfacher Kategorisierungsaufgaben, die im Folgenden als Attribut- und Zielkonzeptaufgabe bezeichnet werden sollen. Die Attributaufgabe erfordert die Klassifikation dargebotener Reize (zum Beispiel positiv und negativ besetzter Adjektive) anhand der Ausprägung eines spezifischen Attributs (etwa ihrer Valenz). Die Zielkonzeptaufgabe erfordert die Zuordnung dargebotener Reize zu einem von zwei Zielkonzepten, die

meist in Form disjunkter Kategorien vorgegeben sind, etwa Blumen und Insekten. Als Reizmaterial werden mehr oder weniger typische Exemplare der vier Antwortkategorien verwendet.

Die Attribut- und die Zielkonzeptaufgabe werden zunächst in zwei voneinander getrennten Übungsphasen eingeübt, die jeweils ein bis zwei Aufgabenblöcke umfassen. In der folgenden dritten Phase sollen beide Aufgaben innerhalb eines Aufgabenblocks gemischt bearbeitet werden. Jeweils zwei der vier Antwortkategorien werden dabei der gleichen Antworttaste zugeordnet. Erfolgt diese Zuordnung beispielsweise so, dass negativ besetzte Zielkonzepte und negative Attribute (etwa Insekten und andere negative Objekte) der einen Antworttaste und positiv besetzte Zielkonzepte und Attribute (etwa Blumen und andere positive Objekte) der anderen Antworttaste zugeordnet sind, so spricht man von einer kompatiblen Zuordnung. Erfolgt die Zuordnung hingegen so, dass negativ besetzte Zielkonzepte und negativ besetzte Attribute (ebenso wie positiv besetzte Zielkonzepte und Attribute) verschiedenen Tasten zugeordnet sind, so spricht man von einer inkompatiblen Zuordnung. Um im Verlauf des Experiments zwischen kompatibler und inkompatibler Tastenzuordnung zu wechseln, wird die Zuordnung der Zielkonzeptaufgabe gewechselt und in einem vierten Block gesondert eingeübt. In der fünften Phase werden dann beide Aufgaben mit der geänderten Tastenzuordnung gemischt bearbeitet. Abbildung 1 zeigt die Tastenzuordnung in der kompatiblen und inkompatiblen Phase eines IAT Blumen-Insekten Einstellungs-IAT.

Ein IAT umfasst also fünf getrennte Phasen. Meist werden zwei bis drei Blöcke der zwei gemischten Phasen bearbeitet, und jeweils ein bis zwei Blöcke der drei Übungsphasen, sodass bis zu zwölf Aufgabenblöcke resultieren. Die einzelnen Aufgabenblöcke umfassen meist zwischen 40 und 60 Durchgängen.

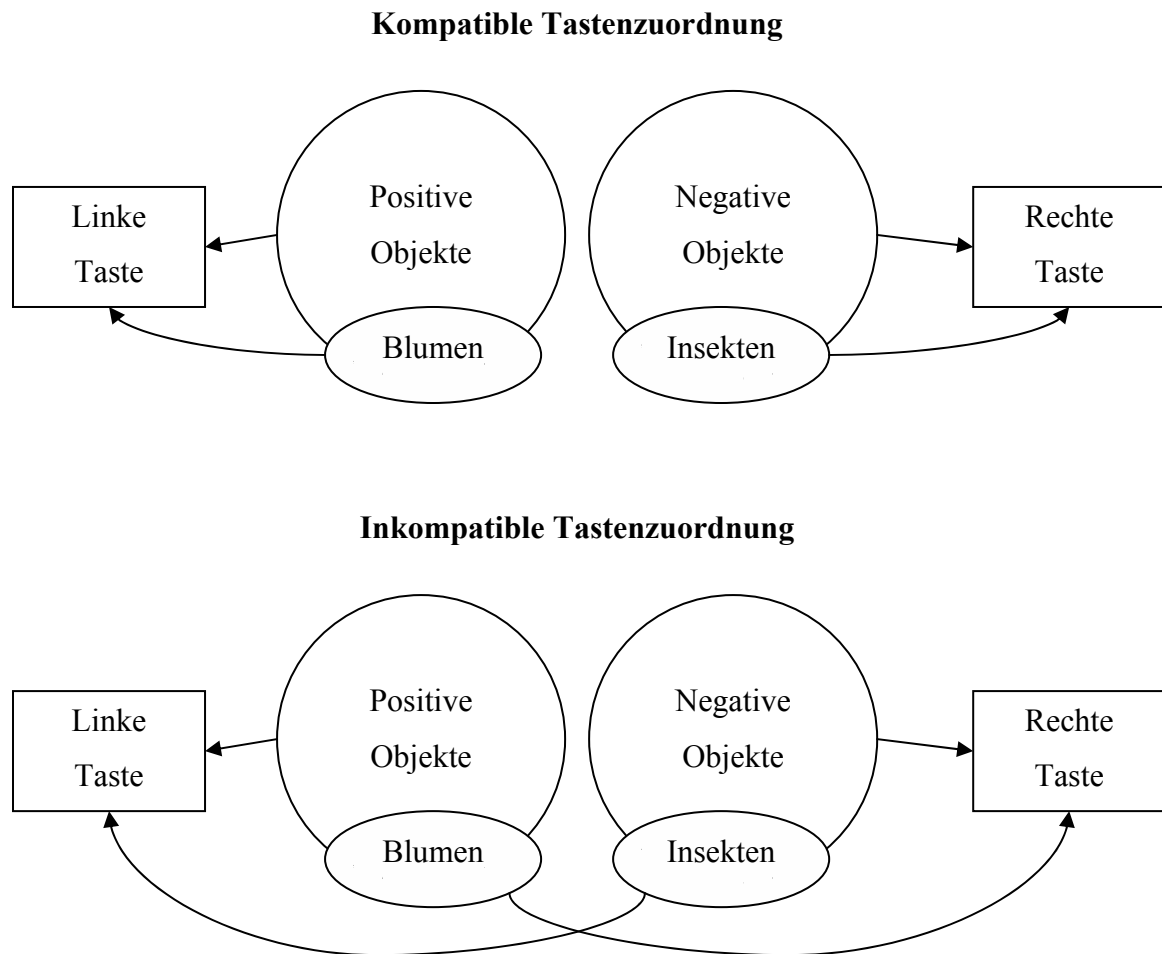


Abbildung 1: Tastenzuordnung in der kompatiblen und inkompatiblen Phase eines Blumen-Insekten Einstellungs-IAT. Bei kompatibler Zuordnung sind alle positiven Konzepte der einen Taste zugeordnet und alle negativen der anderen. Bei inkompatibler Tastenzuordnung sind beiden Antworttasten sowohl positive als auch negative Konzepte zugeordnet.

Der IAT-Effekt

Üblicherweise sind die Antwortlatenzen mit kompatibler Tastenzuordnung schneller als mit inkompatibler (z.B. Greenwald et al., 1998). Die Differenz der Antwortlatenzen in der inkompatiblen minus der kompatiblen Phase wird als IAT-Effekt bezeichnet. In vielen Fällen ist jedoch nicht a priori bekannt, wie einzelne Versuchsteilnehmer die Attribute den Zielkonzepten zuordnen; die Kompatibilität der zwei möglichen Tastenzuordnungen ist in diesem Fall eine abhängige Variable. Der IAT-Effekt wird dann einfach als Differenz zwischen den

Tastenzuordnungen definiert, wobei die Richtung des Effekts sich in dessen Vorzeichen ausdrückt.

Wie weiter oben schon angedeutet, wird der IAT-Effekt als Maß der Stärke assoziativer Verknüpfungen zwischen den Zielkonzepten und den Attributausprägungen interpretiert (Banaji, 2001; Greenwald et al., 2002; Greenwald et al., 1998). Es ist jedoch keineswegs eindeutig, welche Assoziationen dies sind. Mindestens vier unterscheidbare - im experimentellen Design jedoch konfundierte - Assoziationen können zum Auftreten eines IAT-Effekts führen. Im oben angesprochenen Beispiel mit Blumen und Insekten als Zielkonzepten sind dies a) die Assoziation von Blumen mit positiver und b) die Assoziation von Blumen mit negativer Bewertung sowie c) die Assoziation von Insekten mit positiver und d) die Assoziation von Insekten mit negativer Bewertung. Permutiert man die möglichen Assoziationsmuster durch und vergleicht die zu erwartenden Effekte, wird klar, dass eine ganze Reihe von Assoziationsmustern zu gleichen IAT-Effekten führen muss. So ist zum Beispiel anhand der reinen IAT-Effekte nicht unterscheidbar, ob eine Assoziation von Blumen und positiver Bewertung, eine Assoziation von Insekten und negativer Bewertung, oder beides zugleich vorliegt. Offenkundig können IAT-Effekte also nicht absolut interpretiert werden. Es lässt sich lediglich folgern, dass ein Paar von Assoziationen im Vergleich zu einem anderen Paar von Assoziationen *relativ* stärker ausgeprägt ist. Dieser Sachverhalt wird im Folgenden als Relativität des IAT bezeichnet. Dass die Relativität des IAT tatsächlich praktische Probleme aufwirft, zeigen Studien, in denen eine der Zielkonzeptkategorien ausgetauscht (Brendl, Markman & Messner, 2001; Rothermund & Wentura, 2001) beziehungsweise eines der abstrakten Zielkonzepte mit einer spezifischen Bedeutung belegt wird (z.B. Karpinski, 2004). In diesen Studien zeigen sich deutliche Veränderungen der IAT-Effekte.

Unabhängig von dieser Problematik finden sich Hinweise darauf, dass IAT-Effekte auch dann auftreten können, wenn offenkundig keine assoziativen Verknüpfungen vorliegen (Mierke & Klauer, 2003; Rothermund & Wentura, 2001). Wenn assoziative Verknüpfungen nur hinreichend, nicht aber notwendig für das Auftreten der Effekte sind, kann nicht vom Auftreten eines IAT-Effekts auf das Vorliegen bestimmter Assoziationen geschlossen werden. Diese Problematik wird im Zusammenhang mit anderen Fragen im Rahmen der Darstellung alternativer theoretischer Modelle des IAT-Effekts (Brendl et al., 2001; De Houwer, 2001b; Mierke & Klauer, 2001, 2003; Rothermund & Wentura, 2001) eingehender diskutiert. Im

Folgenden sollen zunächst Untersuchungen geschildert werden, die auf eine inhaltliche Validierung des IAT abzielen.

3.1 Befunde zur Validität des IAT

Der IAT wurde in einer großen Zahl experimenteller und differentialpsychologischer Untersuchungen auf unterschiedliche Fragestellungen angewendet und geprüft. Einige dieser Arbeiten sollen im Folgenden dargestellt und diskutiert werden. Dabei werden zunächst experimentelle Befunde dargestellt, bei denen der IAT als abhängige Variable in Experimenten oder Quasi-Experimenten zum Einsatz kam. Im Anschluss daran soll eine Reihe differentialpsychologisch orientierter Studien geschildert werden, bei denen der IAT als Maß interindividueller Unterschiede verwendet wurde. Ein Fokus liegt dabei auf solchen Studien, die sich mit der Erfassung der Testgütekriterien verschiedener konkreter IATs auseinandersetzen haben. Befunde, die der Prüfung theoretischer Modelle des IAT dienen, werden im Anschluss daran in einem gesonderten Teil behandelt.

Experimentelle Befunde zur Validität des IAT

Eine große Zahl von Untersuchungen nutzt den Ansatz bekannter Gruppen zur Validierung spezifischer IATs. Dieser Ansatz setzt voraus, dass die Ausprägung einer abhängigen Variablen für zwei oder mehr spezifische Gruppen bekannt ist oder aus theoretischer Sicht zu erwarten ist. Ist diese Voraussetzung erfüllt, kann aus einer Übereinstimmung zwischen den empirischen gemessenen und den theoretisch zu erwartenden Ergebnissen auf die Validität des Verfahrens geschlossen werden. Eine in diesem Zusammenhang häufig verwendete Strategie basiert auf dem Effekt, dass Individuen dazu neigen, Gruppen, denen sie selbst angehören, positiver zu bewerten. So zeigte sich eine positivere mit dem IAT erfasste Einstellung von japanisch- beziehungsweise koreanischstämmigen Amerikanern gegenüber ihrer jeweiligen Eigengruppe (Greenwald et al., 1998). Da diese Untersuchung in verschiedener Hinsicht prototypisch für quasi-experimentelle Untersuchung im Paradigma bekannter Gruppen ist, sollen die verwendeten Prozeduren und Materialien im Folgenden etwas eingehender dargestellt werden.

An der Untersuchung nahmen 17 koreanischstämmige und 15 japanischstämmige Amerikaner teil. Bei der Konstruktion des IAT wurden neben 25 positiven und 25 negativen Wörtern, die einer Normierung von Bellezza, Greenwald und Banaji (1986) entnommen sind, 25 typisch japanische und 25 typisch koreanische Nachnamen als Stimulusmaterial

verwendet. Da die japanischen Namen deutlich länger waren, als die koreanischen, wurde in einem zweiten, von jedem Versuchsteilnehmer bearbeiteten IAT eine trunkierte Variante dieser Namen verwendet (z.B. „Kawa“ statt „Kawabashi“). Alle Versuchsteilnehmer bearbeiteten in direkter Folge den IAT mit trunkierten und den mit untrunkierten japanischen Namen.

In der ersten Phase wurde die Zielkonzeptaufgabe eingeübt, in diesem Fall die Klassifikation der Nachnamen als „koreanisch“ oder „japanisch“. In der zweiten Phase wurde die Attributaufgabe, also die Bewertung der positiven und negativen Wörter, eingeübt. In der dritten Phase wurden beide Aufgaben im Wechsel bearbeitet, auf einen Durchgang mit der Attributaufgabe folgte immer ein Durchgang mit der Zielkonzeptaufgabe und umgekehrt. In der vierten Phase wurde schließlich die Zielkonzeptaufgabe mit umgekehrter Tastenzuordnung eingeübt, und in der fünften Phase erneut mit der Attributaufgabe im Wechsel bearbeitet. Die Übungsphasen umfassten jeweils zwei Aufgabenblöcke mit je 50 Durchgängen. Die Phasen, in denen beide Aufgaben gemischt bearbeitet wurden, umfassten jeweils vier Aufgabenblöcke mit 50 Durchgängen. Das Response-Stimulus-Intervall (RSI) zwischen der Reaktion im letzten Durchgang und der Darbietung des nächsten Reizes lag bei 250 ms. Die Tastenzuordnung wurde so ausbalanciert, dass für eine Hälfte der Versuchsteilnehmer in der ersten gemischten Phase koreanische Nachnamen und positive Wörter der gleichen Taste zugeordnet waren (koreanisch + positiv) und die beiden verbleibenden Kategorien der anderen, während für die zweite Hälfte der Versuchsteilnehmer zunächst japanische Nachnamen und positive Wörter auf einer Taste lagen (japanisch + positiv).

IAT-Effekte wurden durch die Differenz zwischen den Antwortlatenzen in der Bedingung koreanisch + positiv minus den Latenzen in der Bedingung japanisch + positiv berechnet und durch die gepoolte Standardabweichung geteilt. Die resultierenden IAT-Effekte für japanischstämmige Versuchsteilnehmer ($d = .159$ für untrunkierte und $d = .139$ für trunkierte Nachnamen) unterschieden sich signifikant von denen koreanischstämmiger Versuchsteilnehmer ($d = -.144$ für untrunkierte und $d = -.119$ für trunkierte Nachnamen) und zeigen somit den erwarteten Unterschied zwischen den Gruppen. Die Scores beider IATs korrelierten signifikant mit einem Gefühlsthermometer, nicht jedoch mit einem semantischen Differential zur Einstellung gegenüber den Ethnien. Erwartungsgemäß ließ sich also der bekannte Effekt der Ingroup-Präferenz auch für den IAT als abhängige Variable nachweisen, gleichwohl die Ergebnisse des IAT nur teilweise mit den expliziten Maßen zusammenhingen.

Vergleichbare Unterschiede zeigten sich auch in einigen anderen Untersuchungen, etwa in Form von Unterschieden zwischen Ost- und Westdeutschen (Kühnen et al., 2001), Bayern und Norddeutschen (Neumann et al., 1998), sowie Versuchsteilnehmern jüdischen und christlichen Glaubens (Rudman, Greenwald, Mellott & Schwartz, 1999). In einer weiteren Untersuchung mit bekannten Gruppen konnten Teachman, Gregg und Woody (2001) demonstrieren, dass ein IAT zur Messung angstbezogener Assoziationen zwischen Spinnen- und Schlangenphobikern diskriminieren kann. Ein IAT zur Erfassung der sexuellen Orientierung zeigte systematische Unterschiede zwischen hetero- und homosexuellen Probanden auf (Banse, Seise & Zerbes, 2001). Ebenso fanden sich erwartungskonforme Geschlechtsunterschiede bezüglich der Assoziation zwischen Geschlecht und Mathematik beziehungsweise Geisteswissenschaften (Nosek, Banaji & Greenwald, 2002).

Obwohl ein Großteil der Ergebnisse mit bekannten Gruppen die Validität spezifischer Varianten des IAT belegt, gibt es auch Studien, in denen die erwarteten Differenzen nicht auftreten. So wiesen Spinnenphobiker und nicht spinnenphobische Probanden unerwarteterweise vergleichbar negative Einstellungen gegenüber Spinnen in einem IAT auf (De Jong et al., 2003). Swanson et al. (2001) fanden eine negative Einstellung von Rauchern gegenüber dem Rauchen. Überraschend ist auch, dass fettleibige Probanden fetthaltige Kost – folgt man den Ergebnissen des IAT - eher ablehnen als normalgewichtige (Roefs & Jansen, 2002). Ferner zeigten ältere und jüngere Probanden in einem IAT eine relativ positivere Einstellung zur Kategorie „jugendlich“, sowie farbige und weiße Probanden eine relativ positivere Einstellung gegenüber weißen Personen (Nosek et al., 2002).

Obwohl eine Vielzahl von Untersuchungen im Paradigma bekannter Gruppen die erwarteten Gruppenunterschiede findet, gibt es also eine große Zahl von Untersuchungen mit unerwartetem Ergebnis. Ob das Ausbleiben der erwarteten Unterschiede darauf zurückgeht, dass die Voraussetzungen für die Anwendung des Ansatzes bekannter Gruppen in den genannten Fällen nicht erfüllt waren, oder ob die Befunde einen wichtigen Beitrag zum Verständnis der untersuchten Phänomene liefern, kann auf Basis der gegenwärtigen Datenlage nicht entschieden werden. Bevor die angesprochenen Befunde jedoch in den jeweiligen Inhaltbereichen integriert und substantiell interpretiert werden können, sollte eine kritische Prüfung erfolgen.

Der IAT als Test

Die von Greenwald et al. (1998) vorgeschlagene Nutzung des IAT als Maß inter-individueller Unterschiede kann auf Basis der Ergebnisse mit bekannten Gruppen nicht gerechtfertigt werden, selbst wenn die Validierung positiv ausfällt, das heißt, die erwarteten Ergebnisse auftreten. Insbesondere eine individuelle Interpretation von Testergebnissen setzt nämlich voraus, dass ein Verfahren auch anderen psychometrischen Anforderungen genügt. Im Folgenden sollen deshalb Untersuchungen geschildert werden, die klassische psychometrische Kriterien an den IAT anlegen.

Reliabilität und Interne Konsistenz

Eine Vielzahl von Arbeiten berichtet Ergebnisse bezüglich der internen Konsistenz und Reliabilität der verwendeten IATs. Die entsprechenden Befunde werden am Ende dieses Abschnitts zusammengefasst. Zunächst soll eine Arbeit von Steffens und Buchner (2003) beschrieben werden, die das Verhältnis von Reliabilität und interner Konsistenz eines IAT genauer betrachtet. In dieser Arbeit wurde die Einstellung gegenüber homosexuellen Paaren an zwei Messzeitpunkten durch einen IAT erfasst. In Experiment 1 betrug das Intervall zwischen den Messzeitpunkten eine Woche, während in Experiment 2 nur zehn Minuten zwischen den Messungen lagen. In dem IAT wurden Paare von Vornamen verwendet, die entweder gleichen oder unterschiedlichen Geschlechts waren, sowie positiv und negativ bewertete Adjektive. Die Versuchsteilnehmer wurden informiert, dass die Namenspaare sich auf hetero- und homosexuelle Paare bezögen. Neben einer evaluativen Entscheidung bezüglich der Adjektive, waren die Versuchsteilnehmer instruiert, die Namenspaare abhängig von der ihrer Geschlechtszusammensetzung als hetero- und homosexuell zu klassifizieren.

Zur Berechnung der internen Konsistenz wurde für jedes der zwanzig verwendeten Stimuluswörter ein IAT-Effekt berechnet. Für die resultierende „Skala“ mit 20 Items ergab sich in beiden Experimenten eine relativ hohe interne Konsistenz zwischen $\alpha = .88$ und $\alpha = .93$. Die Stabilität der Messergebnisse wurde durch eine Strukturgleichungsanalyse geprüft. Dabei wurden drei Strukturgleichungsmodelle kontrastiert. Das Reliabilitätsmodell basiert auf der Annahme, dass alle Messwerte durch eine einzelne latente Variable und eine Fehlerkomponente determiniert sind. Die Kernannahme dieses Modells ist, dass die gemeinsame Varianz über die Messzeitpunkte hinweg äquivalent ist. Im Gegensatz dazu geht das Stabilitätsmodell davon aus, dass die Messwerte durch zwei korrelierte latente Variablen und eine Fehlerkomponente determiniert sind. Dieses Modell erlaubt systematische

Varianzkomponenten, die für einzelne Messzeitpunkte spezifisch sind. Die Korrelation der latenten Variablen ist ein Index für die Stabilität der Messung. Im dritten Modell, dem Konsistenzmodell, wird die gemeinsame Varianz über die Messzeitpunkte durch eine dritte latente Variable erklärt. Für den Fall zweier Messungen an zwei Messzeitpunkten sind Konsistenz- und Stabilitätsmodell datenäquivalent, das heißt, die Modellpassung der strukturell unterschiedlichen Modelle ist für einen gegebenen Datensatz identisch.

Eine Anpassung der drei Modelle an die Daten zeigt, dass das Reliabilitätsmodell weder die Daten von Experiment 1, noch die von Experiment 2 hinreichend erklären kann. Im Gegensatz dazu kann das Stabilitätsmodell die Daten gut erklären. Hier zeigt sich im Vergleich zu den hohen internen Konsistenzen innerhalb der Messzeitpunkte ein beträchtlicher Teil nichtgemeinsamer Varianz zwischen den Messzeitpunkten. Die geschätzte Stabilität der Messung liegt mit $r = .56$ (Experiment 1) und $r = .61$ (Experiment 2) deutlich unterhalb dessen, was aufgrund der hohen internen Konsistenzen erwartet werden müsste. Dies spiegelt sich auch in den Retest-Reliabilitäten von $r = .50$ (Experiment 1) und $r = .55$ (Experiment 2) wider. Insgesamt deuten die Daten von Steffens und Buchner (2003) darauf hin, dass der verwendete Einstellungs-IAT zu einem bedeutenden Teil situationsspezifische Varianz erfasst. Ob die geringe transsituationale Stabilität der Messungen auf das verwendete Messinstrument zurückgeht oder eher inhaltsabhängig ist, lässt sich auf Basis der vorliegenden Daten nicht eindeutig klären.

Eine weitere Untersuchung der psychometrischen Eigenschaften eines Einstellungs-IATs wurde von Cunningham, Preacher und Banaji (2001) vorgelegt. Neben einem IAT zur Erfassung der Einstellung gegenüber weißen und farbigen Personen wurden den Teilnehmern eine Variante dieses IAT mit Antwortfenster und eine Variante des affektiven Primings mit Antwortfenster (Draine & Greenwald, 1998; Greenwald et al., 1996) vorgelegt. Ferner bearbeiteten die Teilnehmer die Modern Racism Scale (McConahay, 1986). Alle vier Einstellungsmaße wurden an vier verschiedenen Testzeitpunkten erfasst, die jeweils zwei Wochen auseinander lagen. Die internen Konsistenzen der vier IATs lagen zwischen .88 zum ersten Messzeitpunkt und .68 bei der letzten Messung. Die interne Konsistenz des IAT sank mit steigender Anzahl der vorhergehenden Sitzungen. Die Retest-Korrelationen liegen zwischen .16 und .50. Sie fallen also, wie in der oben geschilderten Untersuchung von Steffens und Buchner, eher niedrig und deutlich geringer aus als die Schätzungen der internen Konsistenz.

Auch in anderen Untersuchungen bewegen sich die Schätzungen der internen Konsistenz meist im Bereich von $\alpha = .80$, etwa bei der Messung impliziter Einstellungen oder dem Selbstvertrauen (Banse et al., 2001; Bosson et al., 2000; Greenwald et al., 2000; Greenwald & Farnham, 2000; Greenwald & Nosek, 2001), jedoch auch bei der Messung von Konstrukten wie Ängstlichkeit (Egloff & Schmukle, 2002), oder Schüchternheit (Asendorpf et al., 2002). Im Vergleich zu anderen reaktionszeitbasierten Verfahren sind interne Konsistenzen in dieser Höhe durchaus zufrieden stellend (Banse, 2001; Bosson et al., 2000). Nicht geklärt ist gegenwärtig, wie sich verschiedene Strategien zur Generierung von Subtests auf die interne Konsistenz auswirken. In der Literatur finden in diesem Zusammenhang drei verschiedene Techniken Anwendung. Namentlich werden entweder die einzelnen Durchgänge durch zufällige Zuweisung zu Subtests zusammengefasst, die verschiedenen Blöcke des experimentellen Designs als Subtests aufgefasst, oder die verwendeten Stimuli als Subtests verstanden. Insbesondere bei sehr heterogenen Stimulismengen wäre zu erwarten, dass letztere Technik zu geringeren internen Konsistenzen führt, da die beiden erstgenannten Techniken Stimuluseffekte durch die Aggregation verdecken können.

Schätzungen der Stabilität fallen auch in anderen Untersuchungen geringer aus als die internen Konsistenzen. In der Literatur findet sich eine Spanne zwischen .27 (Cunningham et al., 2001) und .69 (Bosson et al., 2000). Ein typischer Befund scheinen Retest-Korrelationen in der Region von .60 zu sein (Banse et al., 2001; Bosson et al., 2000; Greenwald & Farnham, 2000). Welche Variablen diese recht extremen Schwankungen moderieren, ist gegenwärtig ungeklärt. In der Summe zeigen die Ergebnisse jedoch, dass ein sorgfältig konstruierter IAT einen durchaus befriedigenden Anteil systematischer Varianz erfassen kann.

Kriteriumsvalidität des IAT

Die Tatsache, dass mittlere IAT-Effekte mit bekannten Gruppen häufig dem erwarteten Muster folgen, sprechen für eine prinzipielle Validität des IAT. Theoriekonforme Mittelwertsunterschiede sind jedoch keine Garantie dafür, dass ein Verfahren auch auf individueller Ebene brauchbare Ergebnisse liefert, sondern nur eine notwendige Bedingung. Eine individuelle Interpretation von Testscores setzt zusätzlich eine ausreichend hohe Korrelationen mit einem verlässlichen Außenkriterium für das zu messende Konstrukt voraus. Die stichhaltigsten Belege für die Validität eines Verfahrens liefern dabei Verhaltensprognosen. Für den IAT liegt eine – im Vergleich zur Gesamtzahl der Publikationen, die sich mit dem IAT befassen – vergleichsweise kleine Zahl von Untersuchungen vor, die eine

entsprechende Validierung spezifischer IATs dokumentieren. Die entsprechenden Studien sollen im Folgenden geschildert werden.

In der Untersuchung von Asendorpf et al. (2002) sollte geprüft werden, ob Verhalten in einer schüchternheitsinduzierenden Situation durch einen entsprechenden IAT vorhergesagt werden kann. Im Verlauf der Untersuchung interagierten die Teilnehmer mit einer physisch attraktiven, unbekannten Person des jeweils anderen Geschlechts. Das Verhalten der Teilnehmer in dieser Situation wurde hinsichtlich mehrerer Schüchternheitsindikatoren kodiert. Dabei wurde spontanes Verhalten, etwa sich selbst am Kopf kratzen, und kontrolliertes Verhalten, etwa der Redeanteil in der Interaktion, unterschieden. Im Anschluss daran wurde die differentielle Assoziation zwischen Selbst und Schüchternheit beziehungsweise Anderen und Schüchternheit mit einem IAT gemessen. Asendorpf et al. (2002) fanden, dass spontanes schüchternes Verhalten besser durch den IAT vorhergesagt werden kann als durch explizite Maße der Schüchternheit, während die expliziten Maße bei der Vorhersage kontrollierten Verhaltens besser abschnitten. In einer vergleichbaren Studie von Egloff und Schmukle (2002) konnten verschiedene Verhaltensindikatoren für Ängstlichkeit in einer stressreichen Situation durch einen Selbst-Andere Ängstlichkeits-IAT vorhergesagt werden.

In einer Untersuchung von McConnell und Leibold (2001) interagierten die Versuchsteilnehmer in zwei experimentell kontrollierten Situationen mit einem farbigen und einem weißen Experimentator. Das Verhalten der Versuchsteilnehmer während dieser Interaktionen wurde im Anschluss an das Experiment von trainierten Beurteilern und den Experimentatoren selbst bezüglich verschiedener Indikatoren diskriminierendes Verhaltens kodiert, und bezüglich des allgemeinen Grades an Diskriminierung beurteilt. Ein im Verlauf des Experiments erhobener IAT zur Einstellung gegenüber farbigen Personen korrelierte signifikant mit der allgemeinen Beurteilung der Interaktion und einigen der kodierten Verhaltensfacetten. So zeigten sich Korrelationen des IAT mit der allgemeinen Sprechzeit, mit dem Lächeln der Probanden, mit der Zahl der Sprechfehler und mit verstärktem Zögern in der Interaktion mit dem farbigen Experimentator. Die Interaktion mit dem farbigen Versuchsleiter erfolgte jedoch immer nach der Bearbeitung expliziter Vorurteilsmaße und dem IAT. Dies könnte zu einer Aktivierung verschiedener vorurteilsbezogener Gedächtnisinhalte oder zu einer verstärkten Anwendung der Kategorie „farbig“ auf den zweiten Experimentator geführt, und somit das „natürliche“ Verhalten in der Interaktionssituation verändert haben. Abgesehen davon dürfte der Zweck der Untersuchung den Versuchs-

teilnehmern kaum verborgen geblieben sein. Trotz dieser methodischen Probleme liefert die Übereinstimmung von IAT und Verhalten ein starkes Argument für die Validität des IAT.

Weniger optimistisch stimmt eine Untersuchung von Karpinski und Hilton (2001). Ein zentrales Ergebnis dieser Untersuchung ist, dass explizite Maße, nicht jedoch der IAT, in der Lage sind, eine Produktentscheidung der Versuchsteilnehmer zwischen einem Apfel und einem Schokoriegel vorherzusagen. Zwar scheint die untersuchte Thematik trivial, trotzdem gibt es wenig Grund anzunehmen, dass die Teilnehmer nicht willens waren, ihre Einstellungen zu Äpfeln und Schokoriegeln wahrheitsgemäß zu berichten. Die Abwesenheit einer Korrelation zwischen IAT und Produktentscheidung kann nicht durch eine geringe Reliabilität der Produktentscheidung erklärt werden, da in diesem Fall auch die Korrelationen zwischen Verhalten und den expliziten Maßen ausbleiben müsste. Die Autoren folgern deshalb, dass eine Kontaminierung des verwendeten IAT mit einstellungsirrelevanten Assoziationen die Ursache für das Scheitern der Verhaltensvorhersage ist. Eine eingehende Diskussion dieser Folgerung und der angesprochenen Untersuchung findet sich in einem gesonderten Abschnitt zu dem von Karpinski und Hilton (2001) vorgeschlagenen Modell der „Environmental Associations“.

Auf Basis der publizierten Untersuchungen lässt sich also festhalten, dass der IAT in verschiedenen Anwendungsbereichen spontanes Verhalten wenigstens teilweise vorhersagen kann. Die erfolgreiche Validierung eines spezifischen IAT darf aber nicht als Beleg für die Validität der gesamten Verfahrensklasse betrachtet werden, denn verschiedene IATs haben wohl kaum mehr gemeinsam als etwa verschiedene Likert-Skalen. Es lässt sich jedoch folgern, dass ein sorgfältig konstruierter IAT durchaus brauchbare Verhaltensvorhersagen liefern *kann*. Unabhängig davon scheint ein Problem dieser Folgerung jedoch zu sein, dass Untersuchungen, die keinen Zusammenhang zwischen IAT und Verhalten finden, schwer publizierbar sind - sofern sie, wie Karpinski und Hilton (2001), keinen theoretischen Rahmen für ihre Befunde liefern. Die informelle Kommunikation auf Tagungen und Kongressen legt nahe, dass einige solcher Untersuchungen unveröffentlicht in den Schubladen verschiedener Arbeitsgruppen liegen. In dieser Hinsicht ist eine positive Bilanz der publizierten Resultate zum Zusammenhang von IAT und Verhalten also mit Vorsicht zu genießen.

Konvergente Validität des IAT

Die Prüfung von Korrelationen zwischen IATs und alternativen Maßen des gleichen Konstrukts ist eine weitere, zur Validierung des IAT häufig angewandte Strategie.

Entsprechende Befunde zur konvergenten Validität finden sich in vielen Papieren, insbesondere in Form von Korrelationen zwischen IAT und expliziten Verfahren. Eine Ausnahme bilden die schon angesprochenen Untersuchungen von Cunningham et al. (2001) und Bosson et al. (2000), in denen zusätzlich auch andere implizite Maße verwendet wurden. In der Untersuchung von Cunningham et al. (2001) lag die mittlere Korrelation zwischen den verschiedenen impliziten Maßen über die Messzeitpunkte hinweg bei lediglich .19, aber eine Strukturgleichungsanalyse der Daten ergibt deutlich höhere Korrelationen zwischen den messfehlerbereinigten latenten Variablen. Bosson et al. (2000) fanden demgegenüber keine Zusammenhänge zwischen verschiedenen impliziten Maßen des Selbstvertrauens, etwa supra- und subliminalem Priming, einem IAT und einer Name-Letter Präferenztaufgabe. Nullkorrelationen zwischen IAT und Primingverfahren finden sich auch in den Untersuchungen von Marsh et al. (2001) zur Einstellung gegenüber Kondomen sowie von Sherman und Kollegen (2003) zur Einstellung gegenüber dem Rauchen von Zigaretten. Eine Ausnahme bildet die Untersuchung von Gawronski und Kollegen (Gawronski, Ehrenberg, Banse, Zukova & Klauer, 2003), in der die Stärke der Assoziation des Zielkonzepts Geschlecht mit den Attributen Haushalts- vs. Karriereorientierung über einen IAT erfasst wurde. In dieser Untersuchung zeigte sich ein Zusammenhang zwischen stereotypgeleiteten Antworttendenzen im „Who said what?“ Paradigma (Taylor et al., 1978) und der Stärke stereotyper Assoziationen gemäß der Messung mit dem IAT. Je stärker die stereotypen Assoziationen ausfielen, desto eher wurden den Teilnehmern in einer fiktiven Gruppendiskussion stereotypkonsistente Aussagen zugeordnet. Alles in allem scheint die Konvergenz zwischen verschiedenen impliziten Verfahren jedoch eher niedrig zu sein.

Ähnliches gilt für den Zusammenhang zwischen IAT und expliziten Maßen. In der Untersuchung von Cunningham et al. (2001) lagen die Korrelationen zwischen einem IAT und der Modern Racism Scale (McConahay, 1986) zwischen $r = .08$ und $r = .26$. Vergleichbar niedrige Korrelationen zwischen IAT und expliziten Maßen ergaben sich auch in einer Vielzahl anderer Untersuchungen (z.B. Bosson et al., 2000; Egloff & Schmukle, 2002; Greenwald et al., 1998; Karpinski & Hilton, 2001; Rudman et al., 1999; Swanson et al., 2001). Eine noch unveröffentlichte Metaanalyse von Hofmann, Gschwendner, Gawronski, Le, und Schmitt (2003) ergibt für die Korrelation zwischen IAT und expliziten Verfahren über 81 Studien hinweg eine durchschnittliche Effektgröße von .24. Ungefähr ein Drittel der Variabilität zwischen den verschiedenen Studien geht dabei auf Moderatorvariablen zurück. Einen moderierenden Effekt hatten das Untersuchungsthema, die kognitive Elaboration des

Themas, die Reihenfolge impliziter und expliziter Maße, die Kompatibilitätsreihenfolge und die Art des verwendeten expliziten Maßes. Überdurchschnittlich hohe Korrelationen ergaben sich mit IATs zu Produkteinstellungen und Selbstkonzept. Wie erwartet sind die Korrelationen dann größer, wenn die soziale Sensitivität der erfassten Themen geringer ausfällt. Überraschenderweise fallen die Korrelationen offenbar höher aus, wenn die expliziten Maße zeitlich vor dem IAT erhoben werden und wenn die Kompatibilitätsreihenfolge innerhalb eines IAT ausbalanciert wird. Ebenso überraschend ist, dass die Korrelationen mit standardisierten Fragebögen geringer ausfallen als Korrelationen mit Adjektivratings, Gefühlsthermometern und semantischen Differentialen. Letzterer Effekt könnte jedoch auf eine Konfundierung zwischen der Art der verwendeten expliziten Maße und dem Anwendungsbereich zurückgehen. Offenbar ist die häufig niedrige konvergente Validität spezifischer IATs jedoch kein prinzipielles Problem, da vereinzelt auch Korrelationen über .30 berichtet werden (Asendorpf et al., 2002; Banse et al., 2001; McConnell & Leibold, 2001).

Die von Cunningham et al. (2001) berichteten Strukturgleichungsanalysen zeigen, dass IAT, Response-Window Priming und Modern Racism Scale nach der Bereinigung von Messfehlern im Bereich von .26 bis .48 korreliert sind. In Anbetracht der beträchtlichen Divergenz zwischen den Korrelationen auf latenter und beobachteter Ebene bleibt jedoch die Frage, ob der hohe Anteil nichtgemeinsamer Varianz an den Messungen eine individuelle Interpretation der Testscores erlaubt. Der Befund, dass ein beachtlicher Teil der innerhalb eines Verfahrens systematischen Varianz nicht über die Verfahren geteilt ist, gibt Anlass zu Bedenken. Unter anderem soll im Verlauf der vorliegenden Arbeit (Experimente 9 bis 12) geklärt werden, ob und inwiefern dieser verfahrensspezifische systematische Varianzanteil Facetten des zu messenden Konstrukts erfasst.

Die insgesamt heterogene Befundlage zur konvergenten Validität des IAT ist schwer einzuordnen. Einige Autoren argumentieren, dass eine Dissoziation zwischen impliziten und expliziten Maßen als Hinweis auf eine Überlegenheit der impliziten Verfahren zu werten sei (Banaji, 2001; Greenwald et al., 2002; Greenwald et al., 1998). Tatsächlich können solche Dissoziationen, wie vielfach diskutiert, zurückgehen auf ein durch soziale Erwünschtheit oder mangelnde introspektive Zugänglichkeit der zu messenden Konstrukte vermitteltes Versagen der expliziten Maße. Unabhängig davon ist jedoch das Ausbleiben von Zusammenhängen mit Maßen des gleichen Konstrukts kein gutes Argument für die psychometrische Gültigkeit eines Verfahrens. Wenn nicht a priori festgelegt werden kann, ob eine Korrelation zu erwarten ist

oder nicht, führt diese Argumentation zu einer tautologischen Vereinbarkeit mit allen möglichen empirischen Befunden, und somit zu deren Bedeutungslosigkeit. Das Auftreten von Korrelationen zwischen impliziten und expliziten Vorurteilsmaßen in der Untersuchung von Cunningham et al. (2001), die starke Effekte sozialer Erwünschtheit erwarten lassen würde, sollte also ebenso kritisch hinterfragt werden, wie das Ausbleiben von Korrelationen in der Untersuchung von Karpinski und Hilton (2001), in der solche Effekte kaum zu erwarten wären.

Effekte prozeduraler Variablen

Die bisher geschilderten Untersuchungen zeigen, dass der IAT im Vergleich zu anderen impliziten Verfahren vergleichsweise gute psychometrische Qualitäten aufweist. Das auffälligste Kennzeichen der Befunde ist jedoch ihre enorme Uneinheitlichkeit. Dies könnte nicht zuletzt daran liegen, dass die Prozedur des IAT in vielerlei Hinsicht nicht exakt spezifiziert ist, sodass sich die verschiedenen Untersuchungen neben den thematisierten Inhalten auch bezüglich diverser prozeduraler Variationen unterscheiden. Solche Unterschiede finden sich zum Beispiel bezüglich der Zahl und Länge der Aufgabenblöcke, bezüglich des Response-Stimulus Intervalls (RSI) und bezüglich der Reihenfolge der kompatiblen und inkompatiblen Tastenzuordnung. Zwar finden Greenwald et al. (1998) keine Unterschiede zwischen einem IAT mit 750 ms RSI und einem IAT mit 150 ms RSI. Ebenso wenig scheint der IAT davon beeinflusst zu sein, ob als Reaktion auf positive Reize mit der linken oder der rechten Hand reagiert werden soll (Greenwald et al., 1998; Mierke & Klauer, 2001, 2003). Das Ausbleiben solcher Effekte ist jedoch generell schwer interpretierbar. In einer Untersuchung mit größerer Teststärke könnten diese durchaus auftreten.

Ein Effekt der Kompatibilitätsreihenfolge wird in einigen Papieren berichtet (Greenwald et al., 1998; Greenwald, Nosek & Banaji, 2003; Nosek et al., 2002). IAT-Effekte sind tendenziell dann stärker ausgeprägt, wenn die kompatible Phase der inkompatiblen vorausgeht. Leider ist die Kompatibilität der Tastenzuordnung in vielen angewandten Kontexten eine Funktion interindividueller Unterschiede und ist damit nicht für alle Versuchsteilnehmer a priori bekannt. Der vielfach vertretenen Ansicht, dem Problem könne durch eine Konstanthaltung der Reihenfolge beider Tastenzuordnungen begegnet werden (Asendorpf et al., 2002; Banse et al., 2001), soll hier widersprochen werden. Erst in diesem Fall führen die Effekte zu einer problematischen Konfundierung zwischen Merkmalsausprägung und Kompatibilitätsreihenfolge. Lässt man in einem Design mit zwei Gruppen, die sich

hinsichtlich der Richtung des IAT-Effekts unterscheiden, alle Teilnehmer die gleiche Reihenfolge der Tastenzuordnungen bearbeiten, so ergibt sich für Teilnehmer mit einem negativen IAT-Effekt systematisch die entgegengesetzte Kompatibilitätsreihenfolge, wie für Teilnehmer mit positivem IAT-Effekt und somit eine unerwünschte Konfundierung zwischen prozeduraler Variable und Merkmalsausprägung.

Scoring-Algorithmus

Ein Versuch den Effekten prozeduraler Variationen und anderer Verzerrungen des IAT-Effekts durch eine Optimierung des Scoring-Algorithmus zu begegnen, wurde von Greenwald et al. (2003) unternommen. Auf der Basis der etwa 1.2 Millionen Datensätze, die auf der IAT-Website der Yale University erhoben wurden, sollte eine Berechnungsvorschrift für den IAT-Effekt generiert werden, die eine Reihe empirisch definierter, wünschenswerter Kriterien erfüllt. Ein zentrales Kriterium war, dass über den ausgewählten Algorithmus möglichst hohe Korrelationen zwischen IAT-Effekten und expliziten Maßen erzielt werden. Diesem Kriterium liegt offenkundig die Annahme zugrunde, dass IAT und explizite Maße das gleiche Konstrukt erfassen. Ferner sollte der neue Algorithmus eine hohe interne Konsistenz der resultierenden Scores aufweisen und eine geringe Korrelation der IAT-Effekte mit den durchschnittlichen Antwortlatenzen. Durch letztgenanntes Kriterium sollte vermieden werden, dass IAT-Effekte von altersabhängigen Unterschieden der reinen Antwortgeschwindigkeit abhängen. Ferner sollten Effekte der Kompatibilitätsreihenfolge und der Testwiederholung minimiert werden.

Der nach diesen Kriterien beste Algorithmus weist einige Unterschiede zu der üblichen Berechnungsvorschrift auf. Das zentrale Merkmal des Algorithmus ist, dass die Differenz zwischen kompatibler und inkompatibler Phase an der mittleren gepoolten Standardabweichung relativiert wird. Das resultierende Maß ähnelt damit dem Effektstärkenmaß d nach Cohen (1977; Cohen, 1992), das ebenfalls als Quotient aus Mittelwertunterschied und Standardabweichung berechnet wird. Durchgänge mit Latenzzeiten von mehr als 10000 ms werden von der Berechnung des IAT-Effekts ausgeschlossen, ebenso die Daten von Probanden, bei denen die Latenzzeiten in mehr als zehn Prozent der Durchgänge unterhalb von 300ms liegen. Die ersten beiden Durchgänge eines Blocks und Durchgänge mit fehlerhaften Antworten werden entgegen der üblichen Praxis nicht von der Auswertung ausgeschlossen.

Ob und inwiefern dieser Algorithmus in der praktischen Anwendung tatsächlich besser abschneidet, ist jedoch zunächst nicht geklärt. Dies hängt unter anderem davon ab, ob die über das Internet gewonnenen Daten eine gute Annäherung an herkömmliche Labordaten liefern, und ob die Optimierungskriterien sinnvoll gewählt sind. Insbesondere die Maximierung der Korrelation mit expliziten Maßen scheint in diesem Zusammenhang nicht unproblematisch. Es ist zumindest denkbar, dass dies Kriterium zur Auswahl eines Scoring-Algorithmus führt, der den Effekt sozialer Erwünschtheit auf die IAT-Scores maximiert. In den weiter unten beschriebenen Experimenten zu methodenspezifischer Varianz im IAT werden einige viel versprechende Ergebnisse mit dem neuen Algorithmus berichtet.

3.2 Varianten des IAT

Im Folgenden soll kurz auf zwei Varianten des IAT eingegangen werden, die vorgeschlagen wurden, um spezifischen Problemen des ursprünglichen IAT Designs zu begegnen. Beide Verfahren versuchen das oben angesprochene Problem der Relativität von Ergebnissen des IAT zu lösen, indem entweder lediglich ein Zielkonzept verwendet wird, wie in der Go/No-Go Assoziationsaufgabe (Nosek & Banaji, 2001), oder aber die Ergebnisse trotz der Verwendung zweier Zielkonzepte voneinander unabhängig sind, wie es in der extrinsisch-affektiven Simonaufgabe (De Houwer, 2003) der Fall zu sein scheint.

Die Go/No-Go Assoziationsaufgabe

Die Go/No-Go Assoziationsaufgabe (GNAT) wurde von Nosek und Banaji (2001) vorgeschlagen und ähnelt dem IAT insofern als Exemplare eines Zielkonzepts und Reize, die Ausprägungen einer Attributdimension repräsentieren, in schneller Folge bearbeitet werden müssen. Im Gegensatz zur Standardprozedur des IAT werden die Teilnehmer instruiert, auf einige der Reize mit einem Tastendruck zu reagieren (go), auf andere jedoch nicht zu reagieren (no-go). Wenn etwa eine assoziative Verknüpfung des Konzepts Insekt mit einer Bewertung durch einen GNAT erfasst werden soll, werden sequentiell positive Wörter, negative Wörter und Insektenbezeichnungen dargeboten. In einer Bedingung soll immer dann eine Taste gedrückt werden, wenn das dargebotene Wort entweder positiv bewertet ist oder die Bezeichnung eines Insekts ist, auf negative Wörter soll nicht reagiert werden. In einer zweiten Bedingung soll immer dann eine Taste gedrückt werden, wenn das dargebotene Wort negativ bewertet ist oder ein Insekt repräsentiert, während auf positive Wörter nicht reagiert werden soll. Die Ergebnisse zeigen, dass die Reaktionen im Vergleich zur jeweils anderen

Bedingung schneller sind, wenn die Konzepte, auf die reagiert werden soll, im Hinblick auf die Attributdimension kompatibel sind, etwa wenn auf Insekten und negative Wörter hin reagiert werden soll, zwei Reizkategorien also, die in Bezug auf Ihre Valenz für die meisten Personen kompatibel sind.

Ein Vorteil des GNAT ist, dass lediglich ein Zielkonzept genutzt wird. Deshalb ist das oben angesprochene Problem der Relativität des IAT für den GNAT weniger oder gar nicht gegeben. Es scheint jedoch, dass Messungen mit dem GNAT nicht in gleicher Weise reliabel sind wie die Effekte eines unmodifizierten IAT (Nosek & Banaji, 2001).

Die extrinsisch-affektive Simonaufgabe

Die extrinsisch-affektive Simonaufgabe (EAST) wurde von De Houwer (2003) entwickelt und basiert auf einer Kombination von Elementen des IAT mit dem affektiven Simoneffekt (De Houwer & Eelen, 1998; De Houwer et al., 1998). Das Material eines EAST besteht aus positiven und negativen Wörtern, die in weißer, grüner oder blauer Farbe dargeboten werden. Die Versuchsteilnehmer werden gebeten, alle weiß dargebotenen Wörter durch Druck einer von zwei Antworttaste als positiv oder negativ zu klassifizieren, sowie die grün oder blau dargebotenen Wörter durch Druck einer der gleichen zwei Antworttasten bezüglich ihrer Farbe. Als blaue und grüne Reize werden Exemplare zweier Zielkonzepte dargeboten.

Die Ergebnisse zeigen, dass die Leistung bei der Klassifikation blauer oder grüner *positiver* Wörter (z.B. Rose) dann besser ist, wenn aufgrund der Tastenzuordnung der Farbe mit der gleichen Taste reagiert werden muss wie auf weiß dargebotene *positive* Wörter. Entsprechend ist die Leistung bei *negativen* blauen oder grünen Wörtern (z.B. Made) dann besser, wenn aufgrund der Farbe die gleiche Taste gedrückt werden muss, wie für weiß dargebotene *negative* Wörter. Für blau oder grün dargebotene Non-Wörter oder Konsonantenketten zeigte sich erwartungsgemäß kein Effekt. Mit dem EAST scheint es also möglich, die Bewertung einzelner Exemplare zu erfassen. Dazu werden die Reaktionszeiten bei der Zuweisung der Exemplare zu Tasten verglichen, die durch die Aufgabenstruktur mit positiven oder negativen Attributausprägungen assoziiert sind.

Beide Varianten des ursprünglichen Designs scheinen eine Lösung des Problems der Relativität von IAT-Messungen zu versprechen. Allerdings sind Reliabilitäten und Mittelwerteffekte des EAST und des GNAT deutlich geringer als die Effekte, die mit herkömmlichen IATs erzielt werden. Des Weiteren steht eine sorgfältige Validierung der

Ergebnisse spezifischer Anwendungen der Verfahren bisher noch aus. Es bleibt also abzuwarten, ob diese sich praktisch bewähren.

3.3 Allgemeine Probleme impliziter Verfahren

Bevor im folgenden Abschnitt eine Reihe theoretischer Modelle zur Erklärung von IAT-Effekten genauer betrachtet wird, sollen noch einige allgemeinere Probleme impliziter Verfahren angesprochen und diskutiert werden. Insbesondere die Interpretation des IAT als Test – namentlich als Maß individueller Unterschiede – birgt solche Probleme. Eine augenfällige Schwierigkeit impliziter Verfahren ergibt sich natürlich aus der Tatsache, dass die Effekte der meisten Untersuchungen in diesem Bereich eher gering ausfallen, vergleichsweise wenig reliabel sind, und eine geringe interne Konsistenz aufweisen (Banse, 2001; Kawakami & Dovidio, 2001). Insbesondere sind, wie schon dargestellt, die Korrelationen verschiedener impliziter Verfahren untereinander häufig wenig befriedigend (Bosson et al., 2000) oder zeigen, dass die Verfahren verschiedene Facetten der zu messenden Konstrukte erfassen (Brauer et al., 2000).

Obwohl diese Kritik auf den IAT, wie weiter oben dargestellt, nur zum Teil zutrifft, verbleiben ungelöste Probleme. So stellt sich die Frage, ob und inwiefern die indirekt gemessenen Konstrukte tatsächlich ein psychologisches *Substrat* für makroskopische Phänomene wie Vorurteile, Stereotypisierung oder das Selbstkonzept darstellen, denn zweifellos besteht das Ziel des IAT weniger darin, Aussagen über Assoziationen zu treffen, als darin, Folgerungen über diese makroskopischen Phänomene der Psychologie zu erlauben (Fiedler, Messner & Blümke, 2003). Selbst *wenn* implizite Verfahren zur Quantifizierung der Stärke assoziativer Verknüpfungen taugen, bleibt also unklar, inwiefern sich auf dieser Basis eine Einstellungs- oder Selbstkonzeptpsychologie betreiben lässt, das heißt, inwieweit die makroskopischen Phänomene der Sozial- oder Persönlichkeitspsychologie sich auf dieses einfache Substrat *reduzieren* lassen. Im besten Fall liefern die neuen Verfahren einfach bessere Methoden zur Erfassung des Altbekanntes. Im schlechtesten Fall jedoch droht die theoretische und empirische Bedeutungslosigkeit der „neuen“ Begriffe, wenn diese sich nämlich auf theoretisch belanglose Phänomene beziehen.

Glücklicherweise kann diese Frage insoweit empirisch beantwortet werden, als die Korrespondenz der impliziten Messungen mit theoretischen Erwartungen geprüft werden kann. Eine solche eingehende Prüfung und Validierung ist unverzichtbar, wenn implizite Verfahren eine ähnlich breite Anwendung finden sollen wie die gängigen Selbstberichtsmaße

oder direkte Verhaltensbeobachtungen, denn nur so kann eine begriffliche Kohärenz sichergestellt werden. Entscheidend ist, dass diese Validierung nicht in zestuös innerhalb bestimmter Verfahrensklassen verbleibt, sondern konkrete Verhaltens- oder Erlebensvorhersagen ableitet und prüft. Dabei stellt sich das Problem, vollständige und bedeutungsvolle Verhaltensstichproben zu definieren, die als Kriterium für implizite Messungen geeignet sind. Erste Ergebnisse deuten darauf hin, dass eine Korrespondenz zwischen Verhaltenskriterien und impliziten Maßen zumindest in bestimmten Bereichen und mit bestimmten Verfahren besteht (Asendorpf et al., 2002; Egloff & Schmukle, 2002; McConnell & Leibold, 2001). Die Einordnung dieser Ergebnisse ist jedoch schwierig. Wie viele Studien, in denen vergleichbare Effekte nicht auftraten, liegen unveröffentlicht in den Schubladen der Scientific Community? Wie viele „Versuche“ sind nötig, um einen einfachen Zusammenhang zwischen den Ergebnissen eines impliziten Verfahrens und konkretem Verhalten zu zeigen? Ohne diese Fragen beantworten zu wollen, kann man sicher sagen, dass nicht jeder - aus inhaltlicher Sicht plausible - Versuch, solche Zusammenhänge nachzuweisen, gelingt.

In alltäglichen Situationen ist menschliches Verhalten meist nicht im engeren Sinne *determiniert*, wie es die Neubestimmung theoretischer Begriffe in Terminus assoziativer Verknüpfungen verheißen mag. Es gibt viele Hinweise darauf, dass Verhalten, explizite Selbsturteile und implizite Messungen nur dann unbewussten Einflüssen ausgeliefert sind, wenn die Beeinflussung durch äußere oder innere Einflussfaktoren - etwa Vorurteile oder die Informationsverarbeitung verzerrende Reizkonstellationen - dem Individuum nicht bewusst wird (Bargh, 1994). Kann man im Hinblick auf diese Tatsache mit Recht davon ausgehen, dass Reaktionstendenzen, die durch bestimmte Eigenschaften der mentalen Repräsentationen bestimmt sind, ein besseres Maß für Vorurteile bilden, als *überlegtes* Handeln im sozialen Kontext? Diese Frage ist offenbar keine genuin empirische. Das zeigt sich schon daran, dass Dissoziationen zwischen den Ergebnissen impliziter und expliziter Maße weder selten noch unerwünscht sind (Karpinski & Hilton, 2001; Nosek et al., 2002; Roefs & Jansen, 2002; Swanson et al., 2001). Die Frage scheint vielmehr zu sein, inwiefern die Perspektive impliziter Verfahren *nützlich* ist. Denn selbst wenn implizite Maße spontanes Verhalten in sozialen Kontexten besser vorhersagen als explizite Maße, bleibt offen, ob der Begriff „Einstellungen“ nunmehr im Sinne von „evaluative Assoziationen“ gebraucht werden sollte (z.B. Fazio & Olson, 2003; Fazio et al., 1986; Greenwald et al., 2002) eine vornehmlich terminologische beziehungsweise konventionelle Frage. Damit sei nicht gesagt, dass empirische Befunde nicht bedeutend zu einer terminologischen Festlegung beitragen können

und müssen. Gerade wegen der großen Distanz zwischen dem beobachteten Verhalten und der inhaltlichen Interpretation, die bei impliziten Verfahren in der Natur der Sache liegen, ist eine gründliche empirische Validierung konkreter Verfahren aber auch allgemeiner Verfahrensklassen unumgänglich.

Die Deklaration eines Verfahrens als Test und die damit verbundene kulturelle und praktische Signifikanz erfordern einen gewissenhaften und verantwortlichen Umgang mit dem jeweiligen Verfahren. Die naive Voraussetzung ungeprüfter Annahmen über die Funktionsweise eines Messinstruments scheint in diesem Zusammenhang unverantwortlich sowohl gegenüber den getesteten Personen als auch der psychologischen Diagnostik als Disziplin. In diesem Zusammenhang sind theoretische Modelle mit funktionalen Annahmen darüber, wann die Validierung eines Verfahrens als erfolgreich gewertet werden kann, was genau gemessen werden soll und wie diese Messung überhaupt zustande kommt, wissenschaftlich wünschenswert aber auch praktisch notwendig. Ob und inwiefern der IAT als Test solchen Kriterien genügen wird kann nicht losgelöst von der Frage betrachtet werden, welche kognitiven und motorischen Prozesse am Zustandekommen der IAT-Effekte beteiligt sind (Fiedler et al., 2003). Im Folgenden sollen deshalb zunächst einige der wichtigsten theoretischen Modelle des IAT-Effekts vorgestellt und diskutiert werden.

3.4 Theoretische Modelle

Theoretische Modelle zum Zustandekommen von IAT-Effekten lassen sich grob in zwei Klassen aufteilen. Zunächst findet sich eine große Zahl von Arbeiten, die sich mit der Frage auseinandersetzen, *was* der IAT misst. Hier reicht die Spanne von Modellen, die substantielle Annahmen über die Repräsentation sozialen Wissens machen (z.B. Greenwald et al., 2002) zu Modellvorstellungen, die die gemessenen Konstrukte auf sehr abstrakter Ebene beschreiben, wie das Figur-Grund Modell (Rothermund & Wentura, 2001). Demgegenüber stehen Modellvorstellungen, die sich mit der Frage befassen, *wie* IAT-Effekte - unabhängig von den gemessenen Konstrukten - zustande kommen. Exemplarisch seien hier das Random-Walk Modell von Brendl und Kollegen (2001), das Stimulus-Response Kompatibilitätsmodell (De Houwer, 2001b) und das Aufgabenwechselmodell des IAT (Mierke & Klauer, 2001, 2003) genannt. Diese zweite Klasse von Modellen macht Annahmen über die *Prozesse*, die am Zustandekommen des IAT-Effekts beteiligt sind, während die erste Klasse die *Inhalte*, das heißt, die zugrunde liegenden mentalen Repräsentationen betrifft. Obwohl die meisten Prozessmodelle Implikationen für die Frage haben, was der IAT misst, und die meisten

Inhaltsmodelle des IAT Implikationen für die Modellierung der beteiligten Prozesse haben, sollen sie in den folgenden Abschnitten zunächst getrennt dargestellt werden.

Was misst der IAT?

Eine von vielen Autoren geteilte Annahme ist, dass der IAT die Stärke assoziativer Verknüpfungen misst (Banaji, 2001; Greenwald et al., 1998; Greenwald & Nosek, 2001). Nach dieser Vorstellung bildet der IAT-Effekt "beides, die Richtung (positiv versus negativ) impliziter Einstellungen und die Stärke der Einstellung (größere Zahlen stehen für größere Differenzen), ..." ab (Banaji, 2001, p. 124. Eigene Übersetzung). Auf dieser Prämisse basiert die so genannte Unified Theory of Social Cognition (Greenwald et al., 2002), die eine Reihe klassischer sozialpsychologischer Phänomene durch ein Repräsentationsmodell sozialen Wissens zu erklären versucht, aber auch eine große Anzahl weiterer Arbeiten (z. B. Asendorpf et al., 2002; Bosson et al., 2000; Maison, Greenwald & Bruin, 2001; Marsh et al., 2001).

Bildet der IAT-Effekt die absolute Assoziationsstärke ab?

Bei genauerer Betrachtung muss man einschränken, dass durch den IAT nicht die Stärke assoziativer Verknüpfungen selbst gemessen wird, sondern Unterschiede der Stärke verschiedener assoziativer Verknüpfungen. Ein starker Effekt in einem Blumen-Insekten Einstellungs-IAT kann somit entweder auf die stärkere Assoziation von Blumen oder Insekten mit einer positiven Bewertung zurückgehen, oder auf die stärkere Assoziation dieser Konzepte mit einer negativen Bewertung. Wie weiter oben schon diskutiert wurde, ist der IAT also ein relatives Maß, auf dessen Grundlage keine Aussage über die absolute Stärke von Verknüpfungen getroffen werden kann (Brendl et al., 2001; Fiedler et al., 2003; Mierke & Klauer, 2001). Selbst wenn vorausgesetzt wird, dass der IAT ausschließlich Assoziationsstärken misst, kann aus starken IAT-Effekten nur gefolgert werden, dass Unterschiede in der Assoziationsstärke der Konzepte mit positiven und negativen Bewertungen vorliegen, während nicht ausgeschlossen werden kann, dass beide Konzepte in absoluter Hinsicht positiv oder beide negativ bewertet sind. Aus dieser Analyse folgt, im Gegensatz zu der oben zitierten Behauptung von Banaji (Banaji, 2001), dass mit einem IAT weder die Richtung noch die Stärke einer einzelnen Einstellung gemessen werden kann, sondern im besten Fall eine Differenz oder ein Verhältnis der Stärke zweier Einstellungen.

Sind Assoziationen notwendig und hinreichend?

Eine weitere Einschränkung im Hinblick auf den Rückschluss von IAT-Effekten auf mentale Repräsentationen folgt aus einigen neueren Befunden, die zeigen, dass assoziative Verknüpfungen keine notwendige Voraussetzung für das Auftreten von IAT-Effekten sind (Brendl et al., 2001; Fiedler et al., 2003; Mierke & Klauer, 2001, 2003; Rothermund & Wentura, 2001). Die zunächst harmlose Erkenntnis, dass auch andere Arten von Relationen und Redundanzen des Reizmaterials zu IAT-Effekten führen können, führt zu dem logischen Dilemma, dass aus dem Auftreten von IAT-Effekten nicht gefolgert werden kann, dass überhaupt Assoziationen vorliegen. Vielmehr ist es jeweils eine empirische Frage, ob Assoziationen erfasst wurden. Im Normalfall kann dieser Problematik durch sorgfältige Auswahl und Kontrolle des Reizmaterials und einer Validierung des konkreten IAT-Verfahrens begegnet werden (z.B. Asendorpf et al., 2002; Egloff & Schmukle, 2002). Es bleibt jedoch immer das Risiko, dass eine unkontrollierte und unerwünschte Konfundierung die empirischen Befunde bedingt.

Kategorie oder Exemplar?

Wenn der IAT, wie oben beschrieben, die Richtung und Stärke assoziativer Verknüpfungen abbildet, so stellt sich die Frage, *welche* Assoziationen hier genau gemeint sind. In theoretischer Hinsicht könnten sowohl Assoziationen auf der Ebene der Konzepte als auch auf Ebene der Reize, die als Exemplare dieser Konzepte verwendet werden, die entscheidende Rolle spielen (Blümke & Friese, 2004; De Houwer, 2001b; Steffens & Plewe, 2001). Dies ist insofern von praktischer Relevanz, als in vielen Bereichen Dissoziationen zwischen der Bewertung einer Kategorie im Allgemeinen und bestimmten Exemplaren dieser Kategorien vorliegen können. Für die meisten praktischen Anwendungen ist ein Lokus auf konzeptueller Ebene eine notwendige Voraussetzung. Wenn etwa Einstellungen gegenüber einer bestimmten Gruppe von Personen erfasst werden sollen, wäre es verfälschend, wenn die Messung durch die Bewertung der dargebotenen spezifischen Exemplare determiniert ist.

De Houwer (2001b) untersuchte diese Frage, indem er die Bewertung der Konzepte und der verwendeten Exemplare orthogonal zueinander manipulierte. In seiner Untersuchung war eine Hälfte der verwendeten Exemplare evaluativ inkongruent zu der Kategorie, die sie repräsentierte (zum Beispiel „A. Einstein“ als positiv bewertetes Exemplar der Kategorie „Ausländer“, „M. Thatcher“ als negativ bewertete Britin), während die andere Hälfte evaluativ kongruent zu der übergeordneten Kategorie war. Wenn der IAT-Effekt auf

Eigenschaften der Bewertung der Zielkonzepte selbst zurückgeht, so sollte ein Kompatibilitätseffekt in Bezug auf die Bewertung der Kategorie auftreten, unabhängig von der Bewertung der Exemplare. Beruht der Effekt jedoch auf den Bewertungen der Exemplare, so sollte ein Kompatibilitätseffekt in Bezug auf die Bewertung der Exemplare auftreten. Die Ergebnisse mit einer britischen Stichprobe zeigen, dass die Kontrastierung (positiv oder negativ) bewerteter britischer Personen und (positiv oder negativ) bewerteter Nicht-Briten zu einem IAT-Effekt auf Basis der Kategoriebewertung führte, während kein Effekt der Bewertung einzelner Exemplare auftrat. Es fand sich also ein IAT-Effekt, der eine positivere Bewertung der Kategorie „Briten“ vermuten lässt, unabhängig von der Bewertung der Exemplare.

In einer Untersuchung von Steffens und Plewe (2001) finden sich im Gegensatz dazu Effekte der verwendeten Exemplare. In dieser Untersuchung wurde zusätzlich zur Bewertung der Attributreize deren semantische Nähe zu den Zielkonzeptreizen manipuliert. Diese systematische Konfundierung zeigte sich deutlich in den resultierenden IAT-Effekten. Die Manipulation beeinflusste jedoch lediglich die Stärke, nicht die Richtung des IAT-Effekts. Die Richtung des IAT-Effekts ist also anscheinend durch die Bewertung der Zielkonzepte determiniert, während die Stärke des Effekts teilweise abhängig von den spezifischen Eigenschaften der verwendeten Exemplare ist. Dies zeigt auch eine weitere Untersuchung von Steffens et al. (2004). In dieser Untersuchung ergab sich eine größere Präferenz für Deutsche gegenüber Ausländern, wenn die Evaluation der verwendeten Zielkonzeptexemplare mit der Evaluation des Konzepts selbst übereinstimmte. Ähnlich wie in der Untersuchung von De Houwer (2001b) wurde ein IAT zur Einstellung gegenüber Deutschen und Ausländern mit den Namen positiv beziehungsweise negativ besetzter bekannter Personen deutscher beziehungsweise nicht-deutscher Herkunft als Stimulusmaterial verwendet. Dass in den Experimenten von Steffens et al. (2004) ein Effekt der Exemplarbewertung auftrat, ist daher überraschend. Der augenfälligste Unterschied zwischen den Untersuchungen besteht darin, dass in dem Experiment von De Houwer die evaluative Kongruenz der Exemplarbewertung innerhalb der IAT-Blöcke manipuliert wurde, während Steffens et al. die Exemplarbewertung zwischen den Probanden manipulierten. Eine mögliche Deutung dieses Befundmusters scheint zu sein, dass eine einheitliche Bewertung der Zielkonzeptexemplare im Sinne eines Kontexteffekts die Verarbeitung der Zielkategorien beeinflusst. In einer unveröffentlichten Arbeit berichten Blümke und Friese (2004) jedoch Effekte irrelevanter Exemplareigenschaften nicht nur zwischen, sondern auch innerhalb der Versuchsbedingungen. Sowohl die

Bewertung einzelner Zielkonzeptkonzeptexemplare als auch stereotyprelevante Konnotationen der verwendeten Attributreize führten jeweils zu einer Verschiebung des resultierenden IAT-Effekts.

In einer neueren Untersuchung konnten Mitchell, Nosek und Banaji (2003) ebenfalls Effekte der verwendeten Exemplare auf den IAT nachweisen. In zwei Experimenten sollten die gleichen prominenten Exemplare der Zielkonzepte „farbige Personen“ und „weiße Personen“ unterschiedlich klassifiziert werden. Im ersten Experiment sollten die Exemplare nach ihrem Berufen klassifiziert werden, wobei alle farbigen Prominenten bekannte Sportler waren und alle weißen Prominenten Politiker. Im zweiten Experiment sollten die gleichen Prominenten anhand ihrer Hautfarbe als „Farbige“ oder „Weiße“ klassifiziert werden. Wurden die Reize bezüglich ihrer Hautfarbe klassifiziert, bevorzugten die Versuchsteilnehmer die weißen Politiker gegenüber den schwarzen Sportlern. Wenn die Exemplare hingegen auf Basis ihrer Beschäftigung als Sportler oder Politiker klassifiziert werden sollten, wurden die (schwarzen) Sportler den (weißen) Politikern deutlich vorgezogen, worauf ein umgekehrtes Vorzeichen des IAT-Effekts in Experiment 1 hinweist. Diese Untersuchung bestätigt offenbar die Annahme, dass eine Änderung des Kontextes, in dem die Exemplare eines IAT präsentiert werden, zu unterschiedlichen Bewertungen führen kann. Eine Änderung dieses Kontextes kann also erreicht werden, in dem die Kategorielabels explizit verändert werden, aber auch indem implizit ein Subtyp vorgegeben wird, etwa durch Darbietung ausschließlich atypischer Kategoriemitglieder.

Kategorie oder Exemplar: Moderatoren

In Anbetracht dieser zunächst widersprüchlichen Befundlage stellt sich die Frage, ob und inwieweit die vorliegenden Ergebnisse materialabhängig sind. Eine Erklärung könnte etwa sein, dass der Einfluss individueller und konzeptueller Bewertung auf die Bewertung eines vorgegebenen Exemplars eine Funktion der Stärke ihrer Kategoriezugehörigkeit und der Stärke der individuellen Assoziation mit einer inkongruenten Bewertung ist. Ergebnisse aus dem Bereich der sozialen Kategorisierung im „Who said what?“ Paradigma (Taylor et al., 1978) zeigen zum Beispiel, dass das Gedächtnis für Verhalten von Personen in einer Gruppendiskussion häufig automatisch mit Bezug auf die Gruppenzugehörigkeit enkodiert wird, während individuelle Eigenschaften der Personen nahezu ignoriert werden (Klauer & Wegener, 1998). Ferner scheinen einige soziale Kategorien kontextabhängig mit höherer Wahrscheinlichkeit angewandt zu werden als andere (Klauer, Ehrenberg & Wegener, 2003;

Klauer, Wegener & Ehrenberg, 2002; Stangor, Lynch, Duan & Glas, 1992). Ein Effekt der Exemplare in einem gegebenen IAT wäre also vor allem dann zu erwarten, wenn die Passung zwischen den Exemplaren und einer verfügbaren Antwortkategorie gering und die Stärke individueller Bewertungen hoch ist. Aus dieser Perspektive wäre zu erwarten, dass Effekte der Exemplare spezifisch für konkrete Reizmaterialien sind, jedoch nicht für den IAT als solchen. Ob und inwiefern in einem konkreten IAT die intendierten Relationen auf konzeptueller Ebene oder andersartige, irrelevante Redundanzen des Reizmaterials (Fiedler et al., 2003) erfasst werden, etwa Figur-Grund Asymmetrien, irrelevante Eigenschaften des Darbietungsformats oder exemplarspezifische Assoziationen, bleibt dessen ungeachtet eine schwer prüfbare empirische Frage, die in vielen publizierten Untersuchungen mit dem IAT nicht angemessen berücksichtigt scheint.

Environmental Associations

Karpinski und Hilton (2001) identifizieren eine solche Konfundierung in ihrem viel beachteten Modell der Environmental Associations. Dieses Modell geht davon aus, dass mit dem IAT nicht nur persönlich akzeptierte Einstellungen erfasst werden, sondern auch kulturell vermitteltes Wissen über Einstellungsobjekte, das von den jeweils individuell akzeptierten Einstellungen abweichen kann. Grundlage dieser Folgerung das schon angesprochene Experiment der Autoren, in dem sich eine Dissoziation zwischen der explizit erfassten Einstellung zu Äpfeln und Schokoriegeln und einem IAT-Maß der gleichen Einstellung ergab. Offenkundig können solche Dissoziationen vielerlei Ursachen haben, etwa eine Tendenz zu sozialer Erwünschtheit. In dem vorliegenden Fall scheint jedoch wenig Grund anzunehmen, dass die Teilnehmer nicht willens waren, ihre Einstellungen zu Äpfeln und Schokoriegeln wahrheitsgemäß zu berichten. Ebenso wenig scheint plausibel, dass die entsprechende Einstellung im vorliegenden Fall introspektiv unzugänglich war, oder dass ein Mangel an Erfahrung mit den Einstellungsobjekten zugrunde liegt. Auf Basis dieser Überlegungen kommen Karpinski und Hilton (2001) zu dem Schluss, dass das Ausbleiben einer Korrelation zwischen IAT und expliziten Maßen nicht auf das Versagen der expliziten Maße zurückgehen kann, sondern vielmehr eine Kontamination des IAT mit kulturell vermitteltem Wissen über Einstellungen widerspiegelt.

Für diese Folgerung spricht auch, dass die mit dem IAT gemessene Einstellung auf eine stärkere Präferenz für Äpfel gegenüber Schokoriegeln hinweist, als die explizit gemessene Einstellung. Karpinski und Hilton (2001) fanden ferner, dass der Apfel-Schokoriegel-IAT

nicht in der Lage war, das Verhalten der Versuchsteilnehmer vorherzusagen, wenn diese am Ende des Experiments die Wahl zwischen einem Apfel und einem Schokoriegel erhielten, während die expliziten Maße mit dem Verhalten korrelierten. Zusammengefasst bestätigen die Ergebnisse die Annahme, dass kulturell geteiltes, aber nicht notwendig individuell akzeptiertes Wissen über die Bewertung von Äpfeln und Schokoriegeln die Ergebnisse im IAT beeinflusst haben.

Eine solche Kontaminierung des IAT mit kulturell vermittelten, einstellungsirrelevanten Assoziationen könnte auch erklären, warum in IAT-Studien ein größerer Anteil weißer Probanden eine negative Einstellung gegenüber Farbigen zeigt als in anderen impliziten Vorurteilsmaßen. Ebenso scheint in diesem Licht plausibel, warum in Untersuchungen mit dem IAT häufig keine Bevorzugung der Ingroup für farbige Versuchsteilnehmer auftritt (z.B. Nosek et al., 2002), während entsprechende Effekte sowohl mit anderen impliziten als auch expliziten Maßen auftreten (vgl. Fazio et al., 1995).

Vertrautheit

Eine weitere, vieldiskutierte Quelle möglicher Verzerrungen des IAT ist die Vertrautheit beziehungsweise Geläufigkeit der verwendeten Reize. Eine Reihe von Arbeiten hat sich mit solchen Effekten befasst. Ein zentrales Ergebnis ist, dass die implizit erfasste Präferenz für ein Zielkonzept größer ausfällt, wenn die Exemplare dieses Zielkonzepts vertrauter sind als die Exemplare des ihm entgegen gesetzten Konzepts (Dasgupta, McGhee, Greenwald & Banaji, 2000; Ottaway, Hayden & Oakes, 2001). Ottaway et al. (2001) konnten weiterhin zeigen, dass die mit dem IAT erfasste Präferenz von Blumen gegenüber Insekten größer ausfällt, wenn das Material vollständig aus vertrauten im Vergleich zu weniger vertrauten Reizen besteht. Es kann jedoch nicht ausgeschlossen werden, dass dieser Effekt auf eine positivere Bewertung der vertrauten Reize - im Sinne eines Mere-Exposure Effekts (z.B. Zajonc, 1968) - zurückgeht. Diese Vermutung bestätigt auch eine Betrachtung der verwendeten Reize: Die vertrauten Blumenexemplare wurden im Vergleich zu den weniger vertrauten als angenehmer eingestuft, während die vertrauten und weniger vertrauten Insektenexemplare als etwa gleich angenehm eingestuft wurden. Aus diesem Blickwinkel könnten die Befunde von Ottaway et al. (2001) einfache Effekte der Exemplarbewertung widerspiegeln (Steffens et al., 2004) und weniger unmittelbare Effekte der Geläufigkeit. Insgesamt scheint der Effekt der Geläufigkeit der verwendeten Reize auf den IAT eher gering, zumindest jedoch nicht die Hauptquelle des Effekts zu sein.

Prozessmodelle des IAT

Neben den oben geschilderten Ansätzen, die schwerpunktmäßig die Frage betreffen, was der IAT misst, wurde auch eine Reihe von Ansätzen vorgeschlagen, die sich damit befassen, welche Mechanismen am Zustandekommen des IAT-Effekts beteiligt sind. Obwohl diese Prozessmodelle des IAT das Zustandekommen von IAT-Effekten durch ein allgemeines und mehr oder weniger inhaltsunabhängiges Modell der vermittelnden Prozesse erklären wollen, haben sie jeweils spezifische Implikationen für die Frage, was der IAT misst. Im Rahmen einer Darstellung der Modelle und der zu ihrer Prüfung durchgeführten empirischen Untersuchungen sollen diese Implikationen im Folgenden diskutiert werden.

Aktivierungsausbreitung

Eine populäre Modellklasse zur Erklärung semantischer und assoziativer Primingeffekte (Neely, 1991) ist das Modell der Aktivierungsausbreitung. Eine Kernannahme dieser Modellklasse ist, dass Wissen in einem assoziativen Netzwerk repräsentiert ist. Einzelne Wissens-elemente werden als Knoten modelliert, die durch assoziative Verknüpfungen unterschiedlicher Stärke verbunden sind (z.B. Collins & Loftus, 1975). Der Mechanismus der Aktivierungsausbreitung postuliert, dass die Aktivierung eines spezifischen Knotens zu einer Ausbreitung der Aktivierung entlang der assoziativen Verknüpfungen führt. Primingeffekte werden in diesem Modell so erklärt, dass die Perzeption eines Primewortes wie „Butter“ zu einer Ausbreitung der Aktivierung auf sinnverwandte und assoziierte Konzepte führt, die dadurch im Folgenden schneller aktiviert werden können, etwa im Rahmen der Verarbeitung eines nachstehenden Targetworts wie „Brot“. Eine analoge Erklärung wurde für affektive Primingeffekte (Fazio et al., 1995; Fazio et al., 1986) vorgeschlagen. Die schnellere und fehlerfreiere Verarbeitung von Wörtern, denen ein affektiv kongruentes Primewort vorangestellt ist, kann so durch die Ausbreitung der Aktivierung entlang evaluativer Assoziationen erklärt werden.

Die Erklärung von IAT-Effekten durch Aktivierungsausbreitung fällt weniger leicht. Dies hängt damit zusammen, dass Aktivierungsausbreitung ein stimulusgetriebener Prozess ist. Während sich die Stimuluszusammensetzung im affektiven Priming bei kongruenten und inkongruenten Wortpaaren unterscheidet, gibt es im IAT keinen stimulusseitigen Unterschied zwischen den Bedingungen. Die Stimuluszusammensetzung der Aufgabenblöcke kann für die kompatible und inkompatible Phase eines IAT identisch sein. Ein rein stimulusgetriebener Mechanismus wie Aktivierungsausbreitung kann die Effekte also prinzipiell nicht erklären

(Mierke & Klauer, 2001). Der entscheidende Unterschied zwischen kompatibler und inkompatibler Phase des IAT liegt ausschließlich in der instruierten Zuweisung der Antworttasten. Ein theoretisches Modell der Effekte im IAT muss also erklären, in welcher Weise die instruierte Tastenzuordnung Einfluss auf die Verarbeitung der präsentierten Reize nimmt.

Das Random-Walk Modell

Das Modell von Brendl, Markmann und Messner (2001) erklärt den IAT-Effekt durch die Kombination eines einfachen Random-Walk Modells der Informationsverarbeitung und die zusätzliche Annahme einer Verschiebung des Antwortkriteriums in der inkompatiblen Phase eines IAT. Die Autoren nehmen an, dass Information über die Attributausprägung des Stimulus und Information über die Zugehörigkeit zu einem der Zielkonzepte in einem Random-Walk Prozess auf einer Entscheidungsdimension akkumuliert werden. Das Modell nimmt an, dass eine Reaktion ausgelöst wird, wenn ein imaginärer Zähler eine bestimmte Antwortschwelle überschreitet. Die Netto-Akkumulationsrate der Information sollte für Exemplare der Zielkonzepte in der inkompatiblen Phase eines IAT niedriger ausfallen, weil Information bezüglich der Kategoriezugehörigkeit und Information bezüglich der Attributausprägung den Zähler in jeweils verschiedene Richtungen bewegen. In der kompatiblen Phase eines IAT bewegen Attributausprägung und Kategoriezugehörigkeit den Zähler hingegen in die gleiche Richtung, weshalb sich eine höhere Netto-Akkumulationsrate in Richtung der korrekten Antwort ergibt. Der IAT-Effekt ist nach diesem Modell das Resultat unterschiedlicher Akkumulationsraten in der kompatiblen und inkompatiblen Phase eines IAT. Brendl et al. (2001) nehmen jedoch weiter an, dass die Versuchsteilnehmer den Antworten in der inkompatiblen Phase ein konservativeres Antwortkriterium zugrunde legen. Als Konsequenz des konservativeren Antwortkriteriums ist eine Verlangsamung der Latenzen für Attribut- und Zielkonzeptreize zu erwarten. Der oben beschriebene einfache Random-Walk Prozess würde IAT-Effekte nur für Exemplare der Zielkonzepte vorhersagen.

Brendl et al. (2001) berichten zwei Experimente, die die Vorhersagen des Modells bestätigen. In der ersten Untersuchung bearbeiteten die Teilnehmer einen einfachen Blumen-Insekten-IAT (Greenwald et al., 1998). Um die Problematik der Relativität von IAT-Scores zu umgehen, wurden getrennte IAT-Scores für Durchgänge mit Blumen- und Insektennamen berechnet. Wie erwartet waren die Reaktionszeiten auf Blumennamen kürzer, wenn diese derselben Taste zugeordnet waren wie die positiven Wörter. Die Reaktionszeiten bei

Durchgängen mit Insektennamen waren, ebenfalls entsprechend der Erwartung, kürzer, wenn diese der gleichen Taste zugeordnet waren wie negative Wörter. Ferner traten Kompatibilitätseffekte auch für Exemplare der Attributkategorien auf. Dies wird von Brendl und Kollegen als Hinweis auf die Verschiebung des Antwortkriteriums gewertet.

In weiteren Experimenten sollte die Annahme geprüft werden, dass ähnliche Ergebnisse auftreten, wenn eines der Zielkonzepte durch evaluativ neutrale Non-Wörter ersetzt wird. Wenn die getrennt berechneten IAT-Scores tatsächlich die absolute Stärke der evaluativen Assoziationen abbilden, sollte der IAT-Effekt für Insekten unabhängig davon sein, ob die Vergleichskategorie Blumen oder Non-Wörter sind, in gleicher Form auftreten. Die Ergebnisse zeigen jedoch, dass die Reaktionszeiten bei den Insektennamen länger ausfallen, wenn diese mit negativen Wörtern gepaart sind und die Non-Wörter mit positiven, als wenn die Non-Wörter mit negativen Wörtern gepaart sind und die Insekten mit positiven. Es zeigten sich also ein deutlicher Effekt der Vergleichskategorie und eine Umkehrung der Ergebnisse des zuerst geschilderten Experiments. Dieser Befund unterstreicht die Relativität von IAT-Scores: Reize mit niedriger Geläufigkeit können ein Befundmuster zur Folge haben, dass auch als Hinweis auf implizite Vorurteile gedeutet werden kann. Abgesehen von der Tatsache, dass IAT-Effekte auch für Exemplare der Attributkategorien auftraten, liefert die Arbeit von Brendl et al. (2001) jedoch keine klare Prüfung der Annahmen des Modells.

Stimulus-Response Kompatibilität

Ein Modell, das den IAT-Effekt über eine bedingungsabhängige Kompatibilität von Stimulus und Response erklärt, wurde von De Houwer (2001b) vorgeschlagen. Es basiert auf der Annahme, dass die Antworttasten bei der Bearbeitung eines IAT eine temporäre evaluative Konnotation erlangen. De Houwer argumentiert, dass diese Konnotation in der kompatiblen Bedingung eines IAT eindeutiger ausfällt, da den Antworttasten Konzepte gleicher Valenz zugeordnet sind, während sie in der inkompatiblen Bedingung uneindeutig bleibt, da einer Antworttaste Konzepte unterschiedlicher Valenz zugeordnet sind. Den IAT-Effekt erklärt das Modell durch eine Kompatibilität zwischen den Reizen und der erworbenen Konnotation der Antworttasten. Das zentrale Element ist dabei eine im Vergleich zur inkompatiblen Phase größere Ähnlichkeit von Eigenschaften des Stimulus und Eigenschaften der Response in der kompatiblen Phase eines IAT. Dieser Mechanismus ähnelt den Mechanismen zur Erklärung von Simoneffekten (De Houwer, 1998, 2003; De Houwer & Eelen, 1998; Simon, 1990). Obwohl das Modell von De Houwer IAT-Effekte recht elegant

erklären kann, fehlt derzeit eine stringente Prüfung der Modellannahmen. Die von De Houwer berichteten Befunde zu Exempleffekten berühren lediglich periphere Annahmen des Modells. Die zentrale Annahme, dass der IAT auf einer Kompatibilität von Stimulus und Response beruht, kann auf dieser Basis nicht als bestätigt gelten.

Der Figur-Grund Ansatz

Der Figur-Grund Ansatz (Rothermund & Wentura, 2001) basiert auf der Annahme, dass IAT-Effekte durch Salienzunterschiede zwischen den verwendeten Kategorien vermittelt werden. Das zentrale Konzept des Ansatzes ist eine systematische Asymmetrie der Salienz innerhalb der Zielkonzept- und Attribut-Dimensionen. Es wird angenommen, dass jeweils die Exemplare einer der gegenübergestellten Kategorien salienter sind und infolgedessen als Figur wahrgenommen werden, während die Exemplare der anderen Kategorie weniger salient sind und als Grund wahrgenommen werden. Kompatibilität ergibt sich in diesem Zusammenhang immer dann, wenn die Konzepte, die einer Reaktion zugeordnet sind, bezüglich ihrer Salienz übereinstimmen. Das Modell geht davon aus, dass die beiden dichotomen Klassifikationsaufgaben in diesem Fall zu einer einzigen unipolaren Suchaufgabe reduziert werden, die auf Elemente der salienten Kategorie fokussiert ist. IAT-Effekte ergeben sich, da diese Rekodierung und die damit einhergehende Vereinfachung nur in der kompatiblen Phase eines IAT möglich sind. Damit macht der Figur-Grund Ansatz sowohl Annahmen darüber, was der IAT misst, nämlich Salienzunterschiede, als auch darüber, welcher Mechanismus die Effekte vermittelt.

Eine große Zahl verschiedener Ergebnisse kann durch das Figur-Grund Modell erklärt werden, wenn angenommen wird, dass Asymmetrien der Salienz mit Asymmetrien der Valenz oder Familiarität einhergehen. Die Ergebnisse eines typischen Einstellungs-IAT können etwa dadurch erklärt werden, dass Exemplare negativer Kategorien automatisch Aufmerksamkeit auf sich ziehen (Fox et al., 2000; Pratto & John, 1991) und somit als Figur wahrgenommen werden. Werden die negativ bewerteten Kategorien der gleichen Reaktion zugeordnet, kann die Aufgabenstruktur zu einer unipolaren Suchaufgabe reduziert werden. Ähnliches gilt für den Befund von Brendl et al. (2001), dass die Kontrastierung von Insekten und Non-Wörtern zu einer Umkehrung des IAT-Effekts führt. Dieser Effekt kann durch das Figur-Grund Modell erklärt werden, wenn angenommen wird, dass Non-Wörter im Vergleich zu Insekten mehr Aufmerksamkeit auf sich ziehen. In diesem Fall sollten die Non-Wörter als

Figur wahrgenommen werden, sodass eine Vereinfachung der Aufgabe bei der Paarung der Non-Wörter mit negativen Reizen möglich ist.

Das Figur-Grund Modell macht eine Reihe spezifischer Vorhersagen, die von Rothermund und Wentura (2001) experimentell geprüft wurden. Unter anderem sollten IAT-Effekte auch dann auftreten, wenn lediglich Figur-Grund Asymmetrien vorliegen, jedoch keine evaluativen oder andersartigen Assoziationen. In ihrem ersten Experiment verglichen die Autoren einen alt-jung-Einstellungs-IAT mit einem IAT, in dem die Bewertungsaufgabe durch eine Diskriminationsaufgabe zwischen neutralen Wörtern und Non-Wörtern ersetzt war. Die Autoren nahmen an, dass die Kategorien „alt“ und „Non-Wort“ salienter sind, als die gegenübergestellten Kategorien „jung“ und neutrale Wörter. Gemäß der Vorhersage des Figur-Grund Modells ergab sich ein IAT-Effekt. Die Performanz der Versuchsteilnehmer war deutlich besser, wenn die Kategorien „alt“ und „Non-Wort“ der gleichen Reaktion zugeordnet waren, als wenn „alt“ und neutrale Wörter der gleichen Reaktion zugeordnet waren. Dieser Effekt kann offenbar nicht durch eine Assoziation im semantischen Gedächtnis erklärt werden. In einem zweiten Experiment sollte die Figur-Grund Asymmetrie in einem alt-jung Einstellungs-IAT durch eine vorhergehende Versuchsphase manipuliert werden. Statt der üblichen Übungsphasen bearbeiteten die Versuchsteilnehmer eine go/no-go Aufgabe, in der eine Fokussierung auf die Kategorie „jung“ erzeugt werden sollte. Bei der kleinen Stichprobe von 8 Personen ergab sich in diesem Experiment eine Umkehrung des IAT-Effekts im Vergleich zu Experiment 1.

Die Daten von Rothermund und Wentura (2001) bestätigen die Annahmen des Figur-Grund Modells. Insbesondere in Bezug auf die Frage, was der IAT misst, weisen die Befunde klar darauf hin, dass ein strenges Assoziationsmodell, nachdem evaluative oder semantische Assoziation hinreichend *und* notwendig für das Auftreten von IAT-Effekten sind, nicht zutrifft. Nicht ausgeschlossen ist dagegen, dass semantische Assoziation oder andersartige Relationen zwischen Kategorien *hinreichend* sind, um IAT-Effekte zu produzieren. Die strenge Formulierung des Modells behauptet jedoch, dass Figur-Grund Asymmetrien eine *notwendige* Voraussetzung für das Zustandekommen von IAT-Effekten darstellen. Im Empirieteil der Arbeit soll deshalb versucht werden, IAT-Effekte zu zeigen, für die eine Erklärung über Figur-Grund Asymmetrien ausgeschlossen ist.

3.5 Das Aufgabenwechselmodell

Ein zentrales Merkmal des IAT ist, dass die Attribut- und die Zielkonzeptaufgabe in den entscheidenden Phasen gemischt bearbeitet werden müssen. Das in dieser Arbeit vorgeschlagene Modell basiert auf der Annahme, dass der IAT-Effekt auf kognitive Prozesse zurückgeht, die beim Wechsel dieser Aufgaben innerhalb eines Aufgabenblocks ablaufen. Um die Darstellung des Modells zu vereinfachen, sollen zunächst einige grundlegende Begriffe, Befunde und Theorien geschildert werden, die im Zusammenhang mit Prozessen des Aufgabenwechslens diskutiert werden.

Empirische Untersuchungen zu Aufgabenwechseln

In den letzten Jahren wurde eine Reihe von Untersuchungen publiziert, die sich mit exekutiven Kontrollprozessen beschäftigen, die ablaufen, wenn zwischen verschiedenen kognitiven Aufgaben gewechselt werden muss (Allport, Styles & Hsieh, 1994; Goschke, 2000; Koch, 2003; Meiran, 1996; Meiran, Chorev & Sapir, 2000; Rogers & Monsell, 1995). Viele Studien in diesem Bereich belegen, dass das Wechseln zwischen verschiedenen Aufgaben mit Kosten in Form verlängerter Reaktionszeiten und/oder größerer Fehlerzahlen verbunden sind (Allport et al., 1994; Jersild, 1927; Mayr & Keele, 2000; Rogers & Monsell, 1995; Spector & Biederman, 1976; Wylie & Allport, 2000), die im Vergleich zu Versuchsbedingungen ohne Aufgabenwechsel auftreten.

Zur Untersuchung von Prozessen des Aufgabenwechsels wurde eine Reihe verschiedener Methoden entwickelt. Eine von Jersild (1927) eingeführte Methode besteht darin, reine und gemischte Blöcke von einfachen Aufgaben zu vergleichen (Allport et al., 1994; Jersild, 1927; Spector & Biederman, 1976). In einem gemischten Block könnten zum Beispiel Entscheidungen bezüglich der Parität von Zahlen und Entscheidungen bezüglich der Lautklasse von Buchstaben (Vokal versus Konsonant) im Wechsel gefordert werden (Rogers & Monsell, 1995, Exp. 3). In reinen Aufgabenblöcken bearbeiten die Versuchsteilnehmer bei jedem Durchgang dieselbe Aufgabe, während die zu bearbeitende Aufgabe in gemischten Blöcken von Durchgang zu Durchgang wechselt. In anderen Studien wurde ein Paradigma mit so genannten Alternating Runs (Rogers & Monsell, 1995) verwendet, in dem die Aufgabe bei jedem n-ten Durchgang vorhersagbar wechselt. Ein Vorteil dieses Paradigmas ist, dass Aufgabenwiederholungsdurchgänge (bei denen die Aufgabe des aktuellen und des vorherigen Durchgangs identisch sind) und Aufgabenwechseldurchgänge (bei denen die Aufgabe zwischen aktuellem und vorherigem Durchgang wechselt) innerhalb eines Blocks realisiert

und untersucht werden können. In einer dritten Art von Aufgabenwechselexperimenten wechselt die durchzuführende Aufgabe innerhalb eines Aufgabenblocks zufällig. In diesem Paradigma werden die einzelnen Durchgänge durch so genannte Aufgabenhinweise eingeleitet, die die Versuchsteilnehmer über die durchzuführende Aufgabe instruieren (Meiran, 1996; Sudevan & Taylor, 1987). In den Studien von Meiran (1996) zeigten sich die Aufgabenwechselkosten im Vergleich zu einer Versuchsbedingung mit simultaner Darbietung von Aufgabenhinweis und imperativem Stimulus verringert, wenn die Aufgabenhinweise vor dem zu bearbeitenden Reiz präsentiert wurden. Wie in den Arbeiten von Rogers und Monsell (1995) verbleiben allerdings auch in diesem Paradigma residuale Wechselkosten, die selbst bei einem vergleichsweise langen Vorbereitungsintervall von 600 ms auftreten, und durch eine Verlängerung des Intervalls nicht weiter verringert werden können.

Aufgabenwechselkosten und exekutive Kontrollprozesse

Während der Basisbefund, dass Aufgabenwechsel mit Kosten verbunden sind, empirisch stabil auftritt, herrscht Uneinigkeit darüber, welche kognitiven Prozesse diesen Kosten zugrunde liegen. Einige Autoren gehen davon, dass Aufgabenwechselkosten - im Sinne einer subtraktiven Logik - als Maß für die Dauer eines exekutiven Kontrollprozesses interpretiert werden können (Goschke, 2000; Jersild, 1927; Logan & Gordon, 2001; Meiran, 1996; Rogers & Monsell, 1995; Rubinstein, Meyer & Evans, 2001), der unter anderem die Aktivierung des so genannten Aufgabensets umfasst. Die Bezeichnung Aufgabenset bezieht sich dabei auf die Menge kognitiver Operationen und Zustände, die nötig ist, eine Aufgabe effektiv durchzuführen. Konkreter umfasst die Bezeichnung Aufgabenset die Fokussierung auf ein relevantes Stimulusattribut, die möglichen Kategorisierungen dieses Stimulusattributs, die verfügbaren Antwortmodi und -optionen, sowie vorgegebene Verbindungen zwischen Stimulus- und Antwortklassen (Monsell, Yeung & Azuma, 2000).

Ein Argument für die Beteiligung exekutiver Kontrollprozesse am Zustandekommen von Aufgabenwechselkosten ist dabei, dass Wechselkosten durch informierte Vorbereitung reduziert werden können. Meiran (1996) konnte wie schon angesprochen nachweisen, dass Aufgabenwechselkosten durch die Präsentation von Aufgabenhinweisen im Vorbereitungsintervall reduziert werden können. Die Versuchsteilnehmer sollten in diesen Experimenten entscheiden, ob ein dargebotener Stimulus in der oberen oder unteren beziehungsweise der linken oder rechten Hälfte einer einfachen 2 x 2 Matrix erschien. Die auszuführende Aufgabe wechselte dabei zufällig von Durchgang zu Durchgang. Die Aufgabenhinweise bestanden aus

zwei Pfeilen, die entweder nach links und rechts oder nach oben und unten wiesen. Die Hinweise erschienen entweder während des Vorbereitungsintervalls auf den nächsten Reiz oder simultan mit diesem. Die Länge des Vorbereitungsintervalls selbst blieb für alle Durchgänge konstant. Solche symbolischen Aufgabenhinweise, die die auszuführende Aufgabe durch aufgabenspezifische Reize ankündigen, werden als exogene Hinweise bezeichnet (Koch, 2003), in Abgrenzung zu so genannten endogenen Hinweisen, bei denen die Aufgabe auf Basis einer festen Aufgabensequenz vorhergesagt werden kann (z. B. Koch, 2003; Rogers & Monsell, 1995). Der zentrale Befund ist, dass eine Präsentation der Aufgabenhinweise im Vorbereitungsintervall zu einer statistisch bedeutsamen Reduktion der Aufgabenwechselkosten im Vergleich zu der Bedingung mit simultaner Darbietung führte. Spiegelten die Aufgabenwechselkosten ausschließlich passive Nachwirkungen der vorherigen Aufgabe wider (siehe unten), wären in beiden Bedingungen gleiche Aufgabenwechselkosten zu erwarten.

Allport und Kollegen (Allport et al., 1994; Wylie & Allport, 2000) bezweifeln jedoch, dass eine einfache subtraktive Logik empirisch angemessen ist, um Aufgabenwechselkosten zu dekomponieren. Sie argumentieren, dass Aufgabenwechselkosten neben dem Ablaufen eines exekutiven Kontrollprozesses auch andere psychologische Vorgänge widerspiegeln. So gehen Allport et al. (1994) davon aus, dass Aufgabenwechselkosten vor allem Interferenzen zwischen dem aktuellen Aufgabenset und kürzlich aktivierten alternativen Aufgabensets reflektieren. Diese Annahme wird unter anderem dadurch gestützt, dass das Wechseln von einer schwierigen, wenig eingeübten Aufgabe zu einer leichteren, besser eingeübten Aufgabe (etwa der Wechsel von Farbnennung zum Wortlesen in einer Stroop-Aufgabe) üblicherweise *höhere* Wechselkosten zur Folge hat, als das Wechseln von einer leichten zu einer schwierigen Aufgabe (Allport et al., 1994). Dieses Muster von Wechselkosten steht im Widerspruch zu dem, was zu erwarten wäre, wenn ein exekutiver Kontrollprozess die Kosten allein verursachte, denn die Etablierung einer weniger gut eingeübten Aufgabe sollte kognitiv aufwändiger sein, mehr exekutive Kontrolle erfordern, und folglich mehr Zeit in Anspruch nehmen. Allport und Kollegen erklären diesen Effekt durch eine Trägheit des vorher aktivierten Aufgabensets, die zu einer proaktiven Interferenz bei der Aktivierung des neuen Aufgabensets führt. Das Trägheitsmodell kann Aufgabenwechselkosten folglich ohne die Annahme exekutiver Kontrollprozesse erklären. Andere Studien (Mayr & Keele, 2000) haben jedoch gezeigt, dass die Hemmung kürzlich aktivierter Aufgabensets exekutive Kontrolle erfordern kann, sodass der Befund proaktiver Interferenz nicht zwingend bedeutet, dass

exekutive Kontrollprozesse nicht am Zustandekommen von Aufgabenwechselkosten beteiligt sind.

Eine weitere Hypothese, die Aufgabenwechselkosten ohne direkte Beteiligung exekutiver Kontrollprozesse erklären kann, wurde von Wylie und Allport (2000) vorgeschlagen. Basierend auf dem Befund, dass ihre Versuchsteilnehmer die Farbwörter in einer Stroop-Aufgabe langsamer vorlasen, wenn die Wörter in einer vorherigen Phase des Experiments benannt werden mussten, folgerten sie, dass die Übung beim Vorlesen der Farbwörter zu einem automatischen Priming des entsprechenden Aufgabensets führt. Diese Übung kann den Effekt verlangsamter Farbnennung erklären, da durch Priming des Aufgabensets eine stärkere Interferenz mit der Farbnennung zu erwarten ist. Ob diese Interferenz passiv abklingt, wie Wylie und Allport (2000) vermuten, oder durch aktive exekutive Kontrollprozesse abgebaut wird (Mayr & Keele, 2000), ist gegenwärtig nicht geklärt. Eine jüngere Arbeit von Hsieh (2002) vertritt die Ansicht, dass beim Aufgabenwechseln sowohl passive Trägheitseffekte als auch aktive exekutive Kontrollprozesse beteiligt sind.

Interindividuelle Unterschiede beim Aufgabenwechsel

Eine Reihe von Arbeiten beschäftigt sich mit interindividuellen Unterschieden in Fähigkeitsparametern, die an Aufgabenwechselprozessen beteiligt sind. Diese Untersuchungen zeigen unter anderem, dass Aufgabenwechsel von jüngeren Versuchsteilnehmern schneller durchgeführt werden können als von älteren (Cepeda, Kramer & Gonzalez de Sather, 2001; Kramer, Hahn & Gopher, 1999; Kray & Lindenberger, 2000). Diese altersabhängige Varianz in den Aufgabenwechselkosten scheint zumindest partiell unabhängig von der Varianz anderer kognitiver Variablen wie der perzeptuellen Geschwindigkeit oder der Kapazität des Arbeitsgedächtnisses (Cepeda et al., 2001). Es zeigen sich jedoch auch systematische Zusammenhänge mit anderen kognitiven Fähigkeiten wie etwa der fluiden Intelligenz (Kray & Lindenberger, 2000). Bezüglich der Altersabhängigkeit ergibt sich über die Lebensspanne ein u-förmiger Zusammenhang mit der Leistungsfähigkeit beim Aufgabenwechseln, mit einer vergleichsweise geringeren Leistung von Kindern und älteren Menschen. Insbesondere ältere Menschen scheinen beim Wechseln zwischen Aufgaben von Übung zu profitieren. Dieser Effekt bleibt offenbar über einen Zeitraum von mehreren Monaten erhalten (Kramer et al., 1999). Bei den älteren Teilnehmern ergab sich, anders als bei Kindern, ein bedeutsamer Einfluss des Intervalls zwischen Antwort und Präsentation des nächsten

Aufgabenhinweises. Dies kann als Hinweis darauf interpretiert werden, dass passive Interferenz mit dem vorherigen Aufgabenset bei älteren Menschen schneller abklingt (Cepeda et al., 2001). Für die gegenwärtige Arbeit ist jedoch der grundlegende Befund von besonderem Belang, dass die Leistung beim Wechseln von Aufgaben in starkem Maße auf kognitive Fähigkeitsparameter zurückgeht, die sich systematisch zwischen Menschen unterscheiden.

Aufgabenwechselkosten im IAT

Wie schon angesprochen, lautet eine zentrale Annahme des im Folgenden vorgestellten Modells, dass der IAT-Effekt auf kognitive Prozesse zurückgeht, die mit dem Wechsel zwischen der Attribut- und der Zielkonzeptaufgabe innerhalb der kompatiblen und inkompatiblen Bedingung eines IAT in Zusammenhang stehen. Nach der Einführung einiger Kernannahmen soll das eigentliche Modell genauer dargestellt und eine Reihe von Hypothesen abgeleitet werden, die sich aus dem Modell ergeben.

Unterschiede zu klassischen Kompatibilitätseffekten

Im Gegensatz zu klassischen Kompatibilitätseffekten wie dem affektiven Priming oder Stroop-Effekt beruht der IAT nicht auf einer Manipulation der Kompatibilität von in einem Durchgang zusammentreffender relevanter und irrelevanter Stimulusinformation. Der IAT-Effekt kann also nicht auf einem direkten Konflikt von relevanten und irrelevanten Stimuluseigenschaften auf inhaltlicher Ebene beruhen, wie etwa der Konflikt zwischen Wortbedeutung und Wortfärbung im Stroop-Paradigma (MacLeod, 1991; Stroop, 1935) oder der Konflikt zwischen der Valenz von Prime- und Targetwort im affektiven Priming (Draine & Greenwald, 1998; Fazio et al., 1986; Klauer & Musch, 2001, 2003). Im IAT konkurrieren vielmehr eine relevante und eine irrelevante Aufgabenspezifikation um die Anwendung auf einen gegebenen Stimulus. Die Reize selbst sind in einem IAT weder intrinsisch kompatibel noch inkompatibel, sondern nur in Zusammenhang mit der Gesamtmenge der Tastenzuordnungen eines Aufgabenblocks. Kompatibilität ergibt sich erst aus dem Zusammenhang von *Reizen und Aufgabenspezifikation*.

Wie Brendl et al. (2001) ausführen, weisen die Exemplare der Zielkonzepte irrelevante Attributinformation auf, nicht aber die Exemplare der Attributkategorien eine Zugehörigkeit zu einem der Zielkonzepte. So sind die Blumen eines Blumen-Insekten Einstellungs-IAT positiv, nicht aber die positiven Reize blumig. Dadurch ist prinzipiell die Attributaufgabe auf Exemplare der Zielkonzepte anwendbar, nicht jedoch umgekehrt. Eine zentrale Annahme des

Aufgabenwechselmodells ist, dass die Aufgabensets der Attribut- und Zielkonzeptaufgabe bei der Bearbeitung von Zielkonzeptreizen in Konflikt treten. In der inkompatiblen Phase eines IAT führt dieser Konflikt zu verlangsamter Performanz, in der kompatiblen zu Erleichterung. Die konkreten kognitiven Prozesse, die dieser Erleichterung beziehungsweise Erschwerung zugrunde liegen, sind aus Sicht des Aufgabenwechselmodells im weiteren Sinne exekutive Kontrollprozesse.

Trägheit von Aufgabensets im IAT

In Anlehnung an das Aufgabenwechselmodell von Allport et al. (Allport et al., 1994; Wylie & Allport, 2000) wird davon ausgegangen, dass die Anwendung eines Aufgabensets dessen Aktivierung sowie die Unterdrückung eventuell konkurrierender Aufgabensets umfasst. Die entscheidende Annahme des Modells von Allport et al. ist, dass einmal aktivierte Aufgabensets eine Trägheit (Inertia) aufweisen. Aus Sicht des Modells stellen Aufgabenwechselkosten vor allem Kosten für die Überwindung störender Nachwirkungen des Vorgängerdurchgangs dar. Dabei lassen sich zwei Komponenten unterscheiden: Einerseits erleichternde Nachwirkungen bei der Wiederanwendung einer Aufgabe durch die residuale Aktivierung des entsprechenden Aufgabensets; andererseits störende Nachwirkungen bei der Anwendung einer Aufgabe, die mit der aktuellen Aufgabe interferiert. Im Falle einer Interferenz von Aufgabensets wird angenommen, dass die konfligierende Aufgabenrepräsentation unterdrückt werden muss, sodass ihre Anwendung in nachfolgenden Durchgängen erschwert ist, ein negativer Primingeffekt.

Eine Übertragung dieser Annahmen auf den IAT führt zu einer Reihe interessanter Hypothesen über sequentielle Effekte. So sind in der kompatiblen Phase eines IAT insgesamt keine Interferenzen von Aufgabensets zu erwarten. Damit sollten sich eventuelle Nachwirkungen vorher bearbeiteter Aufgaben nicht störend auswirken. Folglich sind also weder Reaktionszeiteinbußen bei Aufgabenwechseln noch negatives Priming von Aufgabensets zu erwarten. In der inkompatiblen Phase eines IAT sind im Gegensatz dazu Interferenzen der Aufgabensets zu erwarten: Die residuale Aktivierung des Attributaufgabensets sollte sich störend auswirken, wenn ein Exemplar der Zielkonzepte bearbeitet wird. Soll zum Beispiel in einem Durchgang ein negativ besetzter Reiz wie „Wespe“ als Insekt klassifiziert werden, sollte die residuale Aktivierung des Bewertungsaufgabensets mit dieser Klassifikation interferieren, sofern im Vorgängerdurchgang auf negative Valenz mit einer anderen Antworttaste reagiert werden musste, als der zur Klassifikation erforderlichen. Zur Auflösung der

entstehenden Interferenz muss das Bewertungsaufgabenset unterdrückt werden. Bei Wechseln von der Zielkonzeptaufgabe zur Attributaufgabe sind keine Interferenzen zu erwarten. Eine im Vorgängerdurchgang erforderliche Hemmung des Attributaufgabensets sollte jedoch im Sinne eines negativen Primingeffekts nachwirken. Damit sind in der inkompatiblen Phase eines IAT sowohl Kosten für den Wechsel von der Attribut- zur Zielkonzeptaufgabe als auch in umgekehrter Richtung zu erwarten. In der kompatiblen Bedingung eines IAT sind hingegen keine Aufgabenwechselkosten zu erwarten. Dieses Muster von Sequenzeffekten - also das Ausbleiben von Aufgabenwechselkosten in der kompatiblen Bedingung und ihr Auftreten in der inkompatiblen - wird im Folgenden als *Asymmetrie der Aufgabenwechselkosten* bezeichnet.

Abgrenzung zu einfachen Sequenzeffekten

Bei Aufgabenwiederholung, nicht jedoch bei Aufgabenwechsel, können einzelne Durchgänge die Anwendung der gleichen Kategorie-Reaktions-Verknüpfung erfordern. Es ist zu erwarten, dass dies einen einfachen Sequenzeffekt zur Folge hat, der auf eine schnellere Wiederanwendung der gleichen Verknüpfung zurückgeht (Pashler & Baylis, 1991a, 1991b). Zwar stellt diese Art einfacher Sequenzeffekte in gewisser Hinsicht eine Komponente des Aufgabenwechselprozesses dar, gleichzeitig kann ein einfacher Wiederholungseffekt jedoch deutlich sparsamer erklärt werden. Eine weitere zu prüfende Hypothese ist deshalb, dass der Effekt asymmetrischer Aufgabenwechselkosten über den Effekt einer einfachen Wiederholung der gleichen Kategorie-Reaktions-Verknüpfung hinausgeht. Wenn Aktivierung und Hemmung auf das gesamte Aufgabenset wirken, sollten Wechselkosten auch dann auftreten, wenn keine spezifische Verknüpfung wiederholt wird. Es sollte also auch dann einen Vorteil für die Anwendung einer Kategorie-Reaktions-Verknüpfung geben, wenn dem aktuellen Durchgang die Anwendung einer *anderen S-R Verknüpfung des gleichen Aufgabensets* vorausging. Wenn Hemmung und Aktivierung von Aufgabensets im zeitlichen Verlauf passiv zerfallen, wie Allport et al. (1994) vermuten, sollte diese Asymmetrie der Wechselkosten ferner geringer ausfallen, je größer das zeitliche Intervall zwischen den einzelnen Durchgängen ausfällt. Diese Hypothese wird in Experiment 1 geprüft.

Kompatibilitätseffekt bei der Attributaufgabe

Aus den oben ausgeführten Annahmen folgt, dass IAT-Effekte auch für die Reize der Attributaufgabe auftreten sollten. In einem einfachen Random-Walk Modell wäre dies nicht zu erwarten, da nur die Exemplare der Zielkonzepte ein interferenzträchtiges, irrelevantes

Merkmal aufweisen (Brendl et al., 2001). Entsprechende Befunde von Brendl und Kollegen belegen jedoch, dass dies der Fall ist. Im Rahmen des Random-Walk Modells wird dieser Effekt durch eine Verschiebung des Antwortkriteriums in der inkompatiblen Phase eines IAT erklärt. Dies sollte jedoch zu einer linearen Transformation der Latenzen für Zielkonzept- und Attributreize führen, das heißt, zu einem stärkeren Kompatibilitätseffekt bei den Zielkonzeptreizen. Eine Erklärung dieses Effekts über negatives Priming von Aufgabensets, wie sie hier vorgeschlagen wird, sagt allgemein ein Ansteigen der Latenzen für die Attribut-aufgabe nach inkompatiblen Zielkonzeptreizen vorher. Dieser Effekt ist auch dann zu erwarten, wenn kompatible und inkompatible Zielkonzeptreize innerhalb eines Aufgabenblocks vermischt bearbeitet werden. Eine lineare Abhängigkeit der Latenzen ist hingegen nicht zu erwarten. Diese Hypothese wird in Experiment 3 geprüft.

Vereinfachungs- und Interferenzmechanismus

Zusammenfassend erklärt das hier vorgeschlagene Modell den IAT-Effekt also über die Nachwirkung von Aktivierungs- und Hemmungsprozessen von Aufgabensets, die in der kompatiblen und inkompatiblen Phase eines IAT jeweils unterschiedliche Auswirkungen haben. Dieser Mechanismus soll im Folgenden als *Interferenzmechanismus* bezeichnet werden. Eine einfache Alternativerklärung liefert der so genannte *Vereinfachungsmechanismus* (Mierke & Klauer, 1999, 2001). Die Kernannahme dieses Vereinfachungsmechanismus ist, dass im Lauf der Bearbeitung eines kompatiblen IAT-Blocks eine vereinfachte Aufgabe erlernt wird, bei der lediglich aufgrund der Attributausprägungen geantwortet wird. Die erleichterte Bearbeitung der kompatiblen Bedingung eines IAT würde hiernach auf das Ausbleiben eines aufwändigeren exekutiven Kontrollprozesses (Meiran, 1996; Rogers & Monsell, 1995) zurückgehen. In der inkompatiblen Bedingung ist eine solche Vereinfachung nicht möglich. Die Reaktionszeiten in dieser Bedingung sollten folglich sowohl durch eine spezifische Kostenkomponente, die auf exekutive Kontrollprozesse beim Wechseln der Aufgaben zurückgehen, als auch durch so genannte Mischungskosten (Los, 1996) beeinträchtigt werden, die durch eine höhere Belastung des Arbeitsgedächtnisses bei der Bearbeitung mehrerer Aufgaben zustande kommen.

Aus theoretischer Sicht liegt der Unterschied zwischen dem Interferenz- und dem Vereinfachungsmechanismus vor allem darin, dass letzterer das Ausbleiben von Aufgabenwechselkosten in der kompatiblen Phase über das Ausbleiben der entsprechenden Kontrollprozesse erklärt, während ersterer davon ausgeht, dass die Kontrollprozesse zwar ablaufen,

aber mit deutlich geringeren Kosten verbunden sind (siehe oben). Die zentrale Vorhersage einer Asymmetrie der Aufgabenwechselkosten ergibt sich aber aus beiden Mechanismen, wenn auch aufgrund unterschiedlicher Annahmen. Eine entscheidende Voraussetzung für den Vereinfachungsmechanismus ist, dass die kompatible und inkompatible Bedingung eines IAT in getrennten Phasen realisiert werden, denn die Vereinfachung setzt voraus, dass die Mehrzahl der Reize mit der vereinfachten Aufgabenstruktur korrekt klassifiziert werden kann. Der Interferenzmechanismus setzt dies nicht voraus, da der IAT-Effekt auf lokale Sequenzeffekte zurückgeführt wird, die - wenn auch in schwächerer Form – ebenso dann auftreten, wenn kompatible und inkompatible Zuordnung innerhalb eines Aufgabenblocks gemischt bearbeitet werden. Einen kritischen Test dieser Vorhersage liefert Experiment 3. Einige weitere Unterschiede zwischen den Vorhersagen des Vereinfachungs- und des Interferenzmechanismus werden in der Diskussion der einzelnen Experimente aufgezeigt und geprüft.

Was misst der IAT aus Sicht des Aufgabenwechselmodells?

Neben der Vorhersage asymmetrischer Aufgabenwechselkosten ist eine entscheidende Gemeinsamkeit beider Modelle, dass IAT-Effekte über einen Mechanismus erklärt werden, der Effekte prinzipiell mit einer großen Zahl verschiedenartiger Zusammenhänge oder Redundanzen erwarten lässt (vgl. auch Brendl et al., 2001; Fiedler et al., 2003). Keineswegs legt die Struktur des IAT - als experimentelles Design betrachtet - die Art der erfassten Relationen auf assoziative Verknüpfungen oder gar evaluative Assoziationen fest. Vielmehr sollten IAT-Effekte immer dann auftreten, wenn eine Aufgabe das Reagieren auf Stimulusattribute erfordert, die in einer anderen - im Wechsel zu bearbeitenden Aufgabe - zu störenden Interferenzen führen können. Solche Interferenzen können offenkundig auch dann auftreten, wenn das interferierende Attribut lediglich durch experimentelle Manipulation mit einem dargebotenen Reiz in Zusammenhang gebracht wird, nicht jedoch assoziativ verknüpft ist. Färbt man in einem fiktiven Experiment zum Beispiel einen Teil der Reize, bezüglich derer eine evaluative Entscheidung gefordert ist, rot ein und fordert in einigen Durchgängen eine Farbdiskrimination bezüglich anderer Reize, so sollte dies zu Interferenzen zwischen der Färbung der Reize und der evaluativen Entscheidung führen.

Bedeutsam für diesen Mechanismus ist offenkundig nicht, welche Diskriminationsaufgaben instruiert wurden, sondern nur, welche Aufgabe von den Versuchsteilnehmern ausgeführt wird. Es wäre durchaus denkbar, dass ein bezüglich der Instruktion vollständig aufgabenirrelevantes Merkmal zu dieser Art von Interferenzen führt. Färbt man etwa bei einer

evaluativen Entscheidungsaufgabe alle positiven Reize blau ein und alle negativen rot, so könnten die Versuchsteilnehmer dazu übergehen, die evaluative Entscheidung durch eine Farbdiskrimination zu ersetzen. Mischt man nun diese evaluative Entscheidung mit einer anderen Aufgabe, etwa der Diskrimination zwischen den Namen farbiger und weißer Prominenter, die ebenfalls blau oder rot gefärbt sind, so wären ebenfalls Interferenzeffekte zu erwarten. Aus dieser Perspektive stellt das Figur-Grund Modell von Rothermund und Wentura (2001) eine spezielle Variante des Aufgabenwechselmodells dar: Die Versuchsteilnehmer vereinfachen die Aufgabenstruktur anhand eines dritten, konfundierten Merkmals, der Salienz der Reizkategorien und ihrer Exemplare. Effekte von Figur-Grund Asymmetrien sind also durchaus mit dem hier vorgeschlagenen Ansatz vereinbar, werden jedoch nicht spezifisch vorhergesagt. Eine kritische Prüfung dieser Annahmen sowie eine Abgrenzung zum Figur-Grund Modell liefern die Experimente 2 und 3.

Effekte von Aufgabenhinweisen

Folgt man den geschilderten Ergebnissen zu exekutiven Kontrollprozessen, sollte die Präsentation von Aufgabenhinweisen und die Einführung einfacher und vorhersagbarer Aufgabensequenzen zu einer Verringerung von Aufgabenwechselkosten führen (Koch, 2003; Meiran, 1996; Rogers & Monsell, 1995). Übertragen auf den IAT sollte eine Verringerung der Aufgabenwechselkosten zu einer Verringerung der IAT-Effekte selbst führen, da Aufgabenwechselkosten in der inkompatiblen, nicht jedoch der kompatiblen Phase zu erwarten sind. Damit sollte eine entscheidende Kostenkomponente wegfallen, die für den Unterschied zwischen kompatibler und inkompatibler Phase verantwortlich ist. In den Experimenten 7 und 8 soll diese Vorhersage geprüft werden.

Methodenvarianz

Aus der Literatur ist wie schon angesprochen bekannt, dass Aufgabenwechselkosten interindividuelle Differenzen aufweisen (Cepeda et al., 2001; Kramer et al., 1999; Kray, Li & Lindenberger, 2002; Kray & Lindenberger, 2000). Es wird deshalb angenommen, dass interindividuelle Unterschiede in der Geschwindigkeit mit der zwischen verschiedenen Aufgaben gewechselt werden kann, IAT-Effekte beeinflussen. Wenn diese Effekte stabil und materialunabhängig sind, sollten IATs auch dann korreliert sein, wenn sie komplett unkorrelierte Merkmale erfassen. Methodenspezifische Varianz könnte erklären, warum die internen Konsistenzen und Reliabilitäten von IATs häufig vergleichsweise hoch ausfallen, während konvergente oder kriteriums-basierte Schätzungen der Validität gering ausfallen.

Während bei der Berechnung von internen Konsistenzen und Reliabilitäten nämlich alle eingehenden Scores in gleichem Maße durch methodenspezifische Varianz belastet sind, wäre bei Schätzungen der Validität nur eines der verwendeten Maße, nämlich der IAT, belastet. Aus dieser Perspektive wird deutlich, dass zwischen Truescore-Varianz, im Sinne der gemeinsamen Varianz von Testhälften oder Paralleltests, und inhaltlich gültiger Varianz streng unterschieden werden muss. Die in diesem Zusammenhang bedeutsamen Hypothesen werden in den Experimenten 9 bis 12 geprüft und genauer dargelegt.

Längerfristige Trägheitseffekte und Auswirkungen der Bearbeitung eines IAT

In einigen Untersuchungen von Allport et al. (1994) bearbeiteten die Versuchsteilnehmer Stroop-Aufgaben, bei denen zwischen dem Nennen der Farbe und dem Vorlesen des Farbwortes zu wechseln war. Dabei zeigte sich eine stärkere Interferenz durch die irrelevante Stimuluseigenschaft, wenn diese in einem kürzlich angewendeten Aufgabenset relevant war. Interessanterweise traten solche Nachwirkungen über einen verhältnismäßig langen Zeitraum hinweg auf (Allport et al., 1994; Wylie & Allport, 2000). Unter anderem fanden sich in diesem Paradigma so genannte Reverse-Stroop-Effekte, bei denen die Färbung der Reize deutlich stärker mit der Benennung der Farbe interferierte, wenn in den vorhergehenden Durchgängen die Färbung der Reize benannt werden sollte. In Bezug auf den IAT lassen diese Befunde erwarten, dass die Bearbeitung kompatibler und inkompatibler IAT-Blocks ebenfalls längerfristige Effekte haben könnte. Wenn bei der Bearbeitung von inkompatiblen Blocks über mehrere Durchgänge hinweg die Attributausprägung der Zielkonzeptexemplare unterdrückt werden muss, wäre etwa zu erwarten, dass sich eine Nachwirkung dieser Unterdrückung auf nachfolgende Aufgabenbearbeitung auswirken kann. Unter anderem könnten solche längerfristigen Nachwirkungen Effekte der Kompatibilitätsreihenfolge im IAT erklären, bei denen meist die Bearbeitung der kompatiblen Phase nach einer vorhergehenden inkompatiblen verlangsamt ist. Ebenso könnten solche Effekte erklären, warum sich bei Testwiederholungen häufig geringere IAT-Effekte ergeben, als bei der ersten Testung. Diese Hypothese wird in den Experimenten 5 und 6 geprüft.

4 Experimente zu lokalen Sequenzeffekten

Im Empirieteil der Arbeit soll mit experimentellen und differentialpsychologischen Mitteln versucht werden, die im vorhergehenden Teil der Arbeit abgeleiteten Hypothesen zu prüfen. Die berichteten Experimente lassen sich zu verschiedenen Themenblöcken zusammenfassen. Zunächst soll eine Reihe von Experimenten berichtet werden, die den IAT aus eher allgemeinpsychologischer Perspektive beleuchtet. Dabei stehen lokale Sequenzeffekte im IAT (Experimente 1 bis 3), globale Effekte der Bearbeitung von IATs (Experimente 4 bis 6) und Effekte von Aufgabenhinweisen auf den IAT im Fokus (Experimente 7 und 8). Auch die Konsequenzen der Befunde für die angesprochenen theoretischen Modelle des IAT-Effekts werden jeweils an geeigneter Stelle diskutiert. Im Anschluss an diesen allgemeinpsychologisch orientierten Teil soll eine Serie von Experimenten geschildert werden, die sich mit den spezifischen Folgen dieser Ergebnisse für differentialpsychologische Fragestellungen befasst (Experimente 9 bis 12).

Durch die im Folgenden dargestellten Experimente 1 bis 3 zu lokalen Sequenzeffekten sollen zunächst die Hypothesen geprüft werden, die sich auf Effekte beim Übergang zwischen einzelnen Durchgängen innerhalb eines Aufgabenblocks beziehen. Dabei steht die Frage im Mittelpunkt, wie der Wechsel zwischen Attribut- und Zielkonzeptaufgabe in der kompatiblen und inkompatiblen Phase eines IAT mit Aufgabenwechselkosten in Zusammenhang steht. Lokal sind diese Effekte insofern, als die Antwortlatenzen als Funktion des jeweils aktuellen und des direkt vorhergehenden Durchgangs betrachtet werden, nicht jedoch als Funktion des globalen Kontexts eines Aufgabenblocks.

4.1 Experiment 1

In Experiment 1 sollen die zentralen Hypothesen zu lokalen Sequenzeffekten geprüft werden. Wie weiter oben schon ausgeführt wurde, sieht die Standardprozedur des IAT (Greenwald et al., 1998) eine strikt alternierende Sequenz von Konzept- und Attributaufgabe vor. Um die Vorhersagen des Aufgabenwechselmodells zu prüfen, wurde diese Prozedur in Experiment 1 durch zusätzliche Durchgänge erweitert, bei denen sich die zu bearbeitende Aufgabe wiederholt. Durch den Vergleich von Aufgabenwechsel- und Wiederholungsdurchgängen innerhalb der Aufgabenblöcke können so spezifische Aufgabenwechselkosten erfasst werden (z. B. Kray & Lindenberger, 2000; Meiran, 1996). Gemäß den Annahmen des Aufgabenwechselmodells ergibt sich für diesen experimentellen Faktor die Vorhersage, dass

Performanzunterschiede zwischen kompatibler und inkompatibler Tastenzuordnung bei Wechseln zwischen Attribut- und Kategorisierungsaufgabe maximal sind.

Effekt des RSI

Ferner soll in Experiment 1 eine zentrale Annahme des Trägheitsmodells von Allport und Kollegen (Allport et al., 1994; Wylie & Allport, 2000) geprüft werden. Wie weiter oben beschrieben, geht dieses Modell davon aus, dass Aufgabenwechselkosten primär proaktive Interferenzen durch die vorhergehend bearbeiteten Aufgabensets abbilden, nicht jedoch die Dauer eines exekutiven Kontrollprozesses (Meiran, 1996; Rogers & Monsell, 1995). Diese Modellvorstellungen führen zu unterschiedlichen Vorhersagen für Effekte einer Verlängerung des Response-Stimulus Intervalls (RSI). Während das Trägheitsmodell von Allport et al. (1994) davon ausgeht, dass proaktive Interferenz im Verlauf eines ausreichend langen RSI zumindest partiell abklingen kann, sagt letztere Modellvorstellung vorher, dass Aufgabenwechsel nur dann von einem verlängerten RSI profitieren, wenn die im Folgenden zu bearbeitende Aufgabe während des Intervalls bereits bekannt ist, z.B. durch die Präsentation eines Aufgabenhinweises (Meiran, 1996). Um diese differentiellen Vorhersagen zu prüfen, wurde in Experiment 1 das RSI in zwei Stufen (100 ms versus 1000 ms) manipuliert. Durch die zufällige Abfolge der Aufgaben in den gemischten Blöcken ist ausgeschlossen, dass ein exekutiver Kontrollprozess während des RSI die Bearbeitung der folgenden Aufgabe vorbereiten kann, sodass eine Interaktion zwischen der Länge des RSI und der Wechselkostenasymmetrie auf passiv abklingende proaktive Interferenz sensu Allport zurückgehen muss.

Wiederholung von Reiz-Reaktionsverknüpfungen

Tritt die vorhergesagte Wechselkostenasymmetrie in Experiment 1 auf, soll ferner geprüft werden, ob der Effekt auf die Wiederholung einer spezifischen Verknüpfung von Reiz und Reaktion zurückgeht. Nach dieser einfachen Modellvorstellung besteht die Repräsentation der Aufgaben in einem gemischten Block aus vier voneinander unabhängigen Verknüpfungen zwischen Reizkategorien und Reaktionen. Ein solches einfaches Modell würde vorhersagen, dass Unterschiede zwischen Aufgabenwechseln und Aufgabewiederholungen auf die Wiederholung einer spezifischen Verknüpfung zurückgehen, die nur bei Aufgabewiederholungsdurchgängen auftreten kann. Die residuale Aktivierung dieser spezifischen Verknüpfung führt so zu einer erleichterten Wiederanwendung. Das Aufgabenwechselmodell geht hingegen davon aus, dass solche Effekte auch dann auftreten, wenn nur

solche Aufgabenwiederholungsdurchgänge betrachtet werden, bei denen keine Wiederholung der spezifischen Kategorie-Reaktions-Verknüpfung auftritt, da die Aufgaben in Form zusammengesetzter Aufgabensets repräsentiert sind.

Effekt eines irrelevanten Attributs

Das Aufgabenwechselmodell sagt Performanzunterschiede zwischen kompatibler und inkompatibler Tastenzuordnung sowohl bei Reizen mit irrelevantem Attribut (Zielkonzeptreizen) als auch bei Reizen ohne ein irrelevantes Attribut vorher. Wie weiter oben ausgeführt basiert diese Vorhersage auf der Annahme, dass Aufgabenwechselkosten in der inkompatiblen Bedingung eines IAT auch für den Wechsel von Zielkonzeptreizen zu Attributreizen auftreten sollten, was im Mittel zu einer verlangsamten Bearbeitung dieser Durchgänge führen sollte. Auch das Modell von Brendl et al. (2001), das den IAT-Effekt auf eine Verschiebung des Antwortkriteriums in der inkompatiblen Bedingung des IAT zurückführt, macht diese Vorhersage. Die Antwortlatenzen für Attributreize sollten in dieser Bedingung zwar länger ausfallen, der Kompatibilitätseffekt sollte jedoch überproportional die Zielkonzeptaufgabe betreffen. Diese Vorhersage beruht darauf, dass Unterschiede zwischen den Akkumulationsraten für verschiedene Reize bei einem Random-Walk mit konstanten Akkumulationsraten umso deutlicher zutage treten sollten, je konservativer die Antwortschwelle gesetzt ist. Anders gesagt, muss der Effekt einer Verschiebung des Antwortkriteriums antiproportional zur Akkumulationsrate ausfallen, wie Abbildung 2 verdeutlicht. Je niedriger die Netto-Akkumulationsrate, desto größer der Effekt einer Verschiebung des Antwortkriteriums. Das Aufgabenwechselmodell sagt im Gegensatz dazu vorher, dass der Kompatibilitätseffekt beide Aufgabenarten gleichermaßen betrifft. Es wird folglich keine Interaktion zwischen Kompatibilität und Aufgabenart erwartet.

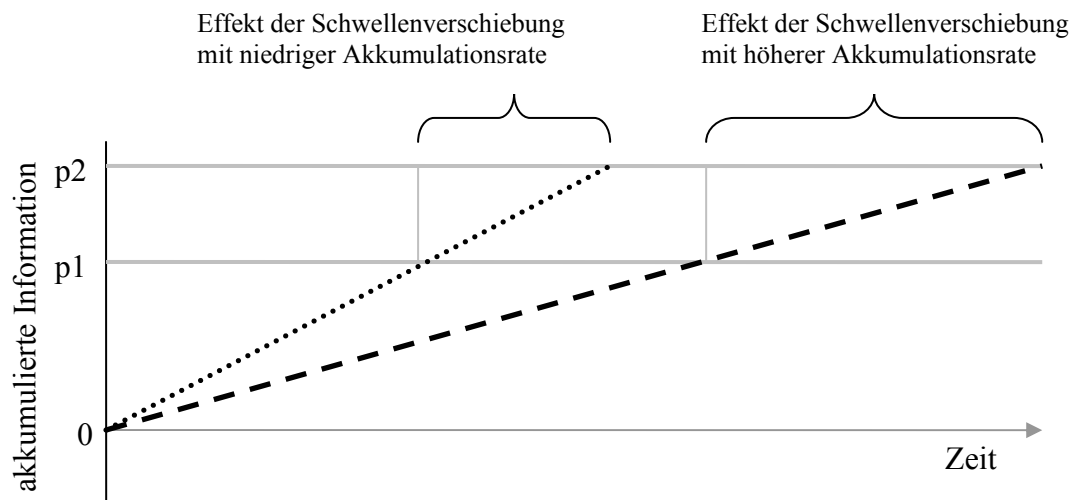


Abbildung 2: Effekt der Schwellenverschiebung von einer liberalen ($p1$) zu einer konservativen Position ($p2$) in Abhängigkeit von der Akkumulationsrate. Bei hoher Akkumulationsrate (gepunktete Linie) ergibt sich ein geringerer Effekt als bei niedriger (gestrichelte Linie). Unterschiede zwischen verschiedenen Akkumulationsraten sollten also umso deutlicher zutage treten, je konservativer die Antwortschwelle ausfällt.

Methode

Die Versuchsteilnehmer von Experiment 1 bearbeiteten einen einfachen Blumen-Insekten Einstellungs-IAT. Die Attributaufgabe ist eine evaluative Entscheidung bezüglich positiver und negativer Substantive. Die Zielkonzeptaufgabe erfordert die Unterscheidung von Blumen- und Insektenbezeichnungen. Die Verwendung des vergleichsweise künstlichen Blumen-Insekten Materials hat den Vorteil, dass die Valenz der Kategorien und ihrer Exemplare über die Stichprobe hinweg weitgehend einheitlich sein sollte. Dadurch lassen sich kompatible und inkompatible Reaktionszuordnung präexperimentell definieren, was die Prüfung der verschiedenen Interaktionshypothesen erleichtert. Die Kombination Blumen + positiv / Insekten + negativ wird als kompatible Zuordnung kodiert. Zur Prüfung der Annahmen des Trägheitsmodells von Allport et al. (1994) wurde das RSI in zwei Stufen manipuliert, 100 ms und 1000 ms. Ferner wurde die Tastenzuordnung so ausbalanciert, dass eine Hälfte der Versuchsteilnehmer auf positive und die andere Hälfte auf negative Wörter mit der dominanten Hand reagieren musste. Weiterhin wurde ausbalanciert, ob die kompatible oder die inkompatible Bedingung des IAT zuerst zu bearbeiten war.

Stichprobe

An der Untersuchung nahmen 32 Personen teil. Zehn Teilnehmer waren männlichen und 22 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 19 und 37 Jahren ($M = 24.20$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Das Material bestand aus 96 Wörtern, die Insekten, Blumen, positive Objekte und negative Objekte bezeichnen. Um Materialeffekte zu minimieren, wurden die Wörter in Quadrupeln zusammengefasst, die bezüglich dreier Kriterien maximal ähnlich waren, namentlich bezüglich der Zahl der Buchstaben, der Auftretenshäufigkeit der Wörter und ihrer Valenz.

Zunächst wurden 800 deutlich positiv und negativ besetzte Wörter auf Basis einer Wortnormierungsstudie deutscher Substantive (Schwibbe, Räder, Schwibbe, Borchardt & Geiken-Pophanken, 1994) ausgewählt. Die jeweils 200 am deutlichsten positiven und negativen Wörter wurden daraufhin von drei unabhängigen Urteiler begutachtet. Die Urteiler waren instruiert solche Wörter zu entfernen, die ihnen entweder gänzlich unbekannt waren oder ungebräuchlich erschienen. Um eine semantische Überlappung mit den Zielkonzepten des IAT zu vermeiden, wurden des Weiteren Wörter entfernt, die sich auf Pflanzen oder Tiere bezogen. 132 positive und 143 negative Wörter verblieben in der Stichprobe.

Nach den gleichen Kriterien wurde eine Liste mit 186 Exemplaren der Zielkonzepte Blumen und Insekten bewertet, die aus verschiedenen Quellen zusammengestellt wurde. Auf Basis dieser Urteile wurden 46 Blumen- und 62 Insektenbezeichnungen ausgewählt, die daraufhin in einer unabhängigen Pilotstudie zehn weiteren Urteilern vorgelegt wurden. Die Urteiler bewerteten jedes der vorgelegten 108 Wörter bezüglich der gleichen Frage mit dem gleichen Antwortformat, das in der Studie von Schwibbe et al. (1994) benutzt wurde. Eine Schätzung der Auftretenshäufigkeit der Wörter wurde aus der lexikalischen Datenbank Celex (Celex, 1995) gewonnen.

Zur Bildung von maximal ähnlichen Wortquadrupeln wurden die einzelnen Wörter als Punkte in einem 3-dimensionalen Raum projiziert, definiert durch die z-transformierten Valenzurteile, die Häufigkeitsschätzungen und die Wortlänge. Die Bildung der Quadrupel

basierte auf einem Algorithmus, der sequentiell je ein Wort aus jeder der vier Mengen selektiert, sodass die mittlere euklidische Distanz zwischen den Wörtern minimal ist. Die 24 resultierenden Quadrupel mit der geringsten mittleren Distanz zwischen den Wörtern wurden für das Hauptexperiment ausgewählt und sind in Anhang A zu finden.

Prozedur

Die Versuchsblöcke des Experiments umfassten die Präsentation von jeweils 48 einzelnen Wörtern. Die Wörter wurden ohne Zurücklegen aus den Reizlisten gezogen, sodass Wörter aus den Attribut- und Zielkonzeptlisten in zufälliger Reihenfolge erschienen. Dadurch ergaben sich in Blöcken mit gemischten Aufgaben Durchgänge, bei denen die zu bearbeitende Aufgabe im Vergleich zum Vorgängerdurchgang wechselt, und solche bei denen die zu bearbeitende Aufgabe gleich bleibt. Jede der vorkommenden Reizkategorien war gleich häufig zu bearbeiten. Die Wörter wurden in schwarzer Schrift auf hellgrauem Hintergrund in einem 20 mm x 120 mm großen Präsentationsfeld auf einem 17° Bildschirm dargeboten. Das Response-Stimulus Intervall (RSI) betrug - abhängig von der Versuchsbedingung - entweder 100 ms oder 1000 ms. Tastendrucke wurden registriert, sobald das jeweilige Wort sichtbar war. Insgesamt wurden 14 Blöcke präsentiert, sodass jeder Versuchsteilnehmer in Experiment 1 insgesamt 672 separate Durchgänge bearbeitete.

Die Teilnehmer wurden instruiert, evaluative Entscheidungen zu treffen, wenn der präsentierte Stimulus ein positives oder negatives Wort war, und eine Blumen-Insekten Diskrimination vorzunehmen, wenn das Wort eine Blumen- oder Insektenbezeichnung war. Sie wurden angehalten so schnell wie möglich auf die Wörter zu reagieren, gleichzeitig jedoch Antwortfehler zu vermeiden. Um einen Versuchsblock zu starten, mussten zunächst beide Antworttasten gedrückt werden, woraufhin ein kurzer Countdown den Beginn des ersten Durchgangs einleitete.

Das Experiment begann mit zwei Übungsphasen, die aus jeweils zwei Aufgabenblöcken bestanden. In der ersten Übungsphase wurden ausschließlich Blumen- und Insektenbezeichnungen dargeboten und die Versuchsteilnehmer wurden angehalten, die zugehörige Blumen-Insekten-Diskriminationsaufgabe einzuüben. In der zweiten Phase wurde die evaluative Entscheidungsaufgabe eingeübt. In dieser Phase wurden nur positive und negative Substantive präsentiert. Die vier Übungsblöcke wurden von vier gemischten Blöcken gefolgt, in denen beide Aufgaben in zufälliger Reihenfolge zu bearbeiten waren. Die Tastenzuordnung der Attribut- und der Zielkonzeptaufgabe war entweder kompatibel oder inkompatibel, in

Abhängigkeit von der Balancierungsbedingung. Die verbleibenden sechs Blöcke bestanden aus zwei weiteren Übungsblöcken, in denen die Tastenzuordnung für die Zielkonzeptaufgabe umgekehrt wurde, und vier gemischten Blöcken, die mit dieser umgekehrten Tastenzuordnung bearbeitet wurden. Die Kompatibilität der Tastenzuordnung in der letzten Phase des Experiments war somit im Verhältnis zur ersten gemischten Phase umgekehrt. Zu Beginn jeder neuen Phase wurden die Versuchsteilnehmer über die im nächsten Block vorkommenden Reizkategorien und ihre Zuordnung zu den Antworttasten informiert. Die Bearbeitung des Experiments dauerte etwa 25 Minuten.

Ergebnisse

Um einem Effekt antizipatorischer Reaktionen und kurzfristiger Aufmerksamkeitslücken entgegenzuwirken, wurden Antwortlatenzen unter 300 ms zu 300 ms (0.5 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.5 % der Durchgänge) rekodiert (vgl. Greenwald et al., 1998). Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (8.2 % der Durchgänge) nicht in die Analyse aufgenommen.

Zur Prüfung der verschiedenen Hypothesen zu Sequenzeffekten wurden die einzelnen Durchgänge mit Bezug auf Wiederholung oder Wechsel von Merkmalen des direkt vorhergehenden Durchgangs kodiert (vgl. Pashler & Baylis, 1991b). Ein erster Sequenzfaktor betrifft die Wiederholung spezifischer Reaktionen, das heißt, ob der direkte Vorgängerdurchgang die gleiche oder eine andere Reaktion erforderte. Der zweite Sequenzfaktor betrifft das Verhältnis der aktuell und im Vorgängerdurchgang zu bearbeitenden Aufgabe. Wenn im aktuellen und im Vorgängerdurchgang die gleiche Aufgabe zu bearbeiten war, so wird der aktuelle Durchgang als Aufgabenwiederholung kodiert. Erfordern der aktuelle und der direkt vorhergehende Durchgang die Bearbeitung verschiedener Aufgaben - die Bearbeitung der Attributaufgabe im Vorgängerdurchgang und die Bearbeitung der Zielkonzeptaufgabe im aktuellen, oder umgekehrt - so wird der aktuelle Durchgang als Aufgabenwechsel kodiert. Als Resultat dieser Kodierung ergeben sich zwei orthogonale within-subjects Faktoren, die unabhängig von allen weiteren im IAT manipulierten Variablen variieren.

Die mittleren Antwortlatenzen wurden mit einer 2 (Reihenfolge der kompatiblen und inkompatiblen Phasen) x 2 (RSI = 100 ms vs. RSI = 1000 ms) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) x 2 (Wechsel vs. Wiederholung der zutreffenden Antwort

zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse (ANOVA) ausgewertet¹. Die drei letztgenannten Faktoren wurden innerhalb der Teilnehmer manipuliert. Die mittleren Latenzen können Tabelle 1 entnommen werden.

Tabelle 1: Mittlere Latenzen in Millisekunden als Funktion der Kompatibilität, des RSI und der Sequenzfaktoren in Experiment 1.

	Aufgabenwechsel		Aufgabenwiederholung	
	Antwort- wechsel	Antwort- wiederholung	Antwort- wechsel	Antwort- wiederholung
RSI 100 ms				
Kompatibel	798	772	775	713
Inkompatibel	1030	1037	919	801
RSI 1000 ms				
Kompatibel	675	675	690	633
Inkompatibel	819	798	758	715

Wie erwartet sind die mittleren Antwortlatenzen in der kompatiblen Bedingung des IAT bedeutsam kürzer als in der inkompatiblen $F(1, 28) = 49.00, p < .01$. Dies weist auf eine relativ positivere Bewertung der Blumen gegenüber den Insekten hin. Der entsprechende IAT-Effekt beträgt 143 ms. Die Reaktionen erfolgen bei Aufgabenwiederholung signifikant schneller als bei Aufgabenwechseln, wie ein signifikanter Haupteffekt des Aufgabenwechselfaktors zeigt, $F(1, 28) = 48.62, p < .01$. Wie vermutet sind diese Aufgabenwechselkosten in der inkompatiblen Bedingung wesentlich ausgeprägter als in der kompatiblen, $F(1, 28) = 28.51, p < .01$. Dieses Ergebnis wird in Abbildung 3 veranschaulicht. Geplante Vergleiche zeigen jedoch, dass Aufgabenwechselkosten sowohl in der inkompatiblen Bedingung, $F(1, 28) = 57.64, p < .01$, als auch in der kompatiblen Bedingung auftreten, $F(1, 28) = 6.08, p < .05$.

¹ Um die Darstellung überschaubar zu halten, wird auf eine Dokumentation der Ergebnisse mit Antwortfehlern verzichtet. Dies entspricht der gängigen Praxis, und es ergeben sich keine abweichenden Schlussfolgerungen aus den entsprechenden Analysen.

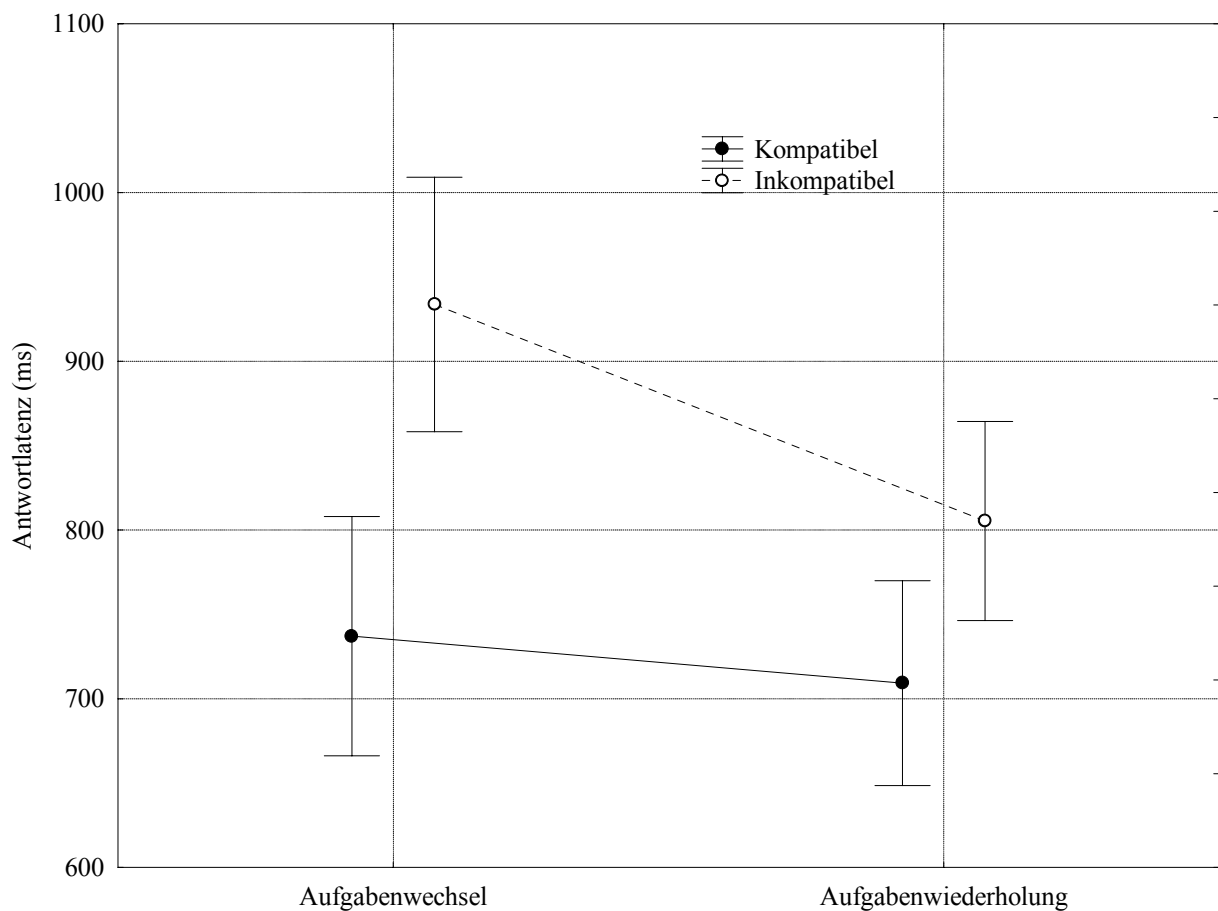


Abbildung 3: Mittlere Antwortlatenzen in Experiment 1 als Funktion der Kompatibilität der Tastenzuordnung für Aufgabenwechsel- und Aufgabenwiederholungsdurchgänge. Die Differenz zwischen kompatibler und inkompatibler Bedingung fällt für Aufgabenwechseldurchgänge signifikant größer aus, innerhalb der kompatiblen Bedingung finden sich nahezu keine Aufgabenwechselkosten.

Insgesamt sind die Latenzen für die Bedingung mit einem RSI von 100 ms kürzer als bei einem RSI von 1000 ms, $F(1, 28) = 5.80, p < .01$. Die Aufgabenwechselkosten fallen bei langem RSI bedeutsam geringer aus, als bei kurzem, $F(1, 28) = 10.05, p < .01$. Eine Zweifach-Interaktion zwischen RSI und der Kompatibilität der Phase verfehlt knapp das Signifikanzniveau, $F(1, 28) = 3.49, p = .056$, ebenso eine Dreifach-Interaktion zwischen RSI, Kompatibilität und dem Aufgabenwechselfaktor, $F(1, 28) = 4.06, p = .054$. Ferner deutet eine marginal signifikante Interaktion zwischen dem IAT-Effekt und der Balancierung der Kompatibilitätsreihenfolge an, dass die Latenzen in den kompatiblen Phasen länger ausfallen, wenn vorher die inkompatible Phase bearbeitet wurde, während die Latenzen in der inkompatiblen Phase unabhängig von ihrer seriellen Position im Experiment ausfallen, $F(1, 28) = 4.03, p < .054$.

Erfordert ein Durchgang das erneute Reagieren mit der gleichen Antworttaste, so fallen die Latenzen geringer aus, $F(1, 28) = 28.20, p < .01$. Diese Erleichterung tritt jedoch nur auf, wenn auch die gleiche Aufgabe bearbeitet werden muss, wie eine Zweifach-Interaktion zwischen dem Aufgabenwechselfaktor und der Antwortwiederholung zeigt, $F(1, 28) = 16.58, p < .01$. Die theoretisch relevante Frage, ob Aufgabenwechselkosten auch dann auftreten, wenn die Antwort des vorherigen Durchgangs nicht wiederholt wird, kann ein entsprechender Kontrast zwischen Aufgabenwechsel- und -wiederholungsdurchgängen klären, in den nur Durchgänge mit ungleicher Antwort eingehen. Auch unter diesen Bedingungen ergeben sich erwartungsgemäß signifikante Aufgabenwechselkosten, $F(1, 28) = 16.35, p < .01$. Es kann also gefolgert werden, dass Aufgabenwechsel nicht auf einen trivialen Wiederholungseffekt der gleichen S-R Zuordnung zurückgehen.

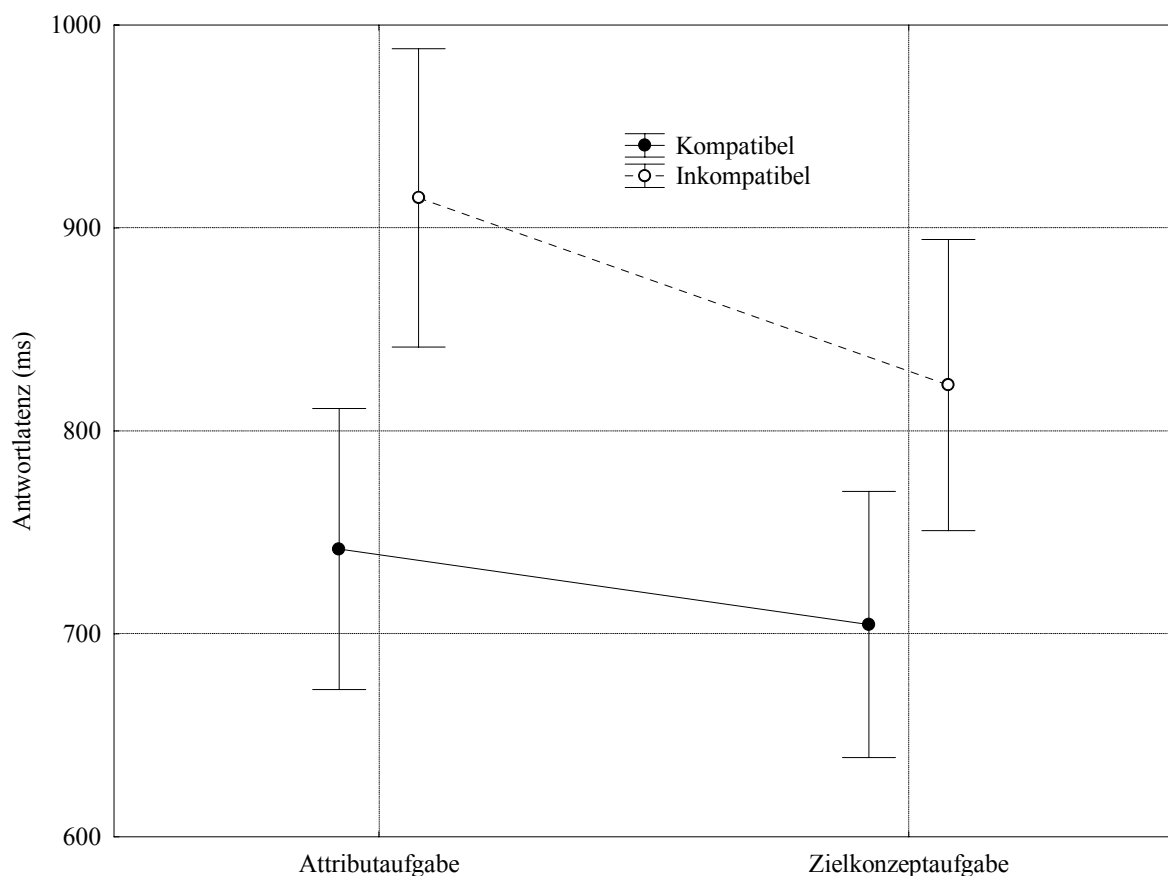


Abbildung 4: Mittlere Antwortlatenzen in Experiment 1 als Funktion der Kompatibilität der Tastenzuordnung für die Attribut- und die Zielkonzeptaufgabe. Entgegen der Vorhersage des Random-Walk Modells fallen die IAT-Effekte für die Attributaufgabe signifikant größer aus als für die Zielkonzeptaufgabe.

Wie weiter oben ausgeführt, lässt das Random-Walk Modell von Brendl et al. (2001) erwarten, dass die Unterschiede zwischen kompatibler und inkompatibler Bedingung für die Zielkonzeptaufgabe größer ausfallen als für die Attributaufgabe. Um diese Vorhersage zu prüfen, wurden die Antwortlatenzen ferner mit einer 2-faktoriellen Varianzanalyse mit den Faktoren Tastenzuordnung (kompatibel versus inkompatibel) und Aufgabentyp (Attribut-versus Zielkonzeptaufgabe) analysiert. Dabei ergibt sich - entgegen der Vorhersage des Random-Walk Modells -, dass der IAT-Effekt für die Attributaufgabe stärker ausfällt als für die Zielkonzeptaufgabe, $F(1, 28) = 11.51, p < .01$, obwohl lediglich für die Zielkonzeptaufgabe Interferenzen im Sinne des Random-Walk Modells zu erwarten sind. Die dieser Interaktion zugrunde liegenden Mittelwerte sind in Abbildung 4 dargestellt.

Diskussion

Die Ergebnisse von Experiment 1 bestätigen den an Allport et al. (1994) angelehnten Trägheitsmechanismus. Einerseits sind die Kompatibilitätseffekte, wie erwartet, bei Aufgabenwechselfurchgängen signifikant größer als bei Aufgabenwiederholung, andererseits verschwinden die Aufgabenwechselkosten in der kompatiblen Phase nicht vollständig. Insbesondere die Tatsache, dass die Aufgabenwechselkosten bei einem kürzeren RSI geringer ausfallen, entspricht den Vorhersagen des Trägheitsmechanismus im Sinne eines passiven Abklingens proaktiver Interferenz innerhalb des RSI, wie ihn das Modell von Allport et al. annimmt: Je länger der vorhergehende Durchgang zeitlich zurückliegt, desto geringer sollte dessen störende Nachwirkung ausfallen.

Die Tatsache, dass Aufgabenwechselkosten auch dann auftreten, wenn Durchgänge mit einer Wiederholung der spezifischen S-R Verknüpfung des vorherigen Durchgangs ausgeschlossen werden, zeigt, dass diese keinen trivialen Wiederholungseffekt widerspiegeln. Dass überhaupt Effekte der Antwortwiederholung entstehen, steht im Einklang mit den Annahmen des Trägheitsmechanismus, da Auswirkungen der Vorläuferdurchgänge explizit erwartet werden. Der entscheidende Befund ist jedoch, dass die Auswirkungen des Vorgängerdurchgangs nicht durch einen trivialen Wiederholungseffekt der motorischen Reaktion oder der Wiederverwendung der gleichen S-R Regel erschöpfend erklärt werden können, sondern weitere aufgabenrelevante Aspekte umfassen müssen.

Im Einklang mit früheren Befunden zeigt Experiment 1 einen marginal signifikanten Effekt der Kompatibilitätsreihenfolge. Wie in früheren Untersuchungen (Greenwald et al., 1998) geht dieser Effekt vor allem auf Performanzunterschiede in der kompatiblen Phase des

IAT zurück, während die Leistung in der inkompatiblen Phase unabhängig von ihrer seriellen Position im Experiment ist. Dies deutet auf die Präsenz negativen Transfers auf die Bearbeitung der kompatiblen Phase hin (Nosek et al., 2002).

Ferner zeigt sich, entgegen den Vorhersagen des Modells von Brendl et al. (2001), dass der IAT-Effekt für die Attributaufgabe größer ausfällt als für die Zielkonzeptaufgabe. Der entscheidende Punkt dieser Analyse ist nicht, dass überhaupt ein IAT-Effekt für die Attributaufgabe auftritt, sondern dass der Kompatibilitätseffekt für die Attributaufgabe größer ausfällt. Zwar kann man erwarten, dass bei einer anfänglich größeren Schwierigkeit der Attributaufgabe eine Verschiebung des Antwortkriteriums auch zu einem größeren Performanzverlust führt. Es ist jedoch kaum plausibel, dass dieser Effekt stärker ausfällt als in der Zielkonzeptaufgabe, für die sowohl eine Verschiebung des Antwortkriteriums als auch eine verminderte Akkumulationsrate die Antwortlatenzen verringern. Auf Basis des Aufgabenwechselmodells kann dieser Effekt so erklärt werden, dass die Lösung einer im Vorgängerdurchgang nötigen Inhibition von Bewertungsprozessen vergleichsweise länger dauert als der betreffende Inhibitionsprozess selbst. Zwar scheint störende Attributinformation vergleichsweise schnell und effizient unterdrückt zu werden, ist sie jedoch in einem nachfolgenden Durchgang relevant, dann ist ihre Verarbeitung mit deutlichen Kosten verbunden.

4.2 Experiment 2

Das Aufgabenwechselmodell nimmt an, dass eine Verlangsamung der Performanz bei der Bearbeitung zweier wechselnder Aufgaben immer dann auftritt, wenn zwei oder mehrere Zuordnungen von Reiz und Reaktion konfligierende Antworten für eine Untermenge der Reize fordern. Im Gegensatz dazu sollte keine Verlangsamung eintreten, wenn die Reaktionszuordnungen in Bezug auf die abzugebenden Antworten für alle Reize übereinstimmen. Interessanterweise ist Kompatibilität in IAT-ähnlichen Aufgaben also keine Eigenschaft einzelner Durchgänge, wie etwa in der Stroop-Aufgabe (MacLeod, 1991) oder beim affektiven Priming (Fazio et al., 1986), sondern ein *Kontexteffekt*, der auf einer Interaktion der Aufgabenspezifikation und der verwendeten Reize basiert. Genauer gesagt geht die Kompatibilität darauf zurück, dass die Attributaufgabe - mehr oder minder offensichtlich - auch auf die Exemplare der Zielkonzeptreize angewandt werden kann. Für sich genommen sind die Reize in der kompatiblen und inkompatiblen Bedingung eines IAT weder kompatibel noch inkompatibel. Die Überlappung der in der Attributaufgabe spezifizierten Kategorien (z.B. positive und negative Reize) mit den Exemplaren der Zielkonzeptkategorien (z.B.

Blumen und Insekten) ist also der entscheidende Faktor. Im Gegensatz dazu ist die Zielkonzeptaufgabe meist nicht auf die Exemplare der Attributreize anwendbar: Die Blumen sind zwar positiv, die positiven Wörter jedoch nicht blumig.

In Experiment 2 soll geprüft werden, ob IAT-Effekte auch dann auftreten, wenn keine assoziativen Verknüpfungen oder Figur-Grund Asymmetrien vorliegen. Als Material dienten Abbildungen einfacher geometrischer Objekte verschiedener Farbe und Größe. Die zu unterscheidenden Zielkonzepte waren rote und blaue Objekte. Als Attributaufgabe sollten Entscheidungen über die Größe andersfarbiger Objekte getroffen werden. Eine Beziehung zwischen Größe und Farbe der Objekte wurde experimentell durch eine Kovariation der Reizeigenschaften erzeugt. In einer Versuchsbedingung waren alle roten Objekte groß und alle blauen Objekte klein, in einer anderen Versuchsbedingung war die Kovariation umgekehrt. Die Größe der andersfarbigen Objekte variierte zufällig. Bei diesen Objekten gab es keine Kovariation von Größe und Farbe.

Eine Zuordnung der Objekte zu Reaktionen ist dann kompatibel, wenn alle roten Objekte klein und der gleichen Taste zugeordnet sind wie die restlichen kleinen Objekte. Eine inkompatible Zuordnung ist entsprechend dann gegeben, wenn alle roten Objekte groß, jedoch der gleichen Taste wie die andersfarbigen kleinen Reize zugeordnet sind. Trotz der offensichtlichen Unterschiede ist diese Anordnung in verschiedener Hinsicht analog zu der in einem herkömmlichen IAT. So gehören die Reize der Attributkonzepte zu keiner der beiden Zielkonzeptkategorien, das heißt, sie waren weder blau noch rot (analog den positiven und negativen Reizen in einem Blumen-Insekten-IAT, die weder Blumen noch Insekten sind). Ferner waren die Zielkonzeptkategorien diskret, wie die Kategorien Blume und Insekt, während das Attribut "Größe" eine kontinuierliche Ausprägung hat, ähnlich der Valenz in einem Einstellungs-IAT. Da präexperimentell nicht von einer Assoziation von Farbe und Größe geometrischer Objekte auszugehen ist, müssen Performanzunterschiede zwischen diesen Bedingungen auf die Kovariation der Reizeigenschaften zurückgehen.

Wenn unter den gegebenen Bedingungen IAT-Effekte auftreten, so spricht dies gegen die starke Hypothese, dass assoziative Verknüpfungen eine notwendige Bedingung für das Auftreten von IAT-Effekten sind. Gleichzeitig wäre jedoch gezeigt, dass IAT-Effekte auch dann auftreten können, wenn die Konzepte selbst unverknüpft sind und lediglich Relationen auf der Ebene der Exemplare vorliegen. Dies steht im Widerspruch zu der Annahme einer zentralen Rolle von Assoziationen auf Ebene der Konzepte und der untergeordneten Rolle von Eigenschaften der Exemplare.

Das Figur-Grund Modell von Rothermund und Wentura (2001) lässt für das geometrische Material andere Ergebnisse erwarten als das Aufgabenwechselmodell. Wie weiter oben erläutert basiert das Figur-Grund Modell auf der Annahme, dass die Aufgabenstruktur bei einer kompatiblen Tastenzuordnung auf Basis der Salienz der dargebotenen Reize rekodiert wird. Zur Erklärung eines IAT-Effekts mit dem beschriebenen geometrischen Material müsste man folglich annehmen, dass zum Beispiel rote und kleine Wörter jeweils als saliente Figur auf dem Grund der entgegengesetzten Kategorie verarbeitet werden. Dies scheint eher unwahrscheinlich. Ein noch stärkeres Argument gegen das Figur-Grund Modell ergibt sich, wenn die Richtung der IAT-Effekte von der experimentell erzeugten Kovariation abhängt: Da die Salienz der Reize und Reizkategorien durch eine Änderung der Kovariation nicht beeinflusst wird, sollte die Richtung der IAT-Effekte von der Kovariation unabhängig sein. Das Aufgabenwechselmodell lässt hingegen eine vollständige Abhängigkeit der Richtung der IAT-Effekte von der Kovariation erwarten.

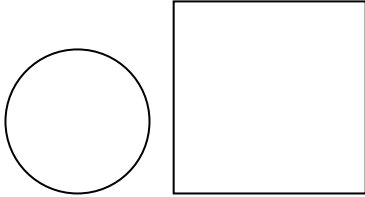

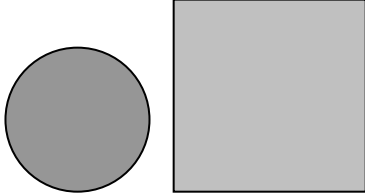

Methode

Die Versuchsteilnehmer bearbeiteten einen einfachen IAT mit geometrischen Objekten. Die Teilnehmer wurden instruiert, große und kleine geometrische Objekte durch Tastendruck zu unterscheiden. Wenn ein Objekt rot oder blau war, sollte dagegen die Größe des Objekts ignoriert werden und eine Unterscheidung anhand der Farbe vorgenommen werden. Für einen Teil der Versuchsteilnehmer waren alle roten Objekte klein und alle blauen Objekte groß. Für einen anderen Teil der Teilnehmer wurde diese Kovariation umgekehrt. Ferner wurde die Tastenzuordnung so ausbalanciert, dass eine Hälfte der Versuchsteilnehmer auf große und die andere Hälfte auf kleine Objekte mit der dominanten Hand reagieren musste. Weiterhin wurde ausbalanciert, ob die kompatible oder die inkompatible Bedingung des IAT zuerst zu bearbeiten war.

Stichprobe

An der Untersuchung nahmen 24 Personen teil. Acht Teilnehmer waren männlichen und 16 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 19 und 33 Jahren ($M = 23.38$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 5,- (ca. € 2,50) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Tabelle 2: Material des IAT mit geometrischen Objekten. Große und kleine Objekte sind pink, gelb oder grün gefärbt.

Blumen-Insekten Einstellungs-IAT	IAT mit geometrischen Objekten
Blumen	<p>Blaue Objekte</p> 
Insekten	<p>Rote Objekte</p> 
Positive Wörter	<p>Große Objekte</p> 
Negative Wörter	<p>Kleine Objekte</p> 

Material

Die verwendeten Reize waren einfache geometrische Objekte verschiedener Farbe (rot, blau, gelb, grün und pink), Größe (groß versus klein) und Form (Rechtecke, Kreise und Dreiecke). Gelb, grün und pink gefärbte Objekte dienten als Exemplare der Attributkategorien, die roten und blauen Objekte dienten als Exemplare der Zielkonzepte. Die Farbe der Reize war deutlich unterscheidbar. Die Kantenlänge bzw. der Durchmesser der großen Objekte lag bei 4 bis 5 cm, und bei 1 bis 2 cm für die kleinen Objekte. Die Form der Objekte war aufgabenirrelevant und wurde zufällig manipuliert, um Varianz in der Reizmenge zu produzieren. Eine Relation zwischen der Größe und der Farbe der Reize wurde durch eine Kovariation dieser Eigenschaften hergestellt: In einer Versuchsbedingung waren alle roten Objekte klein und alle blauen Objekte groß, in einer zweiten Bedingung waren alle roten Objekte groß und alle blauen Objekte klein. Die resultierende Beziehung zwischen diesen Eigenschaften beruht ausschließlich auf den verwendeten Exemplaren der Kategorien, nicht

jedoch auf einer Beziehung der Eigenschaften auf konzeptueller Ebene. Beispiele für die verwendeten Reize sind in Tabelle 2 dargestellt.

Prozedur

Der Ablauf und die Darbietungsparameter der einzelnen Durchgänge entsprechen denen von Experiment 1. Jede Versuchsperson bearbeitete insgesamt 6 Übungsblöcke, in denen Attribut- und Zielkonzeptaufgabe getrennt eingeübt wurden und 8 Experimentalblöcke in denen Attribut- und Zielkonzeptaufgabe gemischt zu bearbeiten waren. Das RSI betrug für alle Versuchsteilnehmer 800 ms. Die Bearbeitung des Experiments dauerte etwa 25 Minuten.

Ergebnisse

Um einem Effekt antizipatorischer Reaktionen und kurzfristiger Aufmerksamkeitslücken entgegenzuwirken, wurden Antwortlatenzen unter 100 ms zu 100 ms (0.7 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.2 % der Durchgänge) rekodiert. Diese niedrigere Schwelle von 100 ms ist angemessen, da die Klassifikation der geometrischen Objekte offenbar äußerst schnell möglich war, sodass ca. 8% der Durchgänge zwischen 100 ms und 300 ms liegen. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (7.9 % der Durchgänge) nicht in die Analyse aufgenommen. Die Kompatibilität der Bedingung wurde in Abhängigkeit von Kovariation der Eigenschaften kodiert. Eine Bedingung wurde dann als kompatibel kodiert, wenn rote und blaue Objekte jeweils der Taste zugeordnet sind, die auch bei einer Entscheidung aufgrund des für diese Reize irrelevanten Merkmals Größe gedrückt werden müsste.

Der mittlere IAT-Effekt liegt bei 308 ms. Die Daten wurden mit einer 2 (kompatible vs. inkompatible Bedingung) x 2 (Aufgabenwechsel vs. Aufgabenwiederholung) x 2 (Richtung der Kovariation von Farbe und Größe der Objekte) x 2 (kompatible vs. inkompatible Bedingung zuerst) x 2 (Balancierung der Tastenzuordnung) faktoriellen Varianzanalyse mit Messwiederholungen auf den ersten beiden Faktoren analysiert. Wie erwartet zeigt sich ein bedeutsamer IAT-Effekt in Richtung der Kovariation der Eigenschaften, $F(1, 16) = 52.97, p < .01$, der darauf zurückgeht, dass die Bedingungen, in denen alle Reize einer bestimmten Größe der gleichen Taste zugeordnet sind, schneller bearbeitet wurden. Wie in Experiment 1 wurden Aufgabenwechseldurchgänge insgesamt langsamer bearbeitet, $F(1, 16) = 29.78, p < .01$. Eine bedeutsame Zweifach-Interaktion zwischen Kompatibilität und dem

Aufgabenwechselfaktor, $F(1, 16) = 34.59, p < .01$, zeigt ferner, dass die IAT-Effekte bei Aufgabenwechselfaktoren signifikant größer ausfallen. Alle weiteren Haupteffekte und Interaktionen bleiben über dem Signifikanzniveau von 5%.

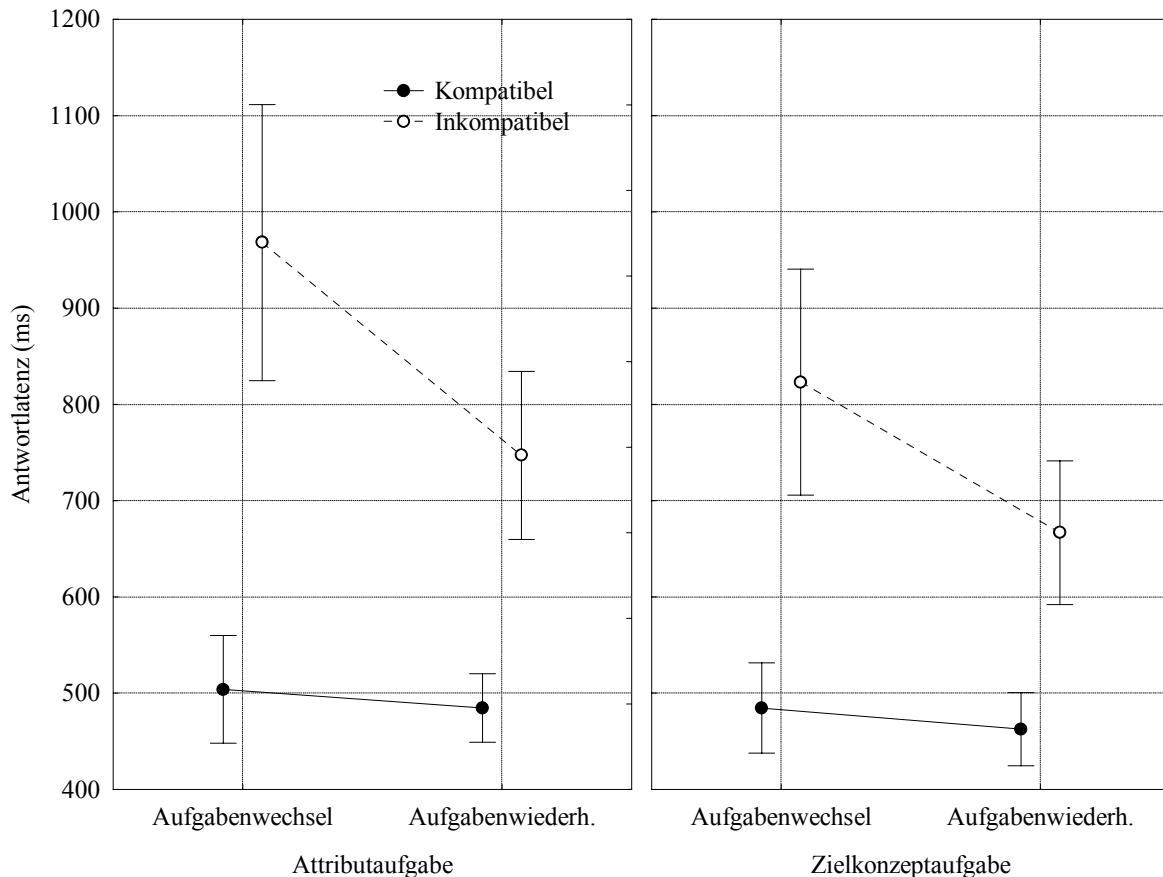


Abbildung 5: Mittlere Antwortlatenzen in Experiment 2 als Funktion der Kompatibilität für Attribut- und Zielkonzeptaufgabe getrennt nach Aufgabenwechsel- versus Aufgabenwiederholungsdurchgängen. Für die kompatible Phase des IAT finden sich nahezu keine Aufgabenwechselkosten. Die Asymmetrie der Aufgabenwechselkosten ist für die Attributaufgabe größer als für die Zielkonzeptaufgabe.

Wie weiter oben beschrieben, sagt das Figur-Grund Modells von Rothermund und Wentura (2001) vorher, dass die IAT-Effekte unabhängig von der Kovariation der Eigenschaften sein sollten, da die Saliens durch eine Änderung der Kovariation nicht betroffen ist. Um diese Hypothese kritisch zu prüfen, wurden die IAT-Effekte für eine zweite Analyse unabhängig von der Kovariation kodiert. Die Kombination blau + groß und rot + klein wurde als eine Bedingung zusammengefasst und die Kombination blau + klein und rot + groß als zweite Bedingung. Wenn die Annahmen des Figur-Grund Modells zutreffen, so sollten die

IAT-Effekte unabhängig von der Kovariation sein. Eine 2 (Kovariation der Eigenschaften) x 2 (Aufgabenwechsel versus -wiederholung) faktorielle ANOVA über die rekodierten IAT-Effekte zeigt jedoch, dass diese in Abhängigkeit von der Kontingenz die Richtung ändern, $F(1, 22) = 58.36, p < .01$. Auch die Wechselkostenasymmetrie ändert, wie erwartet, in Abhängigkeit von der Kovariation die Richtung, $F(1, 22) = 40.20, p < .01$.

In einer weiteren Analyse sollte geprüft werden, ob der in Experiment 1 aufgetretene größere IAT-Effekt für die Attributaufgabe mit dem andersartigen Material von Experiment 2 replizierbar ist. Dazu wurde eine 2 (Kompatibilität) x 2 (Zielkonzept- vs. Attributaufgabe) x 2 (Aufgabenwechsel vs. -wiederholung) faktorielle ANOVA über die mittleren Antwortlatenzen berechnet. Wie schon in Experiment 1 zeigt sich dabei ein größerer Effekt der Kompatibilität auf die Attributaufgabe, $F(1, 23) = 21.89, p < .01$. In dieser Analyse wird sichtbar, dass dieser Effekt erwartungsgemäß vor allem bei Aufgabenwechseldurchgängen auftritt, $F(1, 22) = 7.65, p < .05$. Abbildung 5 zeigt die der Dreifach-Interaktion zugrunde liegenden mittleren Antwortlatenzen.

Diskussion

Experiment 2 repliziert die charakteristische Wechselkostenasymmetrie, die schon in Experiment 1 auftrat, mit einem strukturell vollkommen unterschiedlichen Material. Dies bestätigt die Annahme, dass Aufgabenwechselkosten ein entscheidender und möglicherweise materialunabhängiger Faktor des Designs eines IAT sind. Die Umkehrung der IAT-Effekte durch Umkehrung der Kovariation von Reizeigenschaften kann, wie weiter oben erläutert, nicht durch das Figur-Grund Modell von Rothermund und Wentura (2001) erklärt werden, denn die Salienz der verwendeten Reize oder Reizkategorien bleibt durch die Änderung der Kovariation unverändert. Auch können die IAT-Effekte in den jeweiligen Bedingungen nicht auf Figur-Grund Asymmetrien zurückgehen, wenn der Figur-Grund-Mechanismus die Umkehrung der Effekte nicht vorhersagen kann. Hier liegt ein entscheidender Vorteil des im vorliegenden Experiment verwendeten artifiziellen Materials: Im Gegensatz zu herkömmlichem Material können Zusammenhänge zwischen Reizen bzw. Reizkategorien und ihre semantischen und inhaltlichen Qualitäten gezielt dekonfundiert werden.

Bedeutend ist auch, dass die dargestellten IAT-Effekte ausschließlich auf eine Kovariation von Eigenschaften der verwendeten Reize zurückgehen, nicht aber auf eine Relation auf Ebene der Konzepte. Dies widerspricht den Schlussfolgerungen von Steffens und Kollegen (2004; Steffens & Plewe, 2001) und De Houwer (2001b), die davon ausgehen, dass

Zusammenhänge auf Ebene der verwendeten Exemplare lediglich eine untergeordnete Rolle spielen. Die hier vorgestellten Ergebnisse sprechen eher dafür, dass Effekte der übergeordneten Konzepte und ihrer Exemplare nicht für jeden IAT gleich ausfallen, sondern eine Funktion ihrer spezifischen Merkmale sind. Es ist zwar plausibel, dass die Rolle der spezifischen Exemplare gering ausfällt, wenn sie im Kontext eines starken übergeordneten Konzepts verarbeitet werden, gleichzeitig kann jedoch nicht ausgeschlossen werden, dass die Verarbeitung eines gegebenen übergeordneten Konzepts durch die Zusammensetzung der verwendeten Exemplare, im Sinne eines Kontexteffekts, geprägt wird. In diesem Licht stellen die Exemplareffekte im IAT mehr oder weniger allgemeine Materialeffekte dar, weniger jedoch spezifische Eigenschaften der verwendeten Methode, namentlich des IAT.

Wie schon in Experiment 1 finden sich im Vergleich zur Zielkonzeptaufgabe größere IAT-Effekte bei der Attributaufgabe. Eine entsprechende Analyse zeigt, dass diese Effekte vor allem bei Durchgängen mit Aufgabenwechsel auftreten. Dies bestätigt die Annahme, dass die größeren Effekte bei der Attributaufgabe vor allem auf die Nachwirkung von Inhibition zurückgehen, wie sie nach einem Aufgabenwechsel anfallen sollte. Die Lösung dieser Inhibition scheint mit größeren Kosten verbunden zu sein als der Inhibitionsprozess selbst. Einer vergleichsweise effizienten Hemmung irrelevanter beziehungsweise störender Information steht also eine Schwierigkeit entgegen, diese Information in einem nachfolgenden Durchgang erneut zu beachten und zu verarbeiten.

Die Befunde belegen ferner, dass IAT-Effekte ohne jegliche präexperimentelle Assoziation zwischen den Konzepten auftreten können. Dies Ergebnis ist nicht nur von hoher theoretischer Relevanz, es bestätigt auch, dass ein Schluss vom Auftreten eines IAT-Effekts auf das Vorliegen einer Assoziation im Gedächtnis nicht zulässig ist (z.B. Brendl et al., 2001; Mierke & Klauer, 2003). Sobald andere Beziehungen zwischen den Aufgaben und den verwendeten Reizen nicht ausgeschlossen werden können, können diese Zusammenhänge oder Konfundierungen eine Quelle für das Auftreten von IAT-Effekten sein.

4.3 Experiment 3

Kompatibilität ergibt sich in einem herkömmlichen IAT, wie weiter oben erläutert, aus der Beziehung zwischen der Zielkonzeptaufgabe und der Attributaufgabe. Üblicherweise wird die Kompatibilität dabei auf der Ebene vollständiger Aufgabenblöcke manipuliert. Eine Manipulation der Kompatibilität ist jedoch prinzipiell auch auf der Ebene einzelner Reize möglich. Bei der extrinsisch affektiven Simonaufgabe (De Houwer, 2003) wird zum Beispiel

ein zusätzliches aufgabenrelevantes Merkmal für die Zielkonzeptreize eingeführt, auf das reagiert werden soll. Dadurch kann für einen gegebenen Zielkonzeptreiz die geforderte Antwort derjenigen, die bei einer Anwendung der Attributaufgabe zu geben wäre, entsprechen oder widersprechen, wodurch sich eine extrinsische Kompatibilität der einzelnen Reize ergibt. Eine weitere Möglichkeit bestünde darin, die Tastenzuordnung für die Zielkonzeptreize innerhalb eines Aufgabenblocks zu variieren. Dadurch ergäbe sich jedoch eine äußerst komplexe und möglicherweise verwirrende Aufgabenstruktur. Ein solches Design erscheint somit kaum praktikabel.

Die Frage, welche Effekte die blockweise Manipulation der Kompatibilität auf den resultierenden IAT-Effekt hat, ist von theoretischer Bedeutung für das Aufgabenwechselmodell. Auf Basis des *Trägheitsmechanismus* wäre zu erwarten, dass vor allem solche Durchgänge verlangsamt sind, bei denen die vorhergehende Aufgabe mit der Bearbeitung der aktuellen Aufgabe interferiert. Insbesondere sollte die Bearbeitung eines inkompatiblen Reizes nach Bearbeitung der Attributaufgabe erschwert sein, da die im Vorgängerdurchgang aktivierte Zuordnungsregel in diesem Fall interferiert. Ebenso sollte die Bearbeitung eines Attributreizes nach Bearbeitung eines inkompatiblen Zielkonzeptreizes stärker erschwert sein als nach Bearbeitung eines kompatiblen, wenn die Bearbeitung der inkompatiblen Zielkonzeptreize eine Inhibition der Attributaufgabe erfordert. Unabhängig davon ist zu erwarten, dass die Bearbeitung der gleichen Aufgabe in zwei aufeinander folgenden Durchgängen leichter fällt als ein Wechsel der Aufgabe. Auf Basis des *Vereinfachungsmechanismus* wäre im Gegensatz dazu zu erwarten, dass bei einer gemischten Bearbeitung kompatibler und inkompatibler Durchgänge keine IAT-Effekte auftreten, da keine systematische Vereinfachung der Aufgabe möglich ist.

Methode

Um die Vorhersagen des Trägheits- und des Vereinfachungsmechanismus zu prüfen, wurde die Kompatibilität auf der Ebene einzelner Durchgänge manipuliert. Als Material dienten die gleichen geometrischen Objekte, die schon in Experiment 2 verwendet wurden. Im Gegensatz zu Experiment 2 wurde jedoch keine Kovariation zwischen Farbe und Größe der geometrischen Objekte realisiert. Die Aufgabenblöcke enthalten folglich sowohl Zielkonzeptreize, für die aufgrund von Farbe und Größe die gleiche Reaktion gefordert ist, also auch solche, für die Farbe und Größe verschiedene Reaktionen erfordern. Ein Durchgang wird als kompatibel kodiert, wenn Farbe und Größe eines Objekts die gleiche Reaktion erfordern,

als inkompatibel, wenn Größe und Farbe verschiedene Reaktionen erfordern, und als neutral, wenn ein Attributreiz bearbeitet werden musste. Attributreize sind, wie in herkömmlichen IATs, insofern neutral, als sie keiner der Zielkonzeptkategorien (blau oder rot) angehören. Deshalb ist bei dieser Aufgabe keine direkte Interferenz zu erwarten. Die Tastenzuordnung wurde abermals so ausbalanciert, dass eine Hälfte der Versuchsteilnehmer auf große und die andere Hälfte auf kleine Objekte mit der dominanten Hand reagieren musste.

Stichprobe

An der Untersuchung nahmen 24 Personen teil. 9 Teilnehmer waren männlichen und 15 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 20 und 48 Jahren ($M = 26.67$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Als Material dienten, wie in Experiment 2, geometrische Objekte die bezüglich ihrer Größe (groß versus klein), Farbe (rot, blau, gelb, grün und pink) und Form (Rechtecke, Kreise und Dreiecke) variierten. Es gab in jedem Aufgabenblock sowohl große als auch kleine rote Objekte und sowohl große als auch kleine blaue Objekte.

Prozedur

Die Versuchsteilnehmer bearbeiteten insgesamt 8 Aufgabenblöcke. Zunächst wurde in zwei Blöcken die Klassifikation der roten und blauen Objekte eingeübt, darauf folgend die Attributaufgabe in weiteren zwei Blöcken. Im Anschluss daran wurden die beiden Aufgaben in vier weiteren Aufgabenblöcken gemischt bearbeitet. Der Ablauf der einzelnen Versuchsdurchgänge entspricht dem von Experiment 2. Die Versuchsteilnehmer wurden nach jedem Aufgabenblock über ihre mittlere Reaktionsgeschwindigkeit und ihre Fehlerrate informiert. Die Bearbeitung des Experiments dauerte etwa 20 Minuten.

Ergebnisse

Wie in Experiment 2 wurden Antwortlatenzen unter 100 ms zu 100 ms (0.1 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.2 % der Durchgänge) rekodiert. Die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (5.6 % der Durchgänge) wurden nicht in die Analyse aufgenommen.

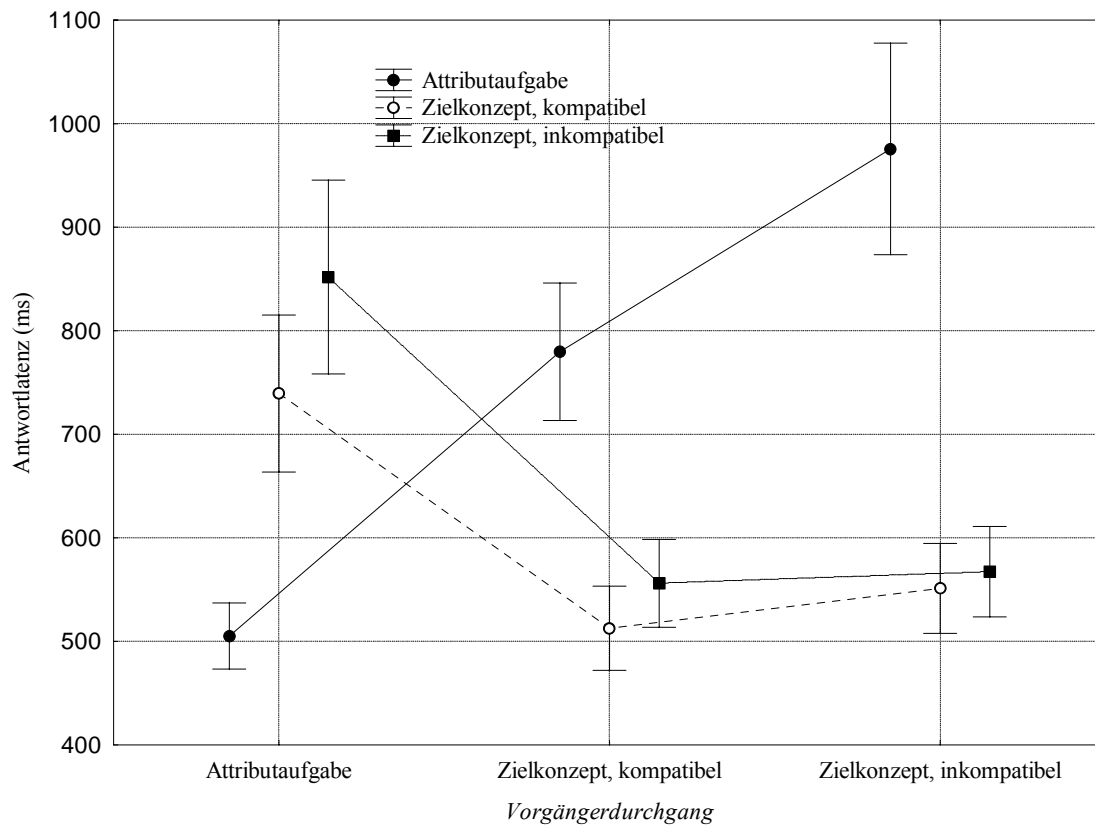


Abbildung 6: Mittlere Antwortlatenzen in Experiment 3, als Funktion des Aufgabentyps im aktuellen und im Vorgängerdurchgang. Unterschiedliche Linien stehen für die im aktuellen Durchgang bearbeitete Aufgabe. Die Performanz bei der Attributauflage hängt wie erwartet von der Art des Vorgängerdurchgangs ab. Die Performanz bei der Zielkonzeptaufgabe ist nur nach der Attributauflage eine Funktion der Kompatibilität, nicht jedoch bei Aufgabenwiederholung.

Die Daten wurden mit einer 3 (kompatibler, inkompatibler oder neutraler Durchgang) x 3 (kompatibler, inkompatibler oder neutraler Vorgängerdurchgang) x 2 (Tastenbalancierung) faktoriellen ANOVA mit Messwiederholung auf den ersten beiden Faktoren analysiert. Es ergibt sich ein signifikanter Haupteffekt für die Kompatibilität des aktuellen Durchgangs, $F(2, 44) = 19.24, p < .01$, der darauf zurückgeht, dass kompatible Durchgänge schneller bearbeitet wurden als inkompatible. Ferner zeigt sich ein bedeutsamer Haupteffekt der Kompatibilität des Vorgängerdurchgangs, $F(2, 44) = 45.98, p < .01$. Unabhängig von der Kompatibilität des aktuellen Durchgangs sind Reaktionen am langsamsten, wenn vorher die Attributauflage bearbeitet werden sollte, und am schnellsten, wenn vorher ein kompatibler Zielkonzeptreiz zu bearbeiten war. Nur ein weiterer Effekt überschreitet das Signifikanzniveau, namentlich eine Zweifach-Interaktion zwischen der Kompatibilität des aktuellen und des Vorgängerdurchgangs, $F(4, 88) = 114.06, p < .01$. Dieser Effekt ist Abbildung 6 dargestellt. Ein Kontrast

belegt, dass die Performanz bei der Attributaufgabe bedeutsam schlechter ist, wenn ein inkompatibler Durchgang vorausging, als wenn ein kompatibler vorausging, $F(1, 22) = 29.83$, $p < .01$. Ferner ist, wie erwartet, die Performanz bei inkompatiblen Durchgängen schlechter als bei kompatiblen, wenn diese der Attributaufgabe folgen, $F(1, 22) = 28.52$, $p < .01$.

Diskussion

Die Ergebnisse von Experiment 3 sprechen für den Trägheitsmechanismus. Kompatibilitätseffekte zeigen sich immer dann am stärksten, wenn die vorhergehend bearbeitete Aufgabe mit der aktuellen Aufgabe durch residuale Aktivierung beziehungsweise residuale Hemmung in Konflikt steht. Ein Hinweis darauf, dass die korrekte Bearbeitung inkompatibler Zielkonzeptreize zu einer Hemmung der Attributaufgabe und zu einem Übertrag (Carryover) dieser Hemmung führt (Allport et al., 1994; Wylie & Allport, 2000) ist vor allem, dass die Bearbeitung der Attributaufgabe nach inkompatiblen Zielkonzeptreizen stärker verlangsamt ist als nach kompatiblen. Dieser Effekt kann nicht unmittelbar auf Interferenzen zurückgehen, da die bei der Attributaufgabe zu bearbeitenden Reize keine interferierende Farbe aufweisen (den Farben der Attributemplare sind in der Zielkonzeptaufgabe keine Reaktionen zugeordnet).

Eine Veränderung globaler Kontrollstrategien, etwa in Form eines konservativeren Antwortkriteriums (Brendl et al., 2001), kann diesen Effekt nicht erklären, da die Attributaufgabe bei Aufgabenwiederholung offenbar ungehindert schnell bearbeitet wird. Das Auftreten von Kompatibilitätseffekten im vorliegenden Experiment spricht ferner gegen den Figur-Grund Ansatz von Rothermund und Wentura (2001), da farb- oder formabhängige Salienzunterschiede im vorliegenden Experiment kaum eine Rolle gespielt haben dürften. Der Vereinfachungsmechanismus lässt für das vorliegende Experiment ein Ausbleiben jeglicher IAT-artiger Effekte erwarten, was dem Befundmuster offenbar nicht entspricht. Weiterhin können die spezifischen Effekte der vorhergehend bearbeiteten Aufgabe durch diesen Mechanismus nicht erklärt werden.

Insgesamt folgen die Kompatibilitätseffekte und das spezifische Muster der Sequenzeffekte genau den Vorhersagen des an Allport et al. (1994) angelehnten Trägheitsmechanismus. Interessant ist, dass in der vorliegenden Untersuchung auch dann Aufgabenwechselkosten auftreten, wenn zu einem kompatiblen Durchgang hin beziehungsweise von einem kompatiblen Durchgang weg gewechselt werden muss. Dies legt nahe, dass Aufgabenwechsel auch in einem IAT kompatibilitätsunabhängige Kosten haben können. Diese

Annahme passt zu den Ergebnissen der Experimente 1 und 2, in denen auch in der kompatiblen Phase des IAT Aufgabenwechselkosten auftraten. Es scheint jedoch plausibel, dass im Verlauf der Bearbeitung eines herkömmlichen IAT durch die geblockte Bearbeitung der kompatiblen und inkompatiblen Phasen zusätzliche Prozesse ablaufen. Auf Basis der vorliegenden Befunde kann zum Beispiel nicht ausgeschlossen werden, dass zusätzlich zu den beschriebenen lokalen Sequenzeffekten globale Änderungen der Bearbeitungsstrategie oder chronische Änderungen des Aktivierungsgrads von Aufgabensets auftreten.

In Anbetracht der Unterschiede zwischen der hier verwendeten Aufgabe und einem herkömmlichen IAT stellt sich die Frage, inwieweit die berichteten Ergebnisse auf herkömmliche IATs übertragbar sind. Entscheidend ist in diesem Zusammenhang, dass die berichteten Kompatibilitätseffekte, wie in herkömmlichen IATs, auf der Kompatibilität von Eigenschaften derjenigen Reize beruhen, die der gleichen Antworttaste zugeordnet sind. Diese Mechanik weicht deutlich von derjenigen „klassischer“ Kompatibilitätseffekte ab. In Paradigmen wie dem Stroop-, dem Simon- oder dem semantischen Primingparadigma liegt den Effekten eine gewissermaßen inhärente Kompatibilität zwischen den Reiz- beziehungsweise Reaktionseigenschaften zugrunde. Der Stroop-Effekt zum Beispiel basiert auf einer Kompatibilität von Farbwort und Farbe, deren Kompatibilität schon präexperimentell gegeben ist. Der hier berichtete Effekt basiert im Gegensatz dazu *ausschließlich* auf der Zuordnung der Reizeigenschaft zu den Reaktionen: Es gibt keine präexperimentelle Verbindung zwischen der Farbe und der Größe geometrischer Objekte. Jeglicher Einfluss der irrelevanten Größe roter und blauer Objekte auf die Antwortlatenzen muss also auf die instruierte Zuordnung der Größe zu einer der Antworttasten zurückgehen.

In diesem Zusammenhang wird deutlich, warum Effekte der irrelevanten Attributinformation nach Aufgabenwechseln deutlich vergrößert sind. Durch die Anwendung der Attributaufgabe wird das entsprechende Aufgabenset aktiviert. Die in diesem Aufgabenset repräsentierte Verbindung zwischen Attribut und Reaktion hat bei einem nachfolgenden Wechsel zur Zielkonzeptaufgabe eine verstärkte Interferenz zur Folge, wenn der Durchgang inkompatibel ist. Die vorliegenden Ergebnisse sprechen für diese Interpretation: Wurde im Vorgängerdurchgang die Attributaufgabe bearbeitet, so zeigt sich ein deutlicher Unterschied zwischen kompatiblen und inkompatiblen Zielkonzeptreizen, der der Asymmetrie der Aufgabenwechselkosten in kompatibler und inkompatibler Phase eines herkömmlichen IATs entspricht. Wurde hingegen im Vorgängerdurchgang ebenfalls die Zielkonzeptaufgabe

bearbeitet, so zeigen sich kaum Unterschiede zwischen kompatiblen und inkompatiblen Durchgängen.

5 Experimente zu globalen Effekten

In den Experimenten 4 bis 6 sollen globale Nachwirkungen der Bearbeitung von IAT-Blöcken untersucht werden. Zwei Phänomene, die eine Auswirkung längerfristiger Nachwirkungen sein können, sind der Effekt der Kompatibilitätsreihenfolge und der Befund, dass die Retest-Reliabilität meist geringer ausfällt als die interne Konsistenz (Asendorpf et al., 2002; Bosson et al., 2000; Egloff & Schmukle, 2002; Greenwald & Farnham, 2000; Greenwald et al., 1998; Nosek, Greenwald & Banaji, 2003; Steffens & Buchner, 2003). Beide Phänomene spiegeln insofern Nachwirkungen wider, als sie auf die vorhergehende Bearbeitung einzelner Blöcke beziehungsweise auf eine komplette Testwiederholung zurückgehen müssen. Durch die hier berichteten Experimente sollen dabei einige Vorhersagen der verschiedenen Prozessmodelle des IAT zu solchen Effekten geprüft und ferner die Grenzen und Bedingungen für ihr Auftreten geklärt werden.

5.1 Experiment 4

In Experiment 4 sollen globale Änderungen der Bearbeitungsstrategie beim Übergang zwischen kompatibler und inkompatibler Phase eines IAT genauer betrachtet werden. Bevor die Vorhersagen der verschiedenen Prozessmodelle des IAT-Effekts dargestellt werden, seien zunächst einige methodische Probleme, die bei der Untersuchung dieses Übergangs in einem herkömmlichen IAT auftreten, beleuchtet. Aus experimenteller Sicht ergeben sich Probleme vor allem als Folge zweier Konfundierungen des Übergangs zwischen den Phasen eines IAT mit theoretisch und inhaltlich wenig relevanten Störvariablen. Zum einen erfordert der Wechsel zwischen kompatibler und inkompatibler Phase in herkömmlichen IATs zwingend ein Umlernen der Tastenzuordnung, was zu einer Konfundierung zwischen dem reinen Lernen der Tastenzuordnung und der Anpassung an die geänderte Kompatibilität führt. Zum anderen werden Effekte des Übergangs zwischen den Phasen eines IAT durch die üblicherweise schlechtere Performanz in den ersten Durchgängen eines Aufgabenblocks belastet.

Eine nahe liegende Lösung dieses Problems wäre, jeweils spezifische Kontrollbedingungen einzuführen, in denen nur der Effekt des Blockbeginns bzw. nur der Effekt des

Umlernens, jeweils ohne Wechsel der Kompatibilität erfasst wird. Das vorliegende Experiment realisiert jedoch eine ökonomischere Lösung, die sich der Flexibilität des in Experiment 2 und 3 verwendeten geometrischen Materials bedient. Anders als in herkömmlichen IATs kann die Kompatibilität mit diesem Material stimulusseitig manipuliert werden, also ohne Änderung der Instruktion. Dadurch kann die Kompatibilität innerhalb eines Aufgabenblocks unabhängig von Blockwechselln und dem Umlernen der Tastenzuordnung geändert werden. Werden zum Beispiel rote und kleine Objekte einer Taste zugeordnet, so können stimulusseitig entweder alle roten Objekte klein oder alle roten Objekte groß sein. Im ersten Fall ergibt sich ein kompatibler Block, im zweiten ein inkompatibler. Es ist leicht zu sehen, dass ein entsprechender Wechsel der Kovariation zwischen Farbe und Größe der Objekte innerhalb eines Blocks erfolgen kann, ohne dass die Versuchsperson darüber informiert werden muss. Die auftretenden Übergangseffekte und ihr Verlauf sollten frei von den oben angesprochenen mehr oder weniger trivialen Lernprozessen und Orientierungsreaktionen sein. Aus den Prozessmodellen des IAT-Effekts lassen sich, einige spezifische Vorhersagen für den Verlauf dieses Übergangs ableiten:

Random-Walk Modell: Das Random-Walk Modell (Brendl et al., 2001) spezifiziert nicht, welche Prozesse zu der angenommenen Aufgabenrepräsentation in Form eines Random-Walk führen. In dieser Hinsicht macht das Modell keine Vorhersagen über globale Veränderungen im Verlauf der Bearbeitung eines IAT. Die angenommene Verschiebung des Antwortkriteriums in der inkompatiblen Phase eines IAT als Reaktion auf die größere Aufgabenschwierigkeit lässt jedoch erwarten, dass die Latenzen in dieser Phase im Verlauf der Bearbeitung ansteigen. Eine globale Veränderung des Antwortkriteriums könnte ferner zur Folge haben, dass die Bearbeitung der kompatiblen Phase nach der inkompatiblen verlangsamt ist.

Vereinfachungsmechanismus: Der Vereinfachungsmechanismus (Mierke & Klauer, 2001, 2003) nimmt an, dass die Bearbeitung der kompatiblen Phase eines IAT durch eine Anwendung der Attributaufgabe auf alle präsentierten Reize vereinfacht wird. Es ist zu erwarten, dass diese Vereinfachung erst nach einer Reihe von Durchgängen erfolgt, vermittelt durch einen impliziten oder expliziten Lernprozess. Es ist also zu erwarten, dass sich die Latenzen in der kompatiblen Phase im Verlauf der Bearbeitung verkürzen. In der inkompatiblen Phase ist keine Veränderung zu erwarten.

Trägheitsmechanismus. Der Trägheitsmechanismus (Mierke & Klauer, 2001, 2003) erklärt den IAT-Effekt durch lokale Sequenzeffekte. Deshalb ist zu erwarten, dass

Unterschiede zwischen kompatibler und inkompatibler Phase unmittelbar auftreten. Bei der geblockten Bearbeitung eines IAT sind jedoch möglicherweise auch längerfristige Nachwirkungen zu erwarten. Solche Effekte werden in den Experimenten 5 und 6 näher untersucht.

Stimulus-Response Kompatibilität: De Houwer (2001b) nimmt an, dass der IAT-Effekt über eine Kompatibilität zwischen Reizbedeutung und einer erworbenen Konnotation der Antworttasten vermittelt ist. Der Erwerb dieser Konnotation erfolge schon während der Aufgabeninstruktion. Eine Erleichterung der kompatiblen Phase wäre somit unmittelbar bei den ersten Durchgängen eines Aufgabenblocks zu erwarten. Da der IAT-Effekt nach De Houwer aber auf der Attributausprägung der Konzepte beruht, ist im vorliegenden Fall überhaupt kein Effekt zu erwarten: Die Zielkonzepte weisen keine präexperimentelle Assoziation mit dem Attribut auf.

Auch das *Figur-Grund Modell* (Rothermund & Wentura, 2001) und das Modell der Environmental Associations (Karpinski & Hilton, 2001) lassen keinerlei IAT-Effekte für das verwendete Material erwarten, da präexperimentell weder systematische Salienzunterschiede noch kulturell bedingte Assoziationen vorliegen dürften, die von diesen Modellen postulierten notwendigen Bedingungen also nicht erfüllt sind.

Methode

Die Teilnehmer von Experiment 4 bearbeiteten insgesamt 8 gemischte Aufgabenblöcke mit kompatibler und inkompatibler Tastenzuordnung, denen vier Übungsblöcke für die Attribut- und Zielkonzeptaufgabe vorangestellt waren. Für eine Hälfte der Versuchsteilnehmer wurde die Kompatibilität innerhalb des vierten gemischten Blocks, der im Folgenden als Übergangsblock bezeichnet wird, umgekehrt. Im Gegensatz zu herkömmlichen IATs wurde dies nicht durch eine Änderung der Tastenzuordnung erreicht, sondern durch Manipulation der Stimuluszusammensetzung. War etwa für rote und kleine Objekte die linke Taste zu drücken, so waren in der kompatiblen Bedingung zunächst alle roten Objekte klein und alle blauen Objekte groß. In der Mitte des Übergangsblocks wurde dann die Stimuluszusammensetzung so geändert, dass alle roten Objekte groß und alle blauen Objekte klein waren. Die schon eingeübte Tastenzuordnung blieb unverändert, war aber nun in Bezug auf die präsentierten Objekte inkompatibel. Die Kompatibilität der Durchgänge vor und nach dem Übergangsblock wurde unabhängig voneinander between-subjects manipuliert. Dadurch ergeben sich vier Arten von Kompatibilitätsübergängen: „kompatibel → kompatibel“,

„kompatibel → inkompatibel“, „inkompatibel → kompatibel“ und „inkompatibel → inkompatibel“.

Stichprobe

An der Untersuchung nahmen 48 Personen teil. 17 Teilnehmer waren männlichen und 31 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 18 und 45 Jahren ($M = 24.17$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Als Material dienten die schon in den Experimenten 2 und 3 verwendeten geometrischen Objekte, die in Größe (groß versus klein), Farbe (rot, blau, gelb, grün und pink) und Form (Rechtecke, Kreise und Dreiecke) variierten.

Prozedur

Die Versuchsteilnehmer von Experiment 4 bearbeiteten insgesamt 12 Aufgabenblöcke mit jeweils 48 Durchgängen. Zu Beginn des Experiments wurden die Zielkonzeptaufgabe und die Attributaufgabe in jeweils zwei Aufgabenblöcken getrennt eingeübt. Es folgten insgesamt acht gemischte Blöcke, in denen die Größen- und die Farbklassifikation gemischt bearbeitet werden sollten. Eine Hälfte der Versuchsteilnehmer bearbeitete die ersten 3 gemischten Blöcke sowie die ersten 24 Durchgänge des Übergangsblocks mit kompatibler Tastenzuordnung, die zweite Hälfte der Versuchsteilnehmer mit inkompatibler. Die Kompatibilität der dargebotenen Zielkonzeptreize wechselte für die Hälfte der Teilnehmer nach den ersten 24 Durchgängen des Übergangsblocks. Für die restlichen Teilnehmer blieb die Kompatibilität über das komplette Experiment konstant. Die Tastenzuordnung blieb in beiden Fällen unverändert. Um kurze zeitliche Verläufe auf möglichst vielen Einzeldurchgängen abzubilden, lag das RSI bei 100 ms. Wie in den Vorgängerexperimenten erhielten die Versuchsteilnehmer abschließend Rückmeldung über ihr persönliches Abschneiden und konnten Informationen über den IAT und die spezifischen Ziele der Untersuchung abrufen.

Ergebnisse

Wie in Experiment 2 wurden Antwortlatenzen unter 100 ms zu 100 ms (0.1 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.3 % der Durchgänge) rekodiert. Die

ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (5.2 % der Durchgänge) wurden nicht in die Analyse aufgenommen. Im Folgenden werden zwei getrennte Analysen für unmittelbare und längerfristige Effekte des Kompatibilitätsübergangs berichtet. Unmittelbare Effekte beziehen sich hierbei auf den Kontrast der ersten und zweiten Hälfte des Übergangsblocks, also der Durchgänge unmittelbar vor und nach dem Übergang. Längerfristige Effekte beziehen sich auf den Kontrast aller experimentellen Durchgänge vor und nach dem Kompatibilitätsübergang.

Unmittelbare Effekte

Die Antwortlatenzen des Übergangsblocks wurden mit einer 4 (Art des Kompatibilitätsübergangs) x 2 (1. vs. 2. Hälfte des Übergangsblocks) faktoriellen Varianzanalyse mit Messwiederholung auf dem zweiten Faktor analysiert. Die vier Faktorstufen des Faktors Kompatibilitätsübergang entsprechen dabei den Kombinationen „kompatibel → kompatibel“, „kompatibel → inkompatibel“, „inkompatibel → kompatibel“ und „inkompatibel → inkompatibel“. Die dieser Analyse zugrunde liegenden mittleren Antwortlatenzen sind in Abbildung 7 dargestellt.

Dabei ergeben sich ein Haupteffekt des Faktors Blockhälfte, $F(1, 44) = 19.11, p < .01$, ein Haupteffekt des Faktors Kompatibilitätsübergang, $F(3, 44) = 14.12, p < .01$, sowie eine Zweifach-Interaktion von Kompatibilitätsübergang und Blockhälfte, $F(3, 44) = 15.03, p < .01$. Diese Interaktionen sind aufgrund der komplexen Struktur des Experiments zunächst schwer interpretierbar. Geplante Vergleiche zeigen, dass dieses Ergebnismuster vor allem auf zwei spezifische Eigenschaften der Daten zurückgeht. Eine dieser Eigenschaften ist, dass die Latenzen in der Bedingung mit kompatiblen Reizen in der ersten und zweiten Experimenthälfte signifikant geringer ausfallen als die Latenzen in den anderen drei Bedingungen, $F(1, 44) = 42.12, p < .01$. Die zweite zentrale Eigenschaft des Datenmusters ist, dass nur in der Bedingung „kompatibel → inkompatibel“ ein unmittelbarer Effekt des Kompatibilitätsübergangs auftritt. Der Unterschied zwischen den kompatiblen Durchgängen zu Beginn des Blocks und den inkompatiblen Durchgängen der zweiten Hälfte beträgt 521 ms und ist mit $F(1, 44) = 63.91, p < .01$, statistisch signifikant. Dieser Effekt ist insbesondere deshalb bemerkenswert, weil er innerhalb einer kurzen Frist von 24 Durchgängen auftritt, und die Aufgabe in den vergangenen 168 Durchgängen in unveränderter Form eingeübt wurde. In keiner der drei anderen Übergangsbedingungen treten vergleichbare Effekte auf. In diesen Bedingungen schwanken die Effekte zwischen 4 ms und 33 ms, mit nicht-signifikanten F -

Werten zwischen $F(1, 44) = 0.00$ und $F(1, 44) = 0.27$. Interessant ist an diesem Befund, dass die 24 kompatiblen Durchgänge in der Bedingung „inkompatibel-kompatibel“ keineswegs schneller bearbeitet werden, als die 24 inkompatiblen Durchgänge, die ihnen vorangingen. Diese Durchgänge werden nicht schneller bearbeitet als die 24 inkompatiblen Durchgänge, denen inkompatible Durchgänge vorausgingen, $F(1, 44) = 0.03$, n.s. und sind im Vergleich zu den Latenzen für kompatible Durchgänge, denen kompatible Durchgänge vorausgingen, deutlich verlangsamt, $F(1, 44) = 16.33$, $p < .01$.

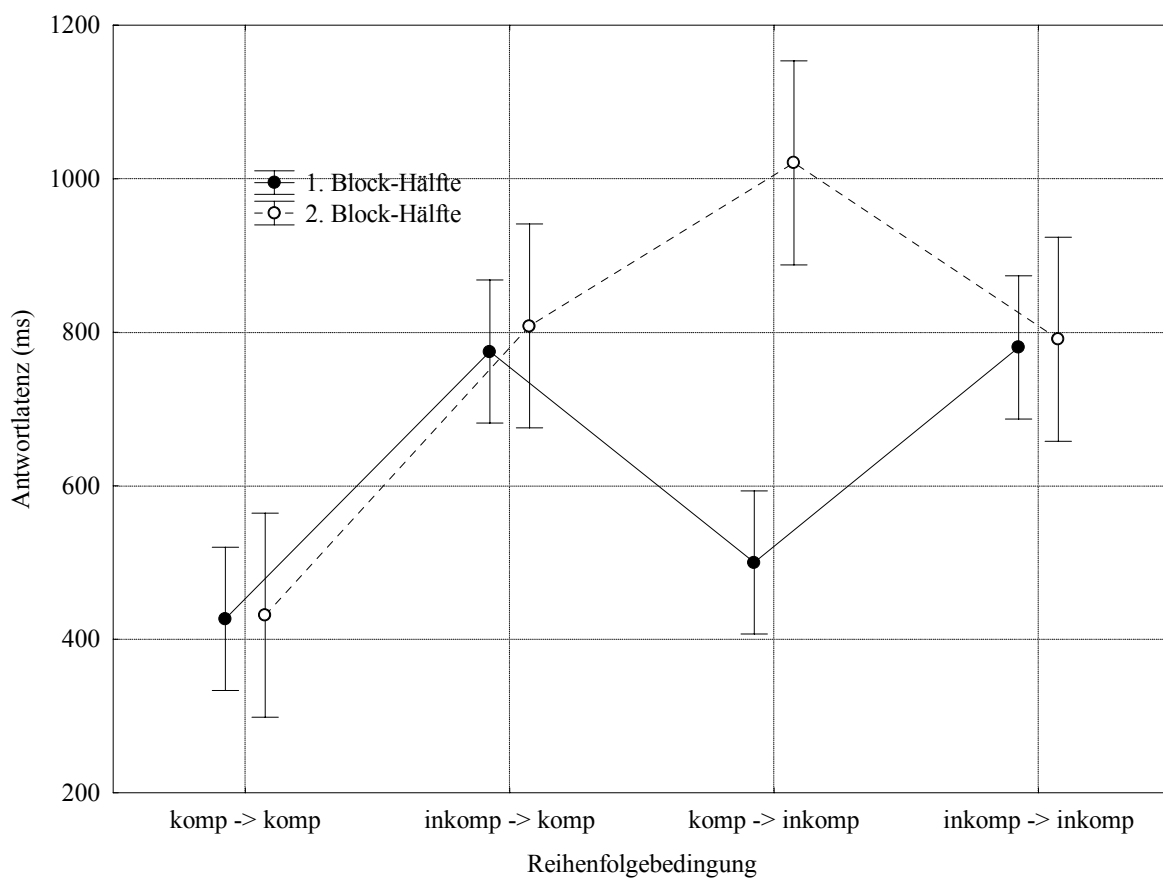


Abbildung 7: Mittlere Antwortlatenzen in der ersten und zweiten Hälfte des Übergangsblocks in Experiment 4 als Funktion der Reihenfolgebedingung. Der Übergang von kompatiblen zu inkompatiblen Reizen zeigt einen unmittelbaren Effekt auf die Performanz. Es findet sich jedoch kein unmittelbarer Effekt beim Übergang von inkompatiblen zu kompatiblen Reizen.

Zusammenfassend zeigt sich also eine unmittelbare Verlangsamung beim Übergang zu inkompatiblen Stimuli, jedoch keine unmittelbare Beschleunigung für den Übergang zu

kompatiblen Reizen. Kompatible Durchgänge werden in diesem Fall genauso langsam bearbeitet wie vorangehende inkompatible Durchgänge in der gleichen Bedingung.

Längerfristige Effekte

Die Antwortlatenzen der acht gemischten Aufgabenblöcke wurden weiterhin mit einer 4 (Art des Kompatibilitätsübergangs) x 2 (1. vs. 2. Hälfte des Experimentes) faktoriellen Varianzanalyse mit Messwiederholung auf dem zweiten Faktor analysiert, um längerfristige Effekte des Kompatibilitätsübergangs zu prüfen.

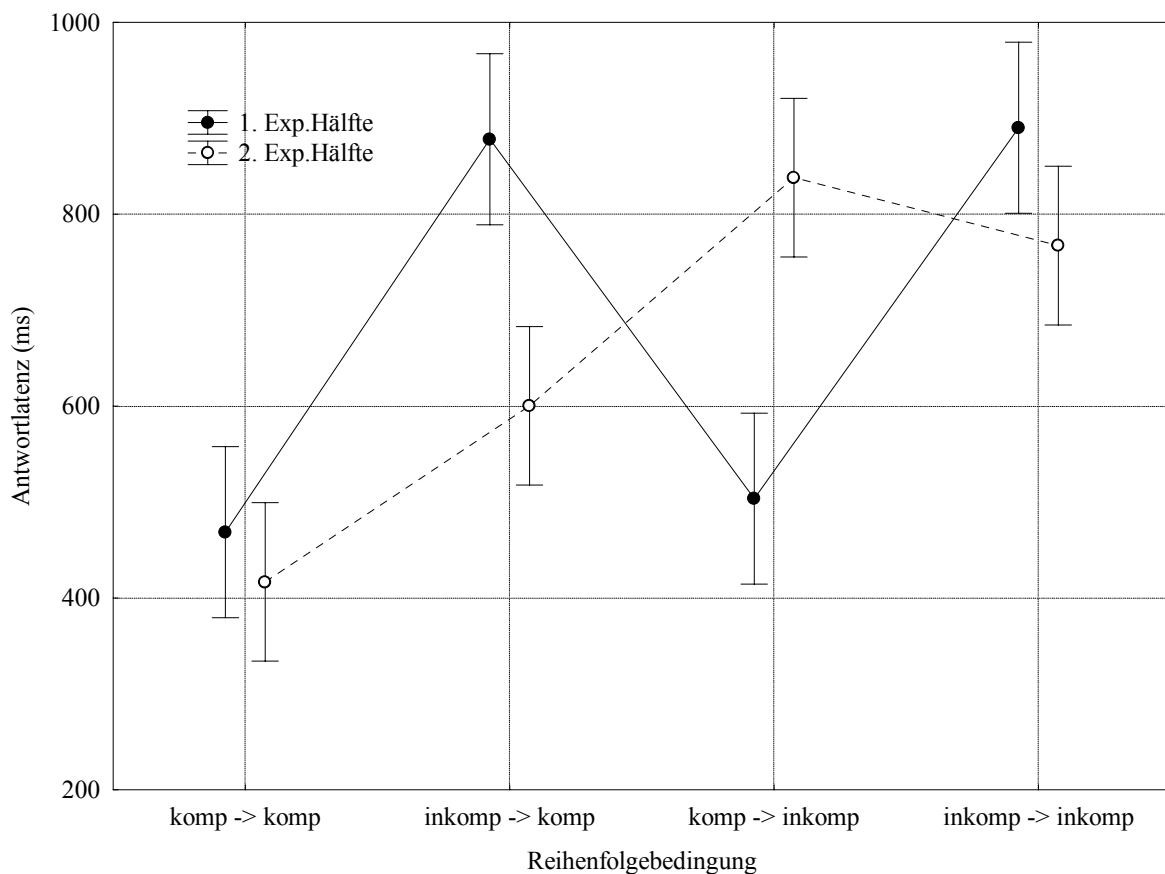


Abbildung 8: Mittlere Antwortlatenzen in der ersten versus zweiten Hälfte von Experiment 4 als Funktion der Reihenfolgebedingung. Längerfristige Effekte zeigten sich sowohl für den Übergang von kompatiblen zu inkompatiblen Reizen, als auch für den Übergang von inkompatiblen zu kompatiblen.

Dabei zeigt sich, dass die Antworten in der zweiten Hälfte des Experiments im Mittel schneller sind als in der ersten, $F(1, 44) = 5.93, p < .05$. Ferner findet sich ein Haupteffekt des Faktors Kompatibilitätsübergang, $F(3, 44) = 16.21, p < .01$, und eine Zweifach-Interaktion

zwischen Experimenthälfte und Kompatibilitätsübergang, $F(3, 44) = 115.85, p < .01$. Das Muster der längerfristigen Effekte unterscheidet sich dabei in verschiedener Hinsicht von dem, der kurzfristigen Effekte. Zunächst finden sich signifikante Übungseffekte sowohl für die Bedingung „kompatibel \rightarrow kompatibel“, $F(1, 44) = 4.58, p < .05$, als auch für die Bedingung „inkompatibel \rightarrow inkompatibel“, $F(1, 44) = 25.77, p < .01$. In beiden Fällen ist die Performanz in der zweiten Hälfte des Experiments schneller. Der Wechsel von inkompatibel zu kompatibel führt längerfristig zu der erwarteten Beschleunigung der Antworten, $F(1, 44) = 131.86, p < .01$. Der Wechsel von kompatibel zu inkompatibel führt, wie schon bei den kurzfristigen Effekten, zu einer Verlangsamung, $F(1, 44) = 191.27, p < .01$. Abbildung 8 zeigt die mittleren Antwortlatenzen in der ersten gegenüber der zweiten Hälfte des Experiments.

Die Latenzen auf kompatible Durchgänge in der zweiten Experimenthälfte sind in der Bedingung „inkompatibel \rightarrow kompatibel“ langsamer als in der Bedingung „kompatibel \rightarrow kompatibel“, $F(1, 44) = 10.07, p < .01$. Die Bearbeitung der kompatiblen Durchgänge ist also abhängig von der Vorgängerphase. Für inkompatible Durchgänge findet sich kein analoger Effekt, $F(1, 44) = 1.49, n.s.$ Der interessanteste Befund bei der Analyse unmittelbarer Effekte ist wohl das Ausbleiben eines Unterschieds zwischen kompatiblen und inkompatiblen Durchgängen, wenn vorher 168 inkompatible Durchgänge bearbeitet wurden. Ein analoger Kontrast über das Experiment hinweg zeigt, dass dieser Effekt längerfristig durchaus auftritt, $F(1, 44) = 8.26, p < .01$. Dies spricht dafür, dass sich die relativ schnellere Bearbeitung der kompatiblen Phase erst nach einer größeren Zahl von Durchgängen einstellt, wie es der Vereinfachungsmechanismus erwarten lässt.

Diskussion

Die Ergebnisse von Experiment 4 liefern einige interessante Einblicke in den Prozess des Übergangs zwischen den verschiedenen Phasen eines IAT. Es zeigt sich, dass inkompatible Durchgänge unmittelbar schwerer zu bearbeiten sind als kompatible, wie es unter anderem auf Basis des Trägheitsmechanismus zu erwarten wäre. Ginge dieser Effekt auf eine Anpassung des Antwortkriteriums zurück (Brendl et al., 2001), müsste diese schon während der ersten 24 Durchgängen der inkompatiblen Phase erfolgt sein, um die vorliegenden Ergebnisse erklären zu können. Die Erleichterung bei der Bearbeitung kompatibler Durchgänge scheint erst nach mehr als 24 Durchgängen aufzutreten. Dass kompatible Durchgänge zunächst nicht schneller bearbeitet werden als inkompatible, passt zu den Annahmen des Vereinfachungsmechanismus. Die Tatsache, dass die Bearbeitung inkompatibler Durchgänge

in der zweiten Experimenthälfte unabhängig von der zuvor bearbeiteten Phase ist, spricht gegen eine vergleichbare Anpassung der globalen Bearbeitungsstrategie bei der inkompatiblen Tastenzuordnung.

Aufgrund der stimulseitigen Manipulation des Kompatibilitätsübergangs in Experiment 4 kann ferner ausgeschlossen werden, dass der Effekt der Kompatibilitätsreihenfolge (siehe oben) mit dem Umlernen der Tastenzuordnung zusammenhängt. Vielmehr scheinen die strukturellen Unterschiede zwischen kompatibler und inkompatibler Tastenzuordnung hinreichend zu sein, um eine unmittelbare Verlangsamung der inkompatiblen Phase und eine längerfristige Erleichterung der kompatiblen Phase zu bewirken. Auf Basis der vorliegenden Befunde ist anzunehmen, dass Effekte der Kompatibilitätsreihenfolge erst im Verlauf mehrerer Blöcke abklingen. Dies passt zu der Folgerung von Nosek, Greenwald und Banaji (2003), dass Reihenfolgeeffekte durch verlängerte Übungsphasen verringert werden können.

Eine Rekodierung der Aufgabe, wie sie der Vereinfachungsmechanismus postuliert, oder das Ausnutzen einer Kompatibilität zwischen den Reizen und Reaktionen, wie es das Modell von De Houwer (De Houwer, 2001b) annimmt, muss nicht auf bewusste Prozesse zurückgehen. Eine solche Anpassung könnte auf ähnlichen Prozessen beruhen, wie der Konsistenzanteilseffekt im affektiven Priming (Klauer, Rossnagel & Musch, 1997; Musch, 2000). Dieser Effekt basiert auf einer stufenweisen Manipulation des Anteils evaluativ konsistenter Wortpaare in einem Aufgabenblock. Die Stärke des Primingeffekts steigt dabei proportional mit dem relativen Anteil konsistenter Wortpaare. Die durchmischte Darbietung kompatibler und inkompatibler Reize in Experiment 3 entspricht dabei einem Konsistenzanteil von 50%, da jeweils die Hälfte der Durchgänge kompatibel beziehungsweise inkompatibel ist. Die geblockte Bearbeitung kompatibler und inkompatibler Durchgänge in Experiment 4 und herkömmlichen IATs entspricht einem Konsistenzanteil von 100%, da jeweils alle Durchgänge eines kompatiblen IAT-Blocks kompatibel ausfallen.

Neben stärkeren Primingeffekten in Aufgabenblöcken mit hohem Konsistenzanteil weist der Konsistenzanteilseffekt noch einige andere interessante Eigenschaften auf (Musch, 2000). Zum Beispiel bleibt er auch dann noch erhalten, wenn nach einigen Blocks mit hohem Konsistenzanteil wieder gleich viele konsistente wie inkonsistente Wortpaare dargeboten werden. Dies spricht für eine globale Änderung der Bearbeitungsstrategie, die mit einer gewissen Trägheit an die aktuelle Blockstruktur angepasst wird. Der Konsistenzanteilseffekt tritt interessanterweise nicht auf, wenn die Primewörter über das Experiment hinweg visuell

maskiert werden. Offenkundig setzt der Effekt also eine Bewusstheit der Primewörter voraus. Werden die Primewörter jedoch erst nach einer Reihe von Aufgabenblöcken maskiert dargeboten, treten die Effekte auch mit maskierten Primewörtern auf. Während die Anpassung an einen höheren Konsistenzanteil also offenkundig eine teilweise Bewusstheit der Primewörter erfordert, tritt der eigentliche Effekt auch mit nicht bewusst wahrnehmbaren Primewörtern auf. Musch (2000) erklärt dieses Muster durch eine kontextabhängige Gewichtung der irrelevanten Primeinformation, die im Verlauf der Bearbeitung angepasst wird. In gewisser Hinsicht entspricht diese Erklärung dem Vereinfachungsmechanismus des IAT-Effekts, der von einer stark positiven Gewichtung der irrelevanten Attributausprägung der Zielkonzeptexemplare ausgeht.

Nimmt man die Ergebnisse der Experimente 3 und 4 zusammen, scheint der IAT-Effekt im Zusammenspiel zweier Faktoren zu entstehen. Den einen Faktor bilden lokale Sequenzeffekte, wie sie in Experiment 3 aufgetreten sind. Diese Effekte sind unabhängig davon, ob kompatible und inkompatible Durchgänge geblockt bearbeitet werden oder, wie in Experiment 3, innerhalb eines Blockes gemischt. Solche lokalen Effekte können die unmittelbar erhöhte Schwierigkeit der inkompatiblen Phase erklären, da sie auch auftreten, wenn der Übergang zwischen kompatiblen und inkompatiblen Durchgängen mehrfach innerhalb eines Aufgabenblocks erfolgt. Der zweite Faktor sind längerfristige globale Änderungen, wie sie höchstwahrscheinlich den verzögert auftretenden Effekten bei der Bearbeitung der kompatiblen Phase zugrunde liegen. Die Wirksamkeit dieses zweiten Faktors setzt voraus, dass kompatible und inkompatible Durchgänge geblockt bearbeitet werden, wie es in einem herkömmlichen IAT immer der Fall ist.

5.2 Experiment 5

Experiment 4 zeigt, dass die Bearbeitung kompatibler Durchgänge nach Aufgabenblöcken mit inkompatiblen Durchgängen unmittelbar und längerfristig verlangsamt ist. Die Leistung in der inkompatiblen Phase selbst scheint im Gegensatz dazu unabhängig von ihrer Position im Experiment zu sein. Dieses Befundmuster lässt sich wie oben diskutiert über den Vereinfachungsmechanismus erklären. Eine ebenso plausible Erklärung wäre jedoch, dass die Effekte vor allem auf längerfristige Nachwirkungen der inkompatiblen Phase zurückgehen. In Experiment 5 sollen deshalb drei alternative Hypothesen zur Ursache des Effekts geprüft werden, die auch dem vielfach beobachteten Effekt der Kompatibilitätsreihenfolge zugrunde liegen könnten.

Die erste Hypothese basiert auf Befunden von Allport et al. (1994), die zeigen, dass Aufgabensets eine längerfristige Trägheit (Inertia) aufweisen. Trägheit beschreibt in diesem Zusammenhang die Tatsache, dass ein einmal aktiviertes Aufgabenset über eine substantielle Zeitspanne hinweg in einem erhöhten Aktivierungszustand verbleibt, während vorhergehend unterdrückte Aufgabensets, *über viele Durchgänge hinweg* schwerer zu aktivieren sind (siehe auch Mayr & Keele, 2000). In den Untersuchungen von Allport und Kollegen bearbeiteten die Versuchsteilnehmer Stroop-Aufgaben, bei denen zwischen dem Nennen der Farbe und dem Vorlesen des Farbwortes zu wechseln war. In diesen Untersuchungen trat konsistent eine stärkere Interferenz durch die irrelevante Stimuluseigenschaft auf, wenn diese in einem kürzlich angewendeten Aufgabenset relevant war. Interessanterweise dauern solche Nachwirkungen über eine verhältnismäßig lange Zeitspanne hinweg an (Allport et al., 1994; Wylie & Allport, 2000). Eine sparsame Erklärung des Befundmusters von Experiment 4 besteht nun darin, dass die wiederholte Hemmung des Attributaufgabensets in der inkompatiblen Phase eines IAT zu einer längerfristigen Hemmung der Anwendung dieser Aufgabe führt, die auf eine folgende kompatible Phase ausstrahlt. Insbesondere sollte die Verarbeitung der Attributausprägung von Zielkonzeptreizen erschwert sein, da vornehmlich bei diesen Interferenzen zu erwarten sind. Wenn diese Hypothese zutrifft, so sollten Effekte der inkompatiblen Phase nur dann auftreten, wenn diese das Wechseln zwischen Aufgaben erfordert. Ferner sollten die Effekte spezifisch die Attributaufgabe betreffen, nicht jedoch andere Aufgabentypen.

Eine weitere Hypothese lässt sich aus dem Random-Walk Modell des IAT (Brendl et al., 2001) ableiten. Wie weiter oben bereits erläutert, kommt es nach diesem Modell bei der Bearbeitung mit inkompatibler Tastenzuordnung zu einer Verschiebung des Antwortkriteriums. Nimmt man zusätzlich an, dass diese Verschiebung über mehrere Durchgänge hinweg erhalten bleibt, so kann der Kompatibilitätsreihenfolgeeffekt recht elegant erklärt werden. Wird die kompatible Phase nach der inkompatiblen bearbeitet, so wirkt sich das konservativere Antwortkriterium auf beide Phasen aus. Wird hingegen die kompatible Phase zuerst bearbeitet, ist offenbar kein Effekt der zeitlich nachgeordneten inkompatiblen Phase zu erwarten. Eine Kernnahe des Modells von Brendl et al. (2001) ist, dass die Verschiebung des Antwortkriteriums nicht aufgabenspezifisch erfolgt. Diese Annahme ist notwendig, um zu erklären, warum auch bei der Attributaufgabe Kompatibilitätseffekte auftreten. Wenn der Reihenfolgeeffekt tatsächlich auf eine Verschiebung des Antwortkriteriums zurückgeht, so sollten entsprechende Effekte auch bei anderen, inhaltlich nicht überlappenden nachgeordneten Aufgaben auftreten. In Experiment 5 wurde deshalb geprüft, ob die Bearbeitung

der inkompatiblen Phase des IAT Auswirkungen auf eine nachgeordnete Farbnennungsaufgabe hat.

Eine dritte Hypothese basiert auf der Annahme, dass die Verknüpfung der irrelevanten Attributausprägung mit den instruierten Antworttasten bei der Bearbeitung der inkompatiblen Phase eines IAT inhibiert wird, um Interferenzen aufzulösen. Diese Inhibition spezifischer Verknüpfungen von Reizen und Reaktionen könnte auf die kompatible Phase ausstrahlen. Der entscheidende Unterschied zur ersten Hypothese ist, dass eine derartige Inhibition sowohl für die verwendeten Reize als auch für die Reaktionen spezifisch sein sollte. Nachwirkungen der inkompatiblen Phase sollten also nur dann auftreten, wenn die nachfolgende Aufgabe sowohl die gleichen Reize als auch die gleichen Reaktionen verwendet. In Experiment 5 wurde deshalb geprüft, ob sich differentielle Nachwirkungen der Bearbeitung von IAT-Blöcken auch dann zeigen, wenn die Antwortmodalität gewechselt wird.

Methode

Es wurden zwei Arten von Aufgabenblöcken realisiert: Einerseits kompatible oder inkompatible IAT-Blöcke, die zur Erzeugung von Nachwirkungen dienen sollten, und andererseits Bewertungsblöcke, mit denen eventuelle Nachwirkungen erfasst werden sollten. In den Bewertungsblöcken wurden in schwarzer Farbe präsentierte Wörter und rot oder blau eingefärbte, sinnlose Konsonantenketten dargeboten. Die Versuchsteilnehmer waren instruiert, die schwarz dargebotenen Wörter durch Aussprechen der Antworten „positiv“ oder „negativ“ zu bewerten beziehungsweise die Farbe der Konsonantenketten zu benennen. Bei den Wörtern handelte es sich um Exemplare der Zielkonzepte aus den IAT-Blöcken, das heißt, um Blumen- und Insektenbezeichnungen. Durch den Wechsel von Tastendrücken zur Benennung der Valenz sollte ausgeschlossen werden, dass die Nachwirkungen auf eine Hemmung spezifischer Verknüpfungen von Reiz und Reaktion zurückgehen. Durch die zusätzliche Farbnennungsaufgabe sollte untersucht werden, ob die Nachwirkungen spezifisch für die in einem IAT-Block vorkommenden Aufgaben sind, oder aufgabenunspezifisch - wie auf Basis des Random-Walk Modells zu erwarten – auf andere Aufgaben ausstrahlen.

Ferner sollte geprüft werden, ob die Nachwirkungen eine Konsequenz der Aufgabenwechsel in den IAT-Blöcken sind. Dazu wurde die Mischung von Bewertungs- und Kategorisierungsaufgabe - und damit die Zahl der Aufgabenwechsel in der jeweiligen Bedingung - in zwei Stufen between-subjects manipuliert. In der entmischten Kontrollbedingung wurde jeder Aufgabenblock in zwei Hälften aufgeteilt. Dabei wurden in der ersten

Hälfte eines Blocks entweder nur Zielkonzeptentscheidungen oder nur Attributentscheidungen gefordert. In der zweiten Hälfte des Blocks musste die jeweils verbleibende Aufgabe bearbeitet werden. Durch diesen Aufbau resultiert lediglich ein Aufgabenwechsel pro Block. In der Mischungsbedingung wurde im Gegensatz dazu für jeden Block eine zufällige Abfolge von Attribut- und Zielkonzeptaufgabe erzeugt. Die Wahrscheinlichkeit, dass ein gegebener Durchgang einen Aufgabenwechsel erfordert, liegt damit bei etwa 50%.

Stichprobe

An der Untersuchung nahmen 49 Personen teil. 17 Teilnehmer waren männlichen und 31 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 16 und 42 Jahren ($M = 26.02$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

In den IAT-Blöcken fanden je 24 bezüglich verschiedener Kriterien gematchte Bezeichnungen für Blumen und Insekten sowie positiv und negativ bewertete Substantive Verwendung. Im Vergleich zu Experiment 1 wurden dabei drei Stimulusquadrupel ersetzt. Diese Änderung erschien sinnvoll, weil die Diskriminationsleistung bei den Wörtern "Beteuerungen", "Eigenmächtigkeit" und "Eigensinn" vergleichsweise gering ausfiel. Um die Ähnlichkeit der Stimulismengen aufrecht zu erhalten, wurden die entsprechenden Quadrupel durch die am stärksten homogenen unbenutzten Quadrupel ersetzt. Die zusätzlichen Reize finden sich in Anhang B. Die Zielkonzeptreize, das heißt, die Exemplare der Kategorien Blumen und Insekten, wurden auch in den Bewertungsblöcken verwendet. Zusätzlich wurden 24 rot und 24 blau eingefärbte Konsonantenketten eingesetzt. Jede dieser zufällig gebildeten Konsonantenketten bestand aus sieben Buchstaben. Die Buchstaben q, x, y und z wurden nicht verwendet.

Prozedur

Insgesamt bearbeitete jeder Teilnehmer 24 Aufgabenblöcke, die aus jeweils 24 Durchgängen bestanden. Ein einzelner Versuchsteilnehmer bearbeitete zwölf Aufgabenblöcke und zwölf Bewertungsblöcke, in denen die Zielkonzeptreize direkt bewertet werden mussten. Bis auf die geringere Zahl der Durchgänge folgten die IAT-Blöcke dem Aufbau des in Experiment 1 verwendeten Blumen-Insekten IATs. Die Bewertungsblöcke bestanden zu

gleichen Teilen aus den Zielkonzeptreizen der IAT-Blöcke und aus Konsonantenketten, die entweder blau oder rot gefärbt waren. Die Versuchsteilnehmer wurden instruiert, die vorkommenden Wortreize durch Aussprechen der Antwortalternativen „positiv“ beziehungsweise „negativ“ zu bewerten. Handelte es sich bei einem der dargebotenen Reize um eine Konsonantenkette, sollte deren Farbe benannt werden. Die Reihenfolge der Durchgänge wurde komplett zufällig bestimmt. Um die Nachwirkungen auf der Ebene von Aufgabenblöcken zu erfassen, wurden die verschiedenen Blocktypen im Wechsel miteinander bearbeitet, das heißt, ohne zwischengeschaltete Übungsphasen. Das Experiment begann immer mit einem IAT-Block. Pro Teilnehmer ergeben sich folglich zwölf Bewertungsblöcke, die direkt auf einen IAT-Block folgen. Das Response-Stimulus-Intervall betrug in allen Bedingungen 1000 ms. Zu Beginn eines jeden Durchgangs wurde ein Fixationsreiz dargeboten, der für 800 ms die Darbietung des nächsten Stimulus ankündigte. Die Reaktionen wurden mithilfe eines Headsets am Computer erfasst. Die Antworten der Versuchsteilnehmer wurden über jeweils 2 Sekunden nach der Präsentation des Stimulus aufgezeichnet.

Ergebnisse

Die Daten dreier Versuchsteilnehmer wurden von der Auswertung ausgeschlossen, da diese die Blumen konsistent als negativ und die Insekten konsistent als positiv klassifizierten, und die Einordnung der Phasen des IAT als kompatibel oder inkompatibel für diese Teilnehmer somit nicht eindeutig ist. Werden diese Teilnehmer in die Auswertung aufgenommen, ergibt sich kein abweichendes Befundmuster. Ausgewertet wurden nur die Durchgänge der Testblocks. Einzelne Durchgänge wurden von der Auswertung ausgeschlossen, wenn die Antworten außerhalb der vorgegebenen Antwortmenge („blau“, „rot“, „positiv“ und „negativ“; 0.4%) oder außerhalb des Antwortintervalls von 2 Sekunden lagen (0.7%).

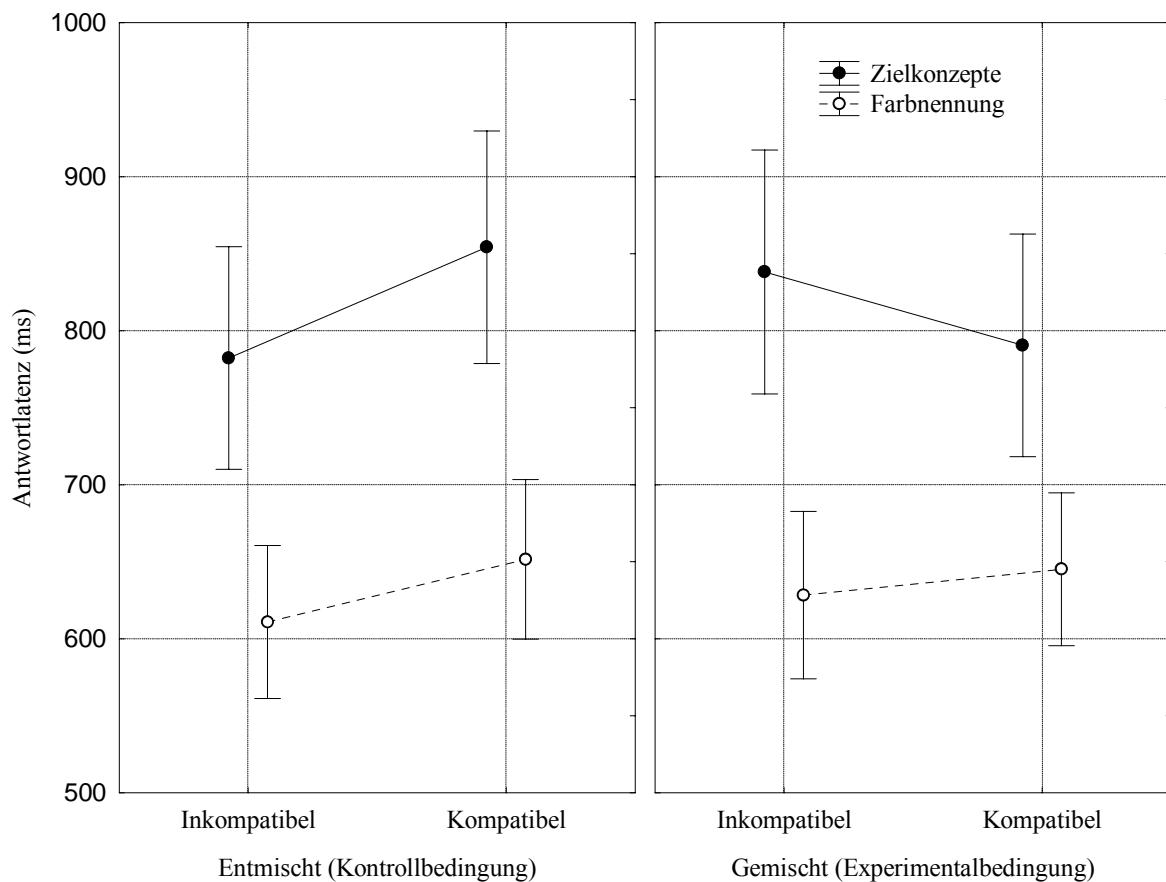


Abbildung 9: Mittlere Antwortlatenzen bei der verbalen Bewertung von Zielkonzeptreizen und bei der Farbnennung von Konsonantenketten nach gemischten und entmischten kompatiblen und inkompatiblen IAT-Blöcken in Experiment 5. Die Antwortlatenzen der IAT-Blöcke können der Abbildung nicht entnommen werden.

Die Antwortlatenzen in den Testblöcken wurden mit einer 2 (Kompatibilität der IAT-Blöcke) x 2 (Mischung vs. Separierung der Aufgabentypen innerhalb der IAT-Blöcke) x 2 (Bewertungsaufgabe vs. Farbnennungsaufgabe) faktoriellen Varianzanalyse ausgewertet. Dabei ergibt sich die erwartete Dreifach-Interaktion der drei Faktoren, $F(1, 41) = 5.27, p < .05$. Die dieser Interaktion zugrunde liegenden Antwortlatenzen sind in Abbildung 9 zu sehen. Effekte der Vorgängerphase finden sich nur nach gemischten IAT-Blöcken und nur für die Bewertungsaufgabe. Geplante Vergleiche zeigen, dass eine Zweifach-Interaktion von Aufgabentyp und Kompatibilität nur nach gemischten IAT-Blöcken auftritt, $F(1, 41) = 4.66, p < .05$, nicht jedoch in der entmischten Kontrollbedingung, $F(1, 41) = 1.15, n.s.$ Ein trivialer Haupteffekt zeigt, dass die Farbnennungsaufgabe schneller bearbeitet wird als die Bewertungsaufgabe, $F(1, 41) = 305.86, p < .01$. Entgegen der Erwartung scheinen beide

Aufgaben nach entmischten inkompatiblen IAT-Blöcken schneller bearbeitet zu werden als nach kompatiblen, dieser Effekt ist jedoch nicht signifikant, $F(1, 41) = 1.80$, n.s.

Diskussion

Die Ergebnisse von Experiment 5 sprechen gegen eine Erklärung von Nachwirkungseffekten durch eine unspezifische Verschiebung des Antwortkriteriums (Brendl et al., 2001), denn der Effekt der Vorgängerphase fällt für die Bewertungsaufgabe signifikant größer aus als für die Farbnennungsaufgabe. Es scheint plausibel, dass die Anwendung der Attribut-aufgabe in den Testblöcken durch die vorhergehende Bearbeitung von inkompatiblen IAT-Blöcken erschwert wurde. Für diese Interpretation spricht auch, dass die Nachwirkungen nicht auf dem Ausstrahlen einer Hemmung spezifischer Verknüpfungen von Reizen und Reaktionen beruhen können, da sie trotz des Wechsels der Antwortmodalität von Tastendrücken zum Aussprechen der Antwort auftreten. Wie auf Basis des Trägheitsmechanismus erwartet, treten spezifische Nachwirkungen nur dann auf, wenn die Aufgaben in den IAT-Blöcken gemischt bearbeitet werden, nicht aber, wenn sie innerhalb eines Aufgabenblocks getrennt werden. Der in Abbildung 9 sichtbare, gleichläufige Effekt entmischter IAT-Blöcke auf Farbnennung und Zielkonzeptbewertung ist unerwartet, erreicht jedoch keine statistische Signifikanz. Ob und inwiefern dies Ergebnis ein replizierbares Phänomen darstellt oder auf ein Versagen der Randomisierungsprozedur zurückgeht, kann zunächst nicht geklärt werden. In der Summe entspricht das Befundmuster den Vorhersagen des Aufgabenwechselmodells.

5.3 Experiment 6

Nach Fazio (Fazio et al., 1986; Fazio & Towles Schwen, 1999) ist die Performanz bei evaluativen Entscheidungen ein Indikator für die Einstellungszugänglichkeit. Die differentiellen Nachwirkungen der inkompatiblen IAT-Blöcke in Experiment 5 könnten also auf eine temporär verringerte Einstellungszugänglichkeit zurückgehen, die durch spezifische Interferenzen der Stimulusbewertung mit der instruierten Aufgabe in dieser Phase zustande kommen. Es kann jedoch nicht ausgeschlossen werden, dass die Ergebnisse von Experiment 5 eine qualitative Änderung der Einstellungsextremität widerspiegeln. Die Verknüpfung einer negativ konnotierten Kategorie und ihrer Exemplare mit einer positiv besetzten Antworttaste könnte, zum Beispiel vermittelt über Prozesse der evaluativen Konditionierung (De Houwer, 2001a; Walther, 2002), einen Effekt auf die Extremität der Einstellung haben.

In Experiment 6 soll geprüft werden, ob die Bearbeitung von kompatiblen und inkompatiblen IAT-Blöcken differentielle Nachwirkungen auf die Extremität der Attributausprägungen hat. Dazu wurde die in Experiment 5 eingesetzte Bewertungsaufgabe durch ein evaluatives Rating ersetzt (siehe unten). Wenn die Nachwirkungseffekte in Experiment 5 auf eine qualitative Änderung der Extremität der Einstellung zurückgehen, so sind analoge Befunde für die qualitative Ausprägung der Ratings und die Ratinglatenzen zu erwarten. Sind die Effekte hingegen eine Folge der Trägheit von Aufgabensets, so sind nur Effekte auf die Ratinglatenzen zu erwarten, jedoch keine Unterschiede in der qualitativen Ausprägung der Ratings.

Methode

Die Methode folgt weitgehend derjenigen von Experiment 5. Alle Versuchsteilnehmer bearbeiteten im Wechsel zwölf IAT-Blöcke und zwölf Testblöcke. Anders als in Experiment 5 sollten die Antworten in den Testblöcken nicht ausgesprochen werden, sondern durch Mausklicks auf einer am Bildschirm dargebotenen sechsstufigen Ratingskala gegeben werden. Für die zu bewertenden Zielkonzeptreize reichte die Skala von „stark negativ“ bis „stark positiv“. Für die farbigen Konsonantenketten reichte die Skala von „gesättigtes rot“ bis „gesättigtes blau“. Beide Skalen wiesen keinen neutralen Punkt auf. Die Teilnehmer wurden angehalten, möglichst schnell denjenigen Punkt auf der Ratingskala anzuklicken, der ihrer Beurteilung am ehesten entspricht. Die Ratings mussten in einem Antwortintervall von 4 Sekunden abgegeben werden, um eine möglichst spontane Beurteilung zu erfassen.

Stichprobe

An der Untersuchung nahmen 80 Personen teil. 41 Teilnehmer waren männlichen und 39 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 18 und 65 Jahren ($M = 34.91$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Prozedur

Die Prozedur entspricht - abgesehen von der geänderten Antwortmodalität - derjenigen von Experiment 5.

Material

Es wurde das gleiche Material verwendet wie in Experiment 5. Zusätzlich wurde jedoch der Sättigungsgrad der blauen und roten Konsonantenketten in jeweils acht Stufen manipuliert, um eine Vergleichbarkeit der Farbratings mit den evaluativen Ratings zu gewährleisten.

Ergebnisse

Die Daten von zwei Versuchsteilnehmern wurden von der Auswertung ausgeschlossen, da diese die Blumen konsistent als negativ und die Insekten konsistent als positiv einstufen. Ausgewertet wurden nur die Durchgänge der Testblocks. Einzelne Durchgänge wurden von der Auswertung ausgeschlossen, wenn die Antworten jenseits des vorgegebenen Antwortfensters von 4 Sekunden lagen (0.9%).

Die Antwortlatenzen in den Testblöcken wurden wie in Experiment 5 mit einer 2 (Kompatibilität der IAT-Blöcke) x 2 (Mischung vs. Separierung der Aufgabentypen innerhalb der IAT-Blöcke) x 2 (Bewertungsaufgabe vs. Farbnennungsaufgabe) faktoriellen Varianzanalyse ausgewertet. Die erwartete Dreifach-Interaktion aller Faktoren verfehlt das Signifikanzniveau, $F(1, 74) = 2.97, p < .09$. Die entsprechenden Antwortlatenzen können Abbildung 10 entnommen werden. Geplante Vergleiche zeigen jedoch, dass die Zweifach-Interaktion zwischen Aufgabentyp und Kompatibilität der Vorgängerphase wie erwartet nach gemischten IAT-Blöcken auftritt, $F(1, 74) = 5.81, p < .01$, nicht jedoch nach entmischten Kontrollblöcken, $F(1, 74) = 0.00, n.s.$ Ein Effekt der Vorgängerphase auf die Ratinglatenzen nach gemischten Blöcken tritt für sich genommen weder für die Farbratings, $F(1, 74) = .07, n.s.$, noch für die Bewertungen auf, $F(1, 74) = 3.05, p < .08$. Deskriptiv geht die Zweifach-Interaktion von Aufgabentyp und Art der Vorgängerphase nach gemischten IAT-Blöcken jedoch offenkundig auf die Bewertungslatenzen zurück. Nach den entmischten Kontrollblöcken findet sich keinerlei Effekt der Kompatibilität auf die einzelnen Aufgaben, jeweils $F(1, 74) < 1, n.s.$

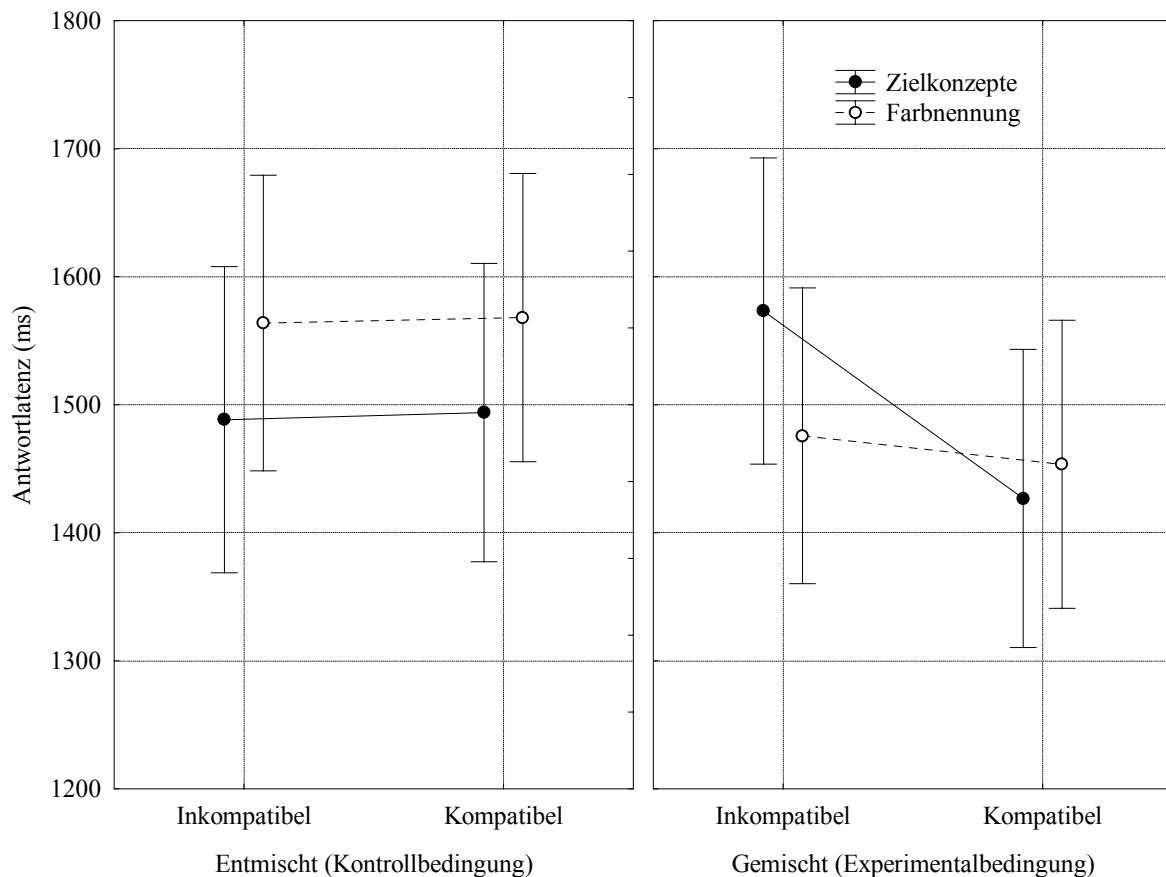


Abbildung 10: Mittlere Antwortlatenzen beim Rating der Zielkonzeptreize und der farbigen Konsonantenketten in Experiment 6 nach gemischten und entmischten kompatiblen und inkompatiblen IAT-Blöcken. Die Antwortlatenzen der IAT-Blöcke können der Abbildung nicht entnommen werden. Ein Effekt der Kompatibilität der Vorgängerphase findet sich nur auf die Ratings der Zielkonzeptexemplare nach gemischten IAT-Blöcken.

Die Ratings der Blumen und Insekten wurden mit einer 2 (Kompatibilität der IAT-Blöcke) x 2 (Mischung vs. Separierung der Aufgabentypen innerhalb der IAT-Blöcke) x 2 (Blumen vs. Insekten) faktoriellen Varianzanalyse ausgewertet. Der einzige signifikante Effekt in dieser Analyse ist ein trivialer Haupteffekt, der darauf zurückgeht, dass die Blumen positiver bewertet werden als die Insekten, $F(1, 74) = 1033.43, p < .01$. Alle anderen Effekte sind nicht statistisch bedeutsam, $F(1, 74) < 1.85, n.s.$ Obwohl die Ratinglatenzen wie erwartet von der Art Vorgängerphase abhängen, finden sich also keine Nachwirkungen der IAT-Blöcke auf die mittleren Ratings.

Diskussion

Die Befunde bei den Ratinglatenzen in Experiment 6 entsprechen weitgehend dem Muster der Aussprechlatenzen in Experiment 5. Die Versuchsteilnehmer bewerteten die Valenz der Zielkonzeptexemplare nach inkompatiblen IAT-Blöcken langsamer als nach kompatiblen, wenn Attribut- und Zielkonzeptaufgabe innerhalb der Blöcke gemischt zu bearbeiten waren. Nach den entmischten Kontrollblöcken traten keinerlei Effekte der Vorgängerphase auf. Dies spricht dafür, dass die Nachwirkungseffekte eine Konsequenz der Aufgabenwechsel darstellen. Eine globale Änderung des Antwortkriteriums kann diesen Effekt ebenso wenig erklären wie das Ausbleiben von Effekten auf die Latenzen der Farbratings. Letztere waren von der Kompatibilität der Vorgängerphase komplett unabhängig. Gleichzeitig finden sich keine Hinweise auf Effekte in der Ausprägung der evaluativen Ratings. Die Nachwirkungseffekte scheinen folglich lediglich die temporäre Zugänglichkeit, nicht jedoch die qualitative Ausprägung der Ratings zu betreffen.

5.4 Zusammenfassung der Befunde zu globalen Effekten

Die Ergebnisse der Experimente 4 bis 6 bieten zusammengenommen ein differenziertes Bild der Effekte beim Übergang zwischen kompatibler und inkompatibler Tastenzuordnung im IAT. Im Folgenden sollen die theoretischen und praktischen Konsequenzen dieser Befunde zusammenfassend beleuchtet werden.

Experiment 4 zeigt, dass die Bearbeitung von Reizen mit inkompatibler Tastenzuordnung unmittelbar erschwert ist, wie es der Trägheitsmechanismus vorhersagt. Dies geht vor allem auf die Interferenz aufgabenirrelevanter Reizeigenschaften zurück, wenn diese in einem kürzlich aktivierten Aufgabenset relevant waren. Die Hemmung von interferierenden Aufgabensets ist ebenso wie die Aktivierung des jeweils relevanten Aufgabensets mit Kosten verbunden und wirkt über eine Reihe von Durchgängen hinweg nach. Damit kann sparsam erklärt werden, warum Kompatibilitätseffekte im IAT vor allem nach Aufgabenwechseln ausgeprägt sind: Die Nachwirkungen der vorhergehenden Aufgabenbearbeitung sind in diesem Fall am stärksten. Solche Effekte sollten auftreten, sobald die irrelevante Eigenschaft in einem Durchgang verarbeitet werden muss, bei stetigem Wechsel zwischen Attribut- und Zielkonzeptaufgabe also unmittelbar. Ein zeitabhängiges, passives Abklingen von Aktivierung und Hemmung der Aufgabensets, wie Allport et al. (1994) es finden, kann ferner den Effekt des Response-Stimulus Intervalls auf die Aufgabenwechselkosten und den IAT-Effekt in Experiment 1 erklären.

Das Ausbleiben einer unmittelbaren Erleichterung der Aufgabe beim Wechsel von inkompatiblen auf kompatible Durchgänge in Experiment 4 lässt sich durch eine längerfristige Trägheit der Aufgabensets erklären: Die Bearbeitung der inkompatiblen Phase führt zu einer Hemmung des Attributaufgabensets, die über viele Durchgänge nachwirkt. Dies erklärt auch die Nachwirkungseffekte in den Experimenten 5 und 6. Gleichzeitig demonstrieren diese Experimente, dass die Durchführung von Aufgabenwechseln eine notwendige und hinreichende Bedingung für das Auftreten der Nachwirkungen darstellt, da keine Nachwirkungen in den entmischten Kontrollbedingungen auftraten. Spezifische Nachwirkungseffekte traten also *dann und nur dann* auf, wenn die Vorgängerphase Aufgabenwechsel umfasste. Wie auf Basis des Trägheitsmechanismus vorhergesagt, sind die Nachwirkungen fernerhin spezifisch für die Attributaufgabe. Dies zeigt sich daran, dass die Latenzen für die Farbnennung beziehungsweise Farbbeurteilung von der Vorgängerphase komplett unabhängig waren. Der zeitliche Verlauf der Nachwirkungen ist jedoch gegenwärtig unklar. Einerseits finden sich Hinweise darauf, dass eine Zeitspanne von einer Sekunde zwischen den Durchgängen zu einer nachweisbaren Verringerung der Nachwirkungen führt, wie der Effekt des RSI in Experiment 1 belegt. Andererseits zeigen die längerfristigen Effekte in den Experimenten 4 bis 6, dass über mehrere Durchgänge hinweg nachweisbare Nachwirkungen erhalten bleiben. Dies könnte auf einen kurvenförmigen Verlauf der Nachwirkungen hindeuten, die zunächst sehr rasch abnehmen, aber trotzdem auf einem niedrigen Niveau längerfristig anhalten.

In der Diskussion zu Experiment 4 wurde eine Analogie zwischen globalen Effekten des Kompatibilitätsübergangs im IAT und Konsistenzanteileffekten im affektiven Priming (Klauer et al., 1997; Musch, 2000) aufgezeigt. Die Nachwirkungen eines hohen beziehungsweise niedrigen Konsistenzanteils auf die Gewichtung der irrelevanten Primeinformation über viele Durchgänge hinweg, selbst dann, wenn der ursprüngliche Kontext sich ändert, scheinen auf Ebene der Effekte weitgehend vergleichbar. In diesem Zusammenhang wurde diskutiert, ob eine im Verlauf der kompatiblen Phase eines IAT zunehmende Gewichtung der irrelevanten Attributinformation den Effekten zugrunde liegen könnte. Zusammengefasst weisen die Daten jedoch darauf hin, dass auch ein graduelles Abklingen der Hemmung des Attributaufgabensets nach der Bearbeitung der inkompatiblen Phase zu diesen Effekten beiträgt. Dafür spricht unter anderem, dass die Bearbeitung von kompatiblen Aufgabenblöcken nach 168 Durchgängen mit inkompatiblen Reizen in Experiment 4 deutlich langsamer ist als die Bearbeitung der kompatiblen Aufgabenblöcke zu Beginn des

Experiments. Nachwirkungen der inkompatiblen Phase sind also mindestens zusätzlich am Zustandekommen der Effekte beteiligt.

Interessant ist in diesem Zusammenhang auch die Frage, welche Aspekte der Aufgabenrepräsentation von den Nachwirkungseffekten betroffen sind. Wie schon angesprochen bezeichnet der Begriff Aufgabenset kein einheitliches psychologisches Konstrukt, sondern vielmehr die Menge kognitiver Operationen und Zustände, die nötig ist, eine Aufgabe effektiv zu bearbeiten. Dazu gehören die Fokussierung auf ein relevantes Stimulusattribut, die möglichen Kategorisierungen dieses Stimulusattributs, die verfügbaren Antwortmodi und -optionen, sowie vorgegebene Verbindungen zwischen Stimulus- und Antwortklassen (Monsell et al., 2000). Auf Basis der vorliegenden Ergebnisse scheinen vor allem solche Prozesse betroffen, die mit der Auswertung der Attributinformation zu tun haben, nicht jedoch globale Kontrollstrategien oder spezifische Verknüpfungen von Reizen und Reaktionen. Ebenso wenig scheint die Auslösung spezifischer motorischer Reaktionen global erschwert zu sein. Ein Locus der Nachwirkungseffekte scheint also auf der Ebene der inhaltlichen Verarbeitung von Attributinformation zu liegen. Die Konsequenzen dieser Ergebnisse für die praktische Anwendung des IAT werden in der Gesamtdiskussion dargelegt.

6 Experimente zur Rolle exekutiver Kontrollprozesse

Exekutive Kontrollprozesse spielen eine zentrale Rolle bei der Handlungskontrolle und der Koordination von automatisch, stimulusgetriebener und intentionaler Verarbeitung (Gopher, 1996; Koch, 2003; Meiran, 1996; Rogers & Monsell, 1995). Man kann wohl davon ausgehen, dass schon die Bearbeitung einfacher experimenteller Aufgaben (ebenso wie eine große Zahl alltäglicher Handlungen) weit davon entfernt ist, exklusiv durch automatische Prozesse und Antwortroutinen gesteuert zu sein. Im Gegenteil ist anzunehmen, dass solchen Handlungen ein keineswegs triviales *Zusammenspiel* exekutiver Kontrollfunktionen und automatischer Prozesse zugrunde liegt.

Auch bei der Bearbeitung eines IAT spielen offenkundig beide Arten kognitiver Prozesse, intentional gesteuerte und automatische, eine Rolle, wie die weiter oben angeführte Argumentation, dass dem IAT kein rein stimulusgetriebener Mechanismus zugrunde liegen kann, offenbart. Dies setzt nicht zwingend voraus, dass entsprechende Prozesse in einem Modell explizit gemacht werden müssen. Das Random-Walk Modell von Brendl et al. (2001) setzt zum Beispiel einfach voraus, dass spezifische Information mit einer spezifischen Gewichtung in Richtung auf eine Antwortalternative akkumuliert wird. Wie diese Gewichtung zustande kommt, ist nicht Gegenstand des Modells.

Das hier vorgeschlagene Aufgabenwechselmodell expliziert die Rolle exekutiver Kontrollprozesse. Es wird angenommen, dass die instruierte Aufgabenstruktur, insbesondere die Zuordnung von Reizklassen zu Antwortoptionen, in Form so genannter Aufgabensets repräsentiert ist. Die Rolle exekutiver Kontrollprozesse besteht im IAT darin, das jeweils angemessene Aufgabenset zu identifizieren, zu aktivieren und anzuwenden, aber auch darin, irrelevante und eventuell störende Aufgabensets zu unterdrücken. Dies betrifft vor allem den Wechsel zwischen Zielkonzept- und der Attributaufgabe innerhalb der gemischten Aufgabenblöcke eines IAT. Selbst wenn Aufgabenwechselkosten kein direkter Indikator für die Dauer exekutiver Kontrollprozesse sind (Allport et al., 1994; Wylie & Allport, 2000), scheint zwingend, dass die Koordination der konfligierenden Aufgabenspezifikationen in der inkompatiblen Phase eines IAT durch exekutive Kontrollfunktionen gesteuert wird.

Wie schon geschildert ergab eine Untersuchung von Meiran (1996), dass die Präsentation von Aufgabenhinweisen zu einer bedeutsamen Reduktion von Aufgaben-

wechselkosten führte. Meiran interpretiert dies als Hinweis darauf, dass exekutive Kontrollprozesse die Bearbeitung einer nachfolgenden Aufgabe zumindest teilweise erleichtern können (vgl. auch Koch, 2003; Meiran et al., 2000). In der Tat kann dieser Effekt nicht durch den von Allport und Kollegen (Allport et al., 1994; Wylie & Allport, 2000) angenommenen Nachwirkungsmechanismus erklärt werden. Wie Hsieh (2002) argumentiert, ist dieser Befund jedoch nicht mit dem Trägheitsmechanismus inkompatibel. Vielmehr scheinen sowohl passive Trägheitseffekte als auch aktive exekutive Kontrollprozesse am Zustandekommen von Aufgabenwechselkosten beteiligt zu sein.

In den Experimenten 7 und 8 soll daher die Rolle exekutiver Kontrollprozesse im IAT untersucht werden. Dabei steht die Frage im Mittelpunkt, ob die im IAT auftretenden Aufgabenwechselkosten durch die Präsentation von Aufgabenhinweisen reduziert werden können. Auf Basis des Aufgabenwechselmodells ist anzunehmen, dass eine Reduktion der Aufgabenwechselkosten zu einer Reduktion des IAT-Effekts selbst führt. Sofern die im nächsten Durchgang anstehende Aufgabe schon im Vorbereitungsintervall bekannt ist, sollten Konflikte zwischen Aufgabensets partiell schon vor der Darbietung des imperativen Stimulus gelöst werden können. In Experiment 8 sollen ferner unterschiedliche Komponenten exekutiver Kontrollprozesse separiert und ihr Effekt auf den IAT untersucht werden.

6.1 Experiment 7

In Experiment 7 sollte der Effekt von exogenen Aufgabenhinweisen (z. B. Koch, 2003; Meiran, 1996) auf einen einfachen Blumen-Insekten-IAT (Greenwald et al., 1998) geprüft werden. Die Aufgabenhinweise wurden in Form einer symbolischen Repräsentation der Aufgabenstruktur vor dem oder gleichzeitig mit dem Stimulusonset dargeboten. Die Aufgabenhinweise bestanden aus vier Symbolen, die die instruierten Antwortkategorien repräsentieren sollten, namentlich ein Daumen nach oben für positive Reize, ein Daumen nach unten für negative Reize, und das Piktogramm einer prototypischen Blume und eines prototypischen Insekts für die Zielkonzepte. Diese Symbole wurden jeweils links und rechts neben dem Stimulus präsentiert, und zwar so, dass die Position des Symbols mit der Position derjenigen Antworttaste auf der Computertastatur übereinstimmte, der die entsprechende Kategorie zugeordnet war.

Die Struktur der Aufgabenhinweise wurde in vier between-subjects Bedingungen manipuliert. In allen Bedingungen lag das RSI bei 800 ms. Die Tastenzuordnung wurde - anders als in den Vorgängerexperimenten - nicht in der oberen Bildschirmhälfte präsentiert. In

der ersten Bedingung wurden die Aufgabenhinweise zum Zeitpunkt des Stimulusonset präsentiert. Es erschienen jeweils nur die Symbole, die den beiden Antwortkategorien der im aktuellen Durchgang geforderten Aufgabe entsprachen. Für die Attributaufgabe wurden die beiden Symbole für positive und negative Reize links und rechts neben dem Stimulus präsentiert. Für die Zielkonzeptaufgabe wurden die Symbole für Blume und Insekt links und rechts neben dem Stimulus präsentiert. In der zweiten Bedingung wurden die gleichen Symbole präsentiert, jedoch 600 ms vor dem Stimulusonset. Ein Aufgabenvorbereitungseffekt sollte also in Form eines Kontrastes zwischen diesen beiden Bedingungen auftreten. In einer dritten Bedingung wurden gleichzeitig mit dem Stimulusonset alle vier Symbole präsentiert, jeweils zwei auf der linken und zwei auf der rechten Seite. In einer weiteren Kontrollbedingung, schließlich wurden keine Aufgabenhinweise präsentiert.

Das beschriebene Design stellt sicher, dass Aufgabenvorbereitungseffekte nicht mit einer Manipulation des RSI zusammenfallen, da in allen Versuchsbedingungen das gleiche RSI realisiert wurde. Die Information der Aufgabenhinweise war ferner in keiner Bedingung mit den zu gebenden Antworten konfundiert, da die korrekte Antworttaste nicht durch die Aufgabenhinweise vorhersagbar war. Um in allen Versuchsbedingungen vergleichbare Voraussetzungen zu schaffen, wurden die Versuchsteilnehmer nicht dazu angehalten, die Aufgabenhinweise zu nutzen. Dadurch sollte ausgeschlossen werden, dass durch die Instruktion schon präexperimentell verschiedene Bearbeitungsstrategien induziert werden.

Auf Basis des Aufgabenvorbereitungsmodells von Meiran (1996) wäre zu erwarten, dass eine Vorbereitung auf die anstehende Aufgabe innerhalb des RSI zu einer Verringerung der Aufgabenwechselkosten führt. Ist dies der Fall, sollten die IAT-Effekte in der Bedingung mit Aufgabenhinweisen vor dem Stimulusonset im Vergleich zu der Bedingung mit gleichzeitiger Darbietung und im Vergleich zu den Kontrollbedingungen verringert sein.

Methode

Der Versuchsaufbau von Experiment 7 entspricht demjenigen von Experiment 1. Die Versuchsteilnehmer bearbeiteten einen einfachen Blumen-Insekten Einstellungs-IAT unter vier verschiedenen Bedingungen, die sich bezüglich der links und rechts neben dem Stimulus präsentierten Aufgabenhinweise unterschieden. Die Reihenfolge von kompatibler und inkompatibler Phase wurde ausbalanciert. Ferner wurde die Tastenbelegung so ausbalanciert, dass die Taste für positiv für eine Hälfte der Probanden mit der dominanten Hand zu drücken war und für die andere Hälfte mit der nichtdominanten Hand.

Stichprobe

An der Untersuchung nahmen 61 Personen teil. 18 Teilnehmer waren männlichen und 43 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 15 und 45 Jahren ($M = 25.11$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Als Material dienten die in den Experimenten 5 und 6 eingesetzten Wörter für Blumen und Insekten, sowie die positiven und negativen Substantive. Als Aufgabenhinweise wurden vier einfache Symbole verwendet, die einen nach oben weisenden und einen nach unten weisenden Daumen, sowie eine prototypische Blume und ein prototypisches Insekt darstellten. Die Symbole wurden in schwarzer Färbung auf einem hellgrauen Hintergrund dargeboten.

Prozedur

Die Aufgabenblöcke bestanden aus 48 sequentiell präsentierten Wörtern, die zufällig und ohne Zurücklegen aus den vier Stimuluslisten gezogen und in zufälliger Reihenfolge dargeboten wurden. Zu Beginn des Experimentes wurden zunächst zwei Übungsphasen bearbeitet, die jeweils zwei Blöcke umfassten. In der ersten Phase wurde die Kategorisierung der Blumen und Insekten eingeübt. In der zweiten Phase wurde die Attributaufgabe eingeübt, in diesem Fall eine evaluative Entscheidungsaufgabe. Danach folgten vier Blöcke, in denen beide Aufgaben gemischt bearbeitet wurden, je nach Balancierungsbedingung entweder mit kompatibler oder mit inkompatibler Tastenzuordnung. Dann wurde die Tastenzuordnung der Zielkonzeptaufgabe umgekehrt und in zwei Übungsblöcken und im Anschluss daran in vier gemischten Blöcken bearbeitet.

In den zwei Experimentalbedingungen wurden Aufgabenhinweise präsentiert, entweder gleichzeitig mit dem Stimulus oder 600 ms vor dem Stimulusonset. Zusätzlich wurden zwei Kontrollbedingungen realisiert, in denen entweder keine oder alle vier Symbole präsentiert wurden. Das RSI betrug für alle Durchgänge 800 ms. In allen Bedingungen wurden die Durchgänge durch einen Fixationsreiz eingeleitet, der 600 ms vor dem Stimulus erschien. Dadurch ist die zeitliche Vorhersagbarkeit des Stimulusonset in allen Versuchsbedingungen identisch.

Ergebnisse

Um einem Effekt antizipatorischer Reaktionen und kurzfristiger Aufmerksamkeitslücken entgegenzuwirken, wurden Antwortlatenzen unter 300 ms zu 300 ms (0.9 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (1 % der Durchgänge) rekodiert. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (7.2 % der Durchgänge) nicht in die Analyse aufgenommen.

Die mittleren Antwortlatenzen wurden mit einer 4 (Aufgabenhinweisbedingung) x 2 (Reihenfolge der kompatiblen und inkompatiblen Phase) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf den letzten beiden Faktoren ausgewertet. Über die Bedingungen hinweg ergibt sich ein signifikanter IAT-Effekt von 125 ms, $F(1, 53) = 62.21, p < .01$. Auch der Unterschied zwischen Aufgabenwechsel- und Aufgabenwiederholungsdurchgängen ist über das Experiment hinweg signifikant, $F(1, 53) = 111.24, p < .01$. Wie in den Vorgängerexperimenten ergibt sich eine Zweifach-Interaktion zwischen Kompatibilität und Aufgabenwechsel, $F(1, 53) = 40.27, p < .01$. Die Kompatibilitätseffekte fallen wie erwartet für Aufgabenwechseldurchgänge signifikant größer aus. Die Balancierung der Kompatibilitätsreihenfolge hat keinerlei Effekt auf die Ergebnisse, alle $F < 1.92, p > .13$.

Eine Zweifach-Interaktion zwischen dem Aufgabenhinweisfaktor und der Größe des IAT-Effekts verfehlt das Signifikanzniveau knapp, $F(3, 53) = 2.54, p = .06$. Werden zusätzliche Faktoren in die Analyse aufgenommen, erreicht diese Interaktion jedoch das Signifikanzniveau (vgl. Mierke & Klauer, 2001). Dies geht vermutlich darauf zurück, dass der Effekt in diesem Fall gegen eine kleinere Fehlerquadratsumme getestet wird. Die vorgezogene Darbietung der Aufgabenhinweise hat im Vergleich zur gleichzeitigen Darbietung keinen Effekt auf die Aufgabenwechselkosten, $F(1, 57) = 1.23, n.s.$, wie ein entsprechender Kontrast zeigt. Auch die IAT-Effekte fallen in diesen Bedingungen gleich aus, $F(1, 53) = 0.02, n.s.$ Ein Vergleich der beiden Bedingungen mit Aufgabenhinweisen mit den beiden Kontrollgruppen zeigt, dass die IAT-Effekte in den Bedingungen mit Aufgabenhinweisen geringer ausfallen als ohne, $F(1, 53) = 4.36, p < .05$. Dieser Effekt ist jedoch offenbar nicht durch verringerte Aufgabenwechselkosten vermittelt, wie ein Vergleich der Interaktion zwischen Kompatibilität und dem Aufgabenwechseln zwischen den Experimental- und Kontrollgruppen zeigt, $F(1, 53) = 0.26, p = .61$. Abbildung 11 zeigt die mittleren

Antwortlatenzen für die kompatible und inkompatible Bedingung des IAT in Abhängigkeit von der Aufgabenhinweisbedingung.

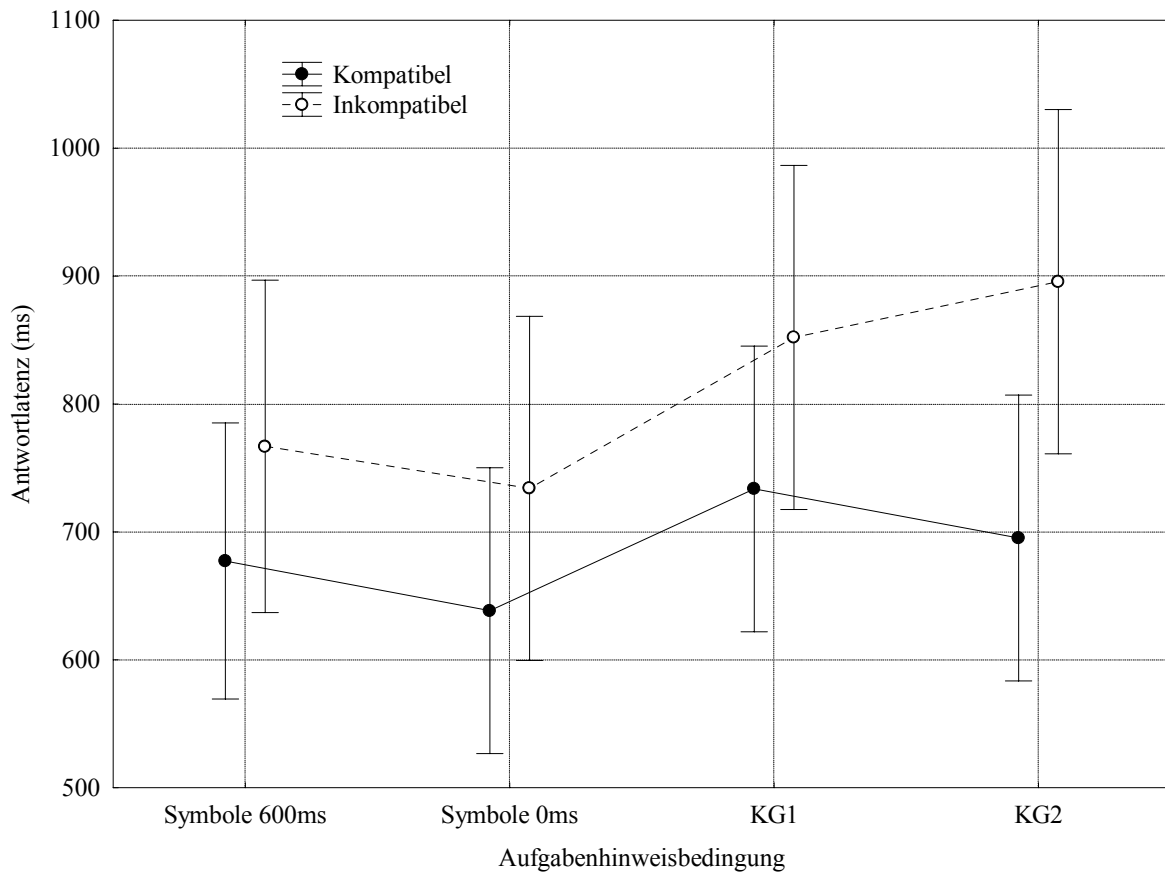


Abbildung 11: Mittlere Antwortlatenzen in der kompatiblen und inkompatiblen Bedingung von Experiment 7 als Funktion der Aufgabenhinweisbedingung. Die Bedingung mit Aufgabenhinweisen vor dem Stimulusonset (Symbole 600 ms) unterscheidet sich nicht von der Bedingung mit simultaner Präsentation von Stimulus und Aufgabenhinweis (Symbole 0 ms). Die größten IAT-Effekte treten auf, wenn alle Symbole (KG2) präsentiert werden, geringere Effekte finden sich, wenn keine Symbole präsentiert werden (KG1).

Diskussion

Experiment 7 repliziert den zentralen Befund asymmetrischer Aufgabenwechselkosten: In der inkompatiblen Phase finden sich über alle Bedingungen hinweg deutlich höhere Aufgabenwechselkosten als in der kompatiblen. Anders gesagt fallen die IAT-Effekte für Aufgabenwechseldurchgänge deutlich größer aus als für Aufgabenwiederholungsdurchgänge. Anders als in der Untersuchung von Meiran (1996) finden sich überraschenderweise keinerlei Effekte der Aufgabenhinweise auf die Aufgabenwechselkosten. Die Tatsache, dass die

Befunde für vorgezogene und simultane Aufgabenhinweise auch deskriptiv nahezu identisch ausfallen, spricht dagegen, dass das Ausbleiben eines Effekts auf eine zu geringe Teststärke zurückgeht. Dennoch zeigt sich ein Effekt der Aufgabenhinweise auf die IAT-Effekte, jedoch unabhängig davon, ob die Hinweise vor dem oder gleichzeitig mit dem Stimulus dargeboten werden. Während ein Effekt der Aufgabenhinweise auf den IAT also mehr oder weniger deutlich auftritt, kann man nicht folgern, dass die Verringerung der IAT-Effekte auf eine Reduktion der spezifischen Aufgabenwechselkosten durch die Aufgabenhinweise zurückgeht.

Eine Reihe von Unterschieden zwischen den Experimenten von Meiran (1996) und Experiment 7 könnten für das überraschende Fehlen eines Vorbereitungseffekts auf die Aufgabenwechselkosten verantwortlich sein. Ein wichtiger Unterschied zwischen den Experimenten scheint zu sein, dass die auszuführende Aufgabe im IAT schon durch den Stimulus selbst spezifiziert wird. Dies liegt daran, dass für die Attribut- und die Zielkonzeptaufgabe notwendigerweise *verschiedene* Stimulussets dargeboten werden. Die durch die Aufgabenhinweise gelieferte Information ist folglich redundant mit einer durch den Stimulus gelieferten Information. In den Experimenten von Meiran (1996) wurden die wechselnden Aufgaben im Unterschied dazu in Bezug auf die *gleiche* Reizmenge bearbeitet. Diese Redundanz von Aufgabenhinweisen und Stimulus könnte dazu geführt haben, dass die Aufgabenhinweise kaum zur Vorbereitung genutzt wurden. Zusammen mit der Tatsache, dass die Versuchsteilnehmer nicht instruiert waren, die Aufgabenhinweise zu nutzen, könnte dies erklären, warum ein zeitliches Vorziehen der Hinweise keine bedeutsame Reduktion der Aufgabenwechselkosten zur Folge hatte.

Aus theoretischer Sicht lassen sich verschiedene Komponenten exekutiver Kontrolle bei Aufgabenwechselprozessen unterscheiden, die jeweils mit spezifischen experimentellen Manipulationen separierbar sein sollen (Koch, 2003; Rogers & Monsell, 1995). Dabei handelt es sich um exogene Kontrollprozesse, die durch die Präsenz exogener Hinweisreize ausgelöst werden, um endogene Kontrollprozesse, die durch einen im Individuum liegenden Impuls ausgelöst werden (Rogers & Monsell, 1995), und um passive, interferenzbedingte Komponenten, die auf Aktivierungs- und Hemmungsprozesse bei der vorhergehenden Aufgabenbearbeitung zurückgehen (Allport et al., 1994; Wylie & Allport, 2000). Letztgenannte Komponente könnte man in Analogie zu den beiden erstgenannten als autogen bezeichnen.

Die in Experiment 7 verwendeten symbolischen Aufgabenhinweise stellen gemäß dieser Klassifikation einen exogenen Hinweisreiz dar. Es wird angenommen, dass solche Hinweis-

reize den Abruf der aufgabenspezifischen Antwortzuordnung durch einen Primingprozess erleichtern (Koch, 2003). Gleichzeitig verbleiben in nahezu allen Experimenten in diesem Bereich residuale Wechselkosten, die durch eine Verlängerung der Vorbereitungsperiode nicht reduzierbar sind (z. B. Koch, 2003; Rogers & Monsell, 1995). Wahrscheinlich gehen diese residualen Wechselkosten auf autogene Interferenz zurück, die erst mit der Darbietung des Stimulus aufgelöst werden kann (Koch, 2003). Das überraschende Fehlen eines Vorbereitungseffekts auf die Aufgabenwechselkosten im vorliegenden Experiment könnte in diesem Sinne ein weiterer Hinweis darauf sein, dass Interferenzen, die auf die kürzliche Verarbeitung eines jetzt irrelevanten Attributs zurückgehen, die Hauptkomponente der Aufgabenwechselkosten im IAT darstellen, während andere Komponenten des Aufgabenwechselprozesses, etwa die Identifikation und Auswahl der zu bearbeitenden Aufgabe und die Änderung des Verarbeitungsziels eine untergeordnete Rolle spielen.

6.2 Experiment 8

In Experiment 8 soll der Anteil exogener, endogener und autogener Komponenten von Aufgabenwechselkosten an der Performanz im IAT experimentell separiert werden. Diese Komponenten exekutiver Kontrollprozesse werden zu diesem Zweck in einem dreifaktoriellen Design unabhängig voneinander manipuliert.

Eine über alle Bedingungen hinweg konstante und vorhersagbare Abfolge der Aufgaben soll dabei eine endogen ausgelöste Vorbereitung ermöglichen. Die Versuchsteilnehmer wurden über das Vorhandensein und die Struktur dieser Abfolge informiert und gebeten diese zu nutzen. Eine über das Experiment hinweg konstante Korrelation zwischen der Position eines Reizes und der auszuführenden Aufgabe sollte exogene Kontrollprozesse ermöglichen. Die Position eines Reizes und damit die auszuführende Klassifikationsaufgabe wurden jeweils im Vorbereitungsintervall durch einen Fixationsreiz angekündigt. Die Versuchsteilnehmer wurden in der Instruktion auch über diese Regelmäßigkeit informiert und entsprechend gebeten diese zu nutzen. Durch diese beiden Manipulationen sollte geklärt werden, ob das Ausbleiben einer Interaktion zwischen Aufgabenhinweisen und Aufgabenwechselkosten im IAT ein replizierbarer Effekt ist, oder auf eine zu geringe Effektstärke zurückgeht. Dabei ist auch von Bedeutung, ob erneut eine Reduktion der IAT-Effekte durch Aufgabenhinweise auftritt oder nicht. Dieses überraschende Befundmuster scheint dafür zu sprechen, dass die Aufgabenhinweise zu einer Reduktion allgemeiner Aufgabenmischungskosten (Los, 1996, 1999) in der inkompatiblen Phase eines IAT beitragen, die sowohl für Aufgabenwechsel- als

auch Aufgabenwiederholungsdurchgänge auftreten. Der Effekt muss also auf eine Komponente des Aufgabenwechselprozesses zurückgehen, die bei beiden Arten von Durchgängen auftritt.

Autogene Effekte wurden durch die Zahl der Aufgabenwechsel innerhalb eines Aufgabenblocks manipuliert. Eine Verringerung der Zahl der Aufgabenwechsel führt dazu, dass die Zahl der Aufgabenwiederholungsdurchgänge sowie die mittlere Länge der Wiederholungssequenzen ansteigen. Der Trägheitsmechanismus lässt erwarten, dass die Aufgabenwechselkosten für längere Wiederholungssequenzen größer ausfallen als für kurze, da eine über mehrere Durchgänge eingeübte Aufgabe stärkere Nachwirkungseffekte produzieren sollte. Einerseits sollten also die Aufgabenwechselkosten in der Bedingung mit langen Aufgabenwiederholungssequenzen geringer ausfallen und zu größeren IAT-Effekten bei den Wechseldurchgängen führen, andererseits sollten die IAT-Effekte insgesamt eher verringert sein, da überwiegend Aufgabenwiederholungsdurchgänge bearbeitet werden.

Methode

In Experiment 8 bearbeiteten die Versuchsteilnehmer den gleichen Blumen-Insekten-IAT, der schon im Vorgängerexperiment verwendet wurde. Als Material dienten die in den Experimenten 5, 6 und 7 eingesetzten Wörter für Blumen und Insekten, sowie die dort verwendeten positiven und negativen Substantive. Exogene, endogene und autogene Komponenten exekutiver Kontrollprozesse wurden in drei zweistufigen Faktoren komplett gekreuzt between-subjects manipuliert. Eine genaue Beschreibung der jeweiligen Manipulationen findet sich in dem Abschnitt zur Prozedur.

Stichprobe

An der Untersuchung nahmen 128 Personen teil. 50 Teilnehmer waren männlichen und 78 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 18 und 55 Jahren ($M = 24,47$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Prozedur

Der Ablauf des Experiments und der Aufbau der Aufgabenblöcke entsprechen, bis auf die im Folgenden dargestellten Abweichungen, demjenigen der Kontrollbedingung ohne

Symbole in Experiment 7. Das RSI betrug in allen Versuchsbedingungen 1000 ms. In den Experimentalgruppen mit exogenen Hinweisreizen wurde eine Kontingenz zwischen der Position des Reizes und der zu bearbeitenden Aufgabe realisiert, während in den zugehörigen Kontrollgruppen keine solche Kontingenz vorlag. Für die Experimentalgruppen wurde das Präsentationsfenster in einen oberen und einen unteren Teil gegliedert. Der eigentliche exogene Hinweisreiz war ein Fixationsreiz (*), der 800 ms vor der Präsentation des imperativen Reizes dessen Position (oben vs. unten) im Präsentationsfenster ankündigte. Eine Pilotstudie (n = 24 Teilnehmer) hatte ergeben, dass eine zufällige Positionierung der Reize in der Kontrollbedingung ohne exogene Hinweisreize die Versuchsteilnehmer offenbar stark ablenkt und zu unspezifischen Performanzeinbußen führt. Im hier berichteten Hauptexperiment wurden die Reize in der Kontrollbedingung deshalb immer an derselben Position eines ungeteilten Präsentationsfensters präsentiert.

Endogene Aufgabenhinweise wurden dadurch manipuliert, dass die Aufgabensequenzen entweder einem festliegenden Muster der Art AABB folgten (EG), oder in zufälliger Abfolge präsentiert wurden (KG). Die Länge dieser Sequenzen wurde durch den dritten Faktor determiniert, durch den Effekte der Nachwirkung vorangegangener Aufgabenbearbeitung manipuliert werden sollten. In einer Bedingung betrug die mittlere Länge der Sequenzen mit Aufgabenwiederholung zwei Durchgänge, in einer zweiten Bedingung waren die Sequenzen im Mittel vier Durchgänge lang. Für die Bedingungen mit festliegender Aufgabensequenz war die Länge der Wiederholungssequenzen exakt zwei (AABB) beziehungsweise vier (AAAABBBB), für die Bedingungen mit zufälliger Reihenfolge wurden die Aufgabenblöcke so gestaltet, dass sich für jeden Block die vorgegebene Zahl von Aufgabenwiederholungen im Mittel ergab, jedoch ohne vorhersagbare Aufgabensequenz. Die Versuchsteilnehmer wurden über die exogenen und endogenen Aufgabenhinweise informiert und gebeten, diese zu nutzen, um sich jeweils auf die nachfolgende Aufgabe vorzubereiten. Am Ende des Experiments konnten die Teilnehmer Informationen über ihr persönliches Abschneiden, den IAT im Allgemeinen und die Hypothesen der Untersuchung abrufen.

Ergebnisse

Antwortlatenzen unter 300 ms wurden zu 300 ms (0.9 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (1 % der Durchgänge) rekodiert. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (7.2 % der Durchgänge) nicht in die Analyse aufgenommen.

Die mittleren Antwortlatenzen in den beiden gemischten Phasen des IAT wurden mit einer 2 (exogene Aufgabenhinweise ja/nein) x 2 (vorhersagbare Aufgabensequenzen ja/nein) x 2 (lange versus kurze Aufgabenwiederholungssequenzen) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf den letzten beiden Faktoren ausgewertet. Über die Versuchsbedingungen hinweg ergeben sich ein statistisch bedeutsamer IAT-Effekt von 83 ms, $F(1, 120) = 110.12, p < .01$, sowie Aufgabenwechselkosten in Höhe von 183 ms, $F(1, 120) = 200.71, p < .01$. Wie in den bisher berichteten Experimenten fallen die Aufgabenwechselkosten in der inkompatiblen Phase größer aus als in der kompatiblen, $F(1, 120) = 65.02, p < .01$, das heißt, die Kompatibilitätseffekte sind für Aufgabenwechseldurchgänge signifikant größer.

Eine vorhersagbare Abfolge von Attribut- und Zielkonzeptaufgabe führt zu einer Reduktion der IAT-Effekte um 43 ms, $F(1, 120) = 7.41, p < .01$. In der Bedingung mit endogen vorhersagbaren Sequenzen findet sich ein mittlerer IAT-Effekt von 62 ms. Mit zufälligen Sequenzen liegt der IAT-Effekt über die restlichen Bedingungen hinweg bei 105 ms. Der Effekt vorhersagbarer Aufgabensequenzen betrifft vor allem die inkompatible Phase. Der Kontrast zwischen den Bedingungen mit und ohne diesen endogenen Aufgabenhinweis wird in der inkompatiblen Phase signifikant, $F(1, 120) = 4.66, p < .05$, und bleibt in der kompatiblen Phase aus, $F(1, 120) = 0.77, n.s.$

Die Präsentation exogener Aufgabenhinweise führt zu einer Verringerung des mittleren IAT-Effekts um 16 ms, $F(1, 120) = 6.10, p < .05$. In der Bedingung ohne exogene Aufgabenhinweise findet sich, über die restlichen Bedingungen gemittelt, ein IAT-Effekt von 80 ms. In der Bedingung mit exogenen Aufgabenhinweisen liegt dieser Effekt bei 64 ms. Deskriptiv geht der Effekt exogener Aufgabenhinweise auf die inkompatible Phase des IAT zurück. Der Kontrast zwischen den Bedingungen mit und ohne exogene Aufgabenhinweise verfehlt in der inkompatiblen Phase jedoch knapp das Signifikanzniveau, $F(1, 120) = 3.77, p = .054$, während der gleiche Effekt in der kompatiblen Phase des IAT nicht einmal andeutungsweise auftritt, $F(1, 120) = 0.66, n.s.$

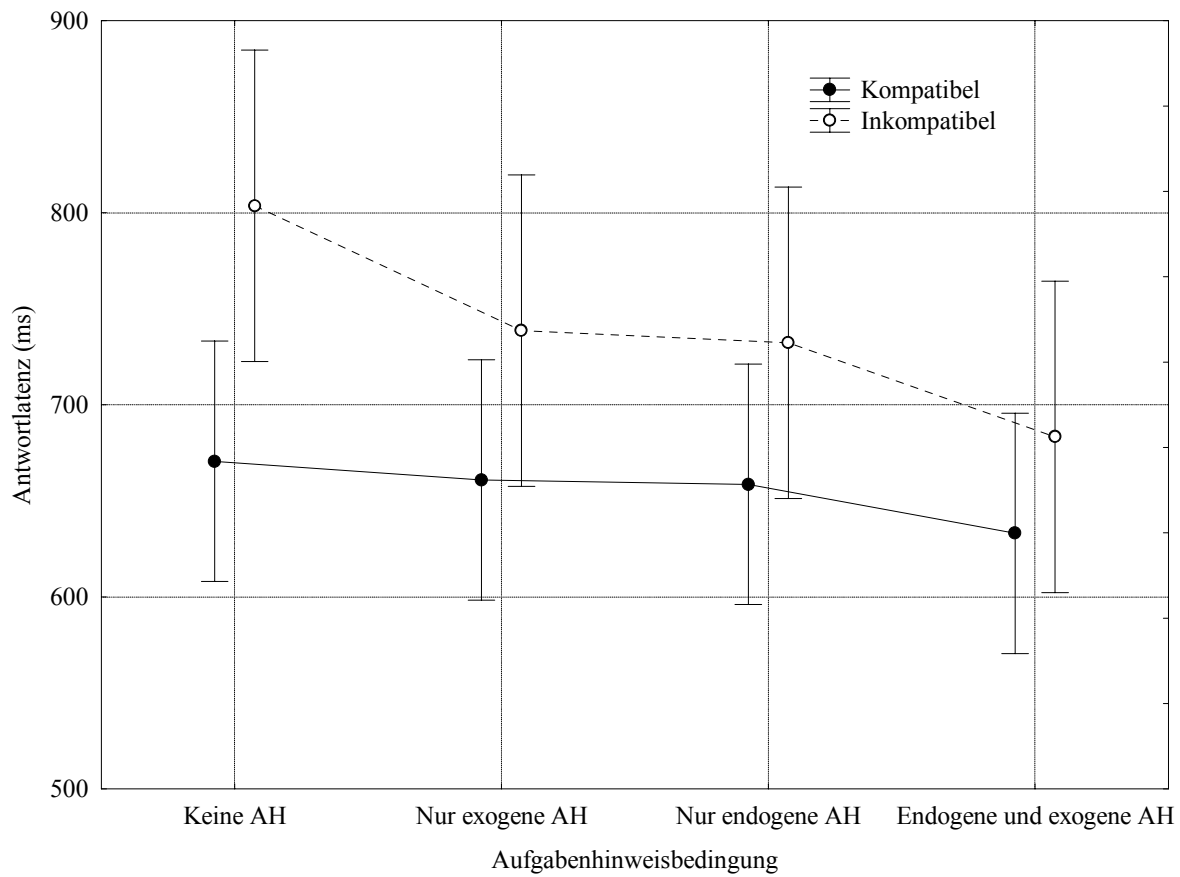


Abbildung 12: Effekte endogener und exogener Aufgabenhinweise (AH) auf die Antwortlatenzen in der kompatiblen und inkompatiblen Bedingung des IAT von Experiment 8. Beide Arten von Aufgabenhinweisen reduzieren die IAT-Effekte. Die Effekte exogener und endogener Aufgabenhinweise sind unabhängig.

Eine Dreifach-Interaktion zwischen den beiden Aufgabenhinweisbedingungen und der Kompatibilität bleibt aus, $F(1, 120) = 0.99$, n.s. Die Effekte sind also offenkundig unabhängig voneinander. Abbildung 12 zeigt die Effekte endogener und exogener Aufgabenhinweise auf die Größe der IAT-Effekte. Wenn beide Arten von Aufgabenhinweisen präsent sind, liegt der mittlere IAT-Effekt bei 50 ms. Der additive Effekt der beiden Arten von Aufgabenhinweisen reduziert den mittleren IAT-Effekt von 132 ms in der Versuchsbedingung ohne jegliche Aufgabenhinweise um 82 ms in der Bedingung mit beiden Arten von Aufgabenhinweisen. Die Präsentation beider Arten von Aufgabenhinweisen hat jedoch wie in Experiment 7 keinen direkten Einfluss auf die Aufgabenwechselkosten ($F(1, 120) = 3.11$, n.s. für die exogenen Aufgabenhinweise und $F(1, 120) = 1.91$, n.s. für die endogenen). Es findet sich wie in Experiment 7 keinerlei Hinweis auf eine Interaktion zwischen den Aufgabenhinweisen, den Aufgabenwechselkosten und der Kompatibilität der Phase, beide $F(1, 120) < 0.52$, n.s.

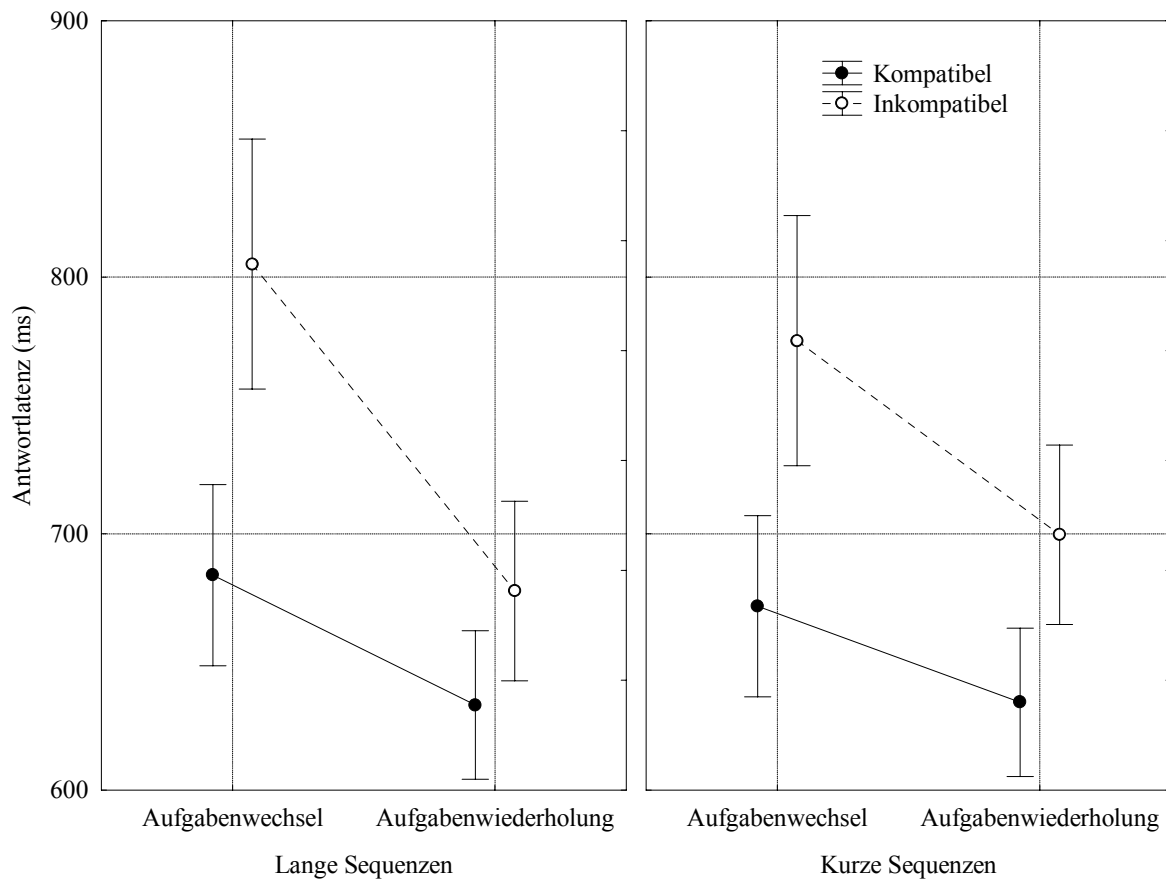


Abbildung 13: Antwortlatenzen in der kompatiblen und inkompatiblen Phase des IAT als Funktion der Länge von Aufgabenwiederholungssequenzen (lange Seq. versus kurze Seq.) für Aufgabenwechsel- und Aufgabenwiederholungsdurchgänge. In der Bedingung mit langen Sequenzen sind exakt 11, in der Bedingung mit kurzen Sequenzen exakt 23 Aufgabenwechsel pro Block zu bearbeiten. Mit langen Sequenzen ergibt sich eine größere Asymmetrie der Aufgabenwechselkosten (linke Hälfte der Abbildung) als in der Bedingung mit kurzen Sequenzen (rechte Hälfte der Abbildung).

Die Manipulation der Länge der Aufgabenwiederholungssequenzen hat den erwarteten Effekt auf die Aufgabenwechselkosten, $F(1, 120) = 9.98, p < .01$. In der Bedingung mit langen Wiederholungssequenzen liegen die Aufgabenwechselkosten bei 80 ms, in der Bedingung mit kurzen Wiederholungssequenzen bei 56 ms. Es findet sich die erwartete Dreifach-Interaktion zwischen der Länge der Wiederholungssequenzen, den Aufgabenwechselkosten und der Kompatibilität der Phase, $F(1, 120) = 7.32, p < .01$. Die Asymmetrie der Aufgabenwechselkosten wird größer, je länger die Wiederholungssequenzen ausfallen. Die Interaktion von Aufgabenwechselkosten und Sequenzlänge betrifft, wie ein entsprechender Kontrast zeigt, die inkompatible Phase, $F(1, 120) = 11.59, p < .01$, nicht

jedoch die kompatible, $F(1, 120) = 2.12$, n.s. Die zugehörigen mittleren Antwortlatenzen können Abbildung 13 entnommen werden.

Diskussion

Die systematische Dekomposition dreier Komponenten exekutiver Kontrolle in Experiment 8 ergibt ein interessantes Befundmuster. Zunächst zeigt sich, dass sowohl exogene als auch endogene Komponenten exekutiver Kontrolle den IAT-Effekt beeinflussen können. Sowohl die Ankündigung der nächsten Aufgabe durch die Präsentation eines externen Reizes als auch die vorhersagbare Sequenz der Aufgaben führten zu einer besseren Performanz in der inkompatiblen Phase. Die Performanz in der kompatiblen Phase war durch beide Manipulationen unbeeinflusst. Interessanterweise geht die Interaktion zwischen der Kompatibilität der Tastenzuordnung und den Aufgabenhinweisen weder im gegenwärtigen Experiment noch in Experiment 7 auf eine Verringerung der spezifischen Aufgabenwechselkosten zurück. Offenbar reduzieren die Aufgabenhinweise also eher allgemeine Kosten für die Mischung von Attribut- und der Zielkonzeptaufgabe in der inkompatiblen Phase des IAT. Aus der Befundlage kann gefolgert werden, dass diese allgemeinen Mischungskosten (Los, 1996, 1999) sowohl Aufgabenwechsel- als Aufgabenwiederholungsdurchgänge betreffen.

Ein Faktor, der die Leistung in einem Block mit gemischten Aufgaben unabhängig von der Art des Durchgangs betreffen sollte, ist das Behalten der Tastenzuordnung für die Attribut- und Zielkonzeptaufgabe. Wenn, wie der deterministische Vereinfachungsmechanismus vorhersagt, in der kompatiblen Phase lediglich die Attributaufgabe angewendet wird, so sollte die Belastung des Arbeitsgedächtnisses in der inkompatiblen Phase des IAT größer ausfallen, da hier die komplette Aufgabenstruktur behalten werden muss. Die in Experiment 7 verwendete symbolische Repräsentation der Aufgabenstruktur könnte nun zu einer Entlastung des Arbeitsgedächtnisses geführt haben, die lediglich in der inkompatiblen Phase zum Tragen kommt. Es scheint jedoch, dass die Effekte der deutlich abstrakteren Aufgabenhinweise in Experiment 8 mit dieser Annahme schlecht erklärbar sind. Insbesondere der Effekt endogener Kontrollprozesse bei vorhersagbarer Aufgabensequenz ist durch diesen Mechanismus schlecht zu erklären. Zum einen enthalten die Sequenzen keinerlei Information über die Aufgabenstruktur, zum anderen sollte die Nutzung solcher Sequenzen eher eine Mehrbelastung des Arbeitsgedächtnisses zur Folge haben, denn eine Vorbereitung kann nur

dann erfolgen, wenn der Versuchsteilnehmer den Ablauf der Sequenzen mitverfolgt und kognitiv repräsentiert.

Ein weiterer Faktor, der die Leistung in gemischten Aufgabenblöcken verringern sollte, ist die Identifikation der auszuführenden Aufgabe. Bei kompatibler Tastenzuordnung sollte die korrekte Identifikation der zu bearbeitenden Aufgabe weniger entscheidend sein als bei inkompatibler, da die Aufgabensets von Attribut- und Zielkonzeptaufgabe in den meisten Fällen die gleiche Antwort für einen gegebenen Stimulus spezifizieren. Es ist deshalb zu vermuten, dass entsprechende Kosten für die Identifikation der Aufgabe in der inkompatiblen Phase eines IAT deutlich ausgeprägter sind. Die Identifikation der auszuführenden Aufgabe muss zwingend vor der Durchführung eines Aufgabenwechsels erfolgen, jedoch für Aufgabenwiederholungs- und -wechseldurchgänge in gleichem Maße. Aus dieser Perspektive könnte plausibel erklärt werden, warum die Aufgabenhinweise in den vorliegenden Experimenten sowohl Aufgabenwechsel- als auch -wiederholungsdurchgänge betreffen - jedoch nur in der inkompatiblen Phase, die eine korrekte Identifikation der Aufgabe erfordert.

Ungeachtet der konkreten Prozesse, die den vorliegenden Ergebnissen zugrunde liegen, scheint es, dass ein Teil der Varianz in den IAT-Effekten, der unabhängig von spezifischen Aufgabenwechselkosten auftritt, dennoch mit Aufgabenwechselprozessen in Verbindung steht. Anders kann kaum erklärt werden, warum die IAT-Effekte in den Experimenten 7 und 8 unabhängig von spezifischen Wechselkosten durch Aufgabenhinweise reduziert waren.

Die Interpretation der Effekte langer versus kurzer Sequenzen mit Aufgabenwiederholung fällt leichter. Wie auf Basis des Trägheitsmechanismus erwartet, fallen die Aufgabenwechselkosten größer aus, wenn von einer Aufgabe weg gewechselt werden muss, die über mehrere Durchgänge hinweg eingeübt wurde. Dies gilt jedoch nur für die inkompatible Phase. Die Spezifität der Effekte für Aufgabenwechseldurchgänge in der inkompatiblen Phase passt gut zu der theoretischen Interpretation der Trägheitseffekte als Kostenkomponente, die mit der *Implementation* von Aufgabensets zu tun hat. Anders als die Identifikation der auszuführenden Aufgabe sollten solche Effekte nur bei Aufgabenwechseldurchgängen auftreten, was sich in den Daten bestätigt.

Die Befunde der Experimente 7 und 8 sprechen zusammengenommen dafür, dass IAT-Effekte zumindest teilweise auf exekutive Kontrollprozesse zurückgehen. Die Wirkung von Aufgabenhinweisen scheint die Präsenz des zu verarbeitenden Stimulus vorauszusetzen, wie Experiment 7 zeigt. Beide Experimente legen nahe, dass Aufgabenhinweise eher allgemeine

Mischungskosten als spezifische Wechselkosten reduzieren. Eine plausible Erklärung dieses Befundmusters scheint zu sein, dass die Aufgabenhinweise zwar die Identifikation der Aufgabe erleichtern, nicht jedoch ihre Implementation. Die spezifischen Aufgabenwechselkosten scheinen vornehmlich autogene Komponenten des Wechselprozesses widerzuspiegeln, die nicht durch Aufgabenhinweise beeinflussbar sind, sondern auf die Trägheit von Aufgabensets zurückgehen. Dafür sprechen die in den Experimenten 4 bis 6 aufgetretenen längerfristigen Trägheitseffekte, die Verringerung der Wechselkostenasymmetrie durch ein längeres RSI in Experiment 1 und die stärkere Asymmetrie der Aufgabenwechselkosten bei längeren Sequenzen mit Aufgabenwiederholung in Experiment 8. Auch der Effekt der Kompatibilität auf die Attributaufgabe, dem keine Interferenz durch ein irrelevantes Attribut zugrunde liegen kann, deutet auf Trägheitseffekte hin.

7 Experimente zu methodenspezifischer Varianz

Zum Abschluss der vorliegenden Arbeit soll eine Serie von vier Experimenten dargestellt werden, die sich mit den praktischen Konsequenzen von Aufgabenwechselkosten im IAT befasst. Ziel dieser Experimente ist es zu prüfen, inwiefern systematische interindividuelle Unterschiede in Fähigkeitsparametern, die an Aufgabenwechselprozessen beteiligt sind, einen systematischen und stabilen Einfluss auf IAT-Ergebnisse haben können. Wie weiter unten genauer ausgeführt wird, hat eine Beteiligung solcher Fähigkeitsparameter an der Varianz des IAT bedeutsame Konsequenzen für die Einschätzung der psychometrischen Qualität des IAT. Aus theoretischer Sicht sind zunächst drei Hypothesen bezüglich der Konfundierung kognitiver Fähigkeitsparameter mit dem IAT-Effekt denkbar, die im Folgenden formuliert werden sollen.

Vermittlungshypothese

Die Vermittlungshypothese behauptet, dass das Ausmaß von Aufgabenwechselkosten von den zugrunde liegenden Assoziationsmustern abhängig ist. Die Befunde der bisher geschilderten Experimente legen dies insofern nahe, als Aufgabenwechselkosten jeweils in der inkompatiblen Phase eines IAT verstärkt auftraten. Dies bestätigt zunächst die elementare Annahme, dass der IAT-Effekt die Richtung der zu messenden Assoziationen abbildet. Die Vermittlungshypothese nimmt zusätzlich an, dass das Ausmaß der Aufgabenwechselkosten eine direkte und monotone Funktion der Assoziationsstärke ist. Aus Sicht des Trägheitsmechanismus scheint diese Annahme berechtigt, da die Trägheit von Aufgabensets nur dann Auswirkungen haben sollte, wenn ein irrelevantes Attribut der Zielkonzeptreize zu Interferenzen bei der Aufgabenbearbeitung führt. Dies ist offenbar nur dann der Fall, wenn bei der Verarbeitung der Zielkonzeptreize irrelevante Attributinformation aktiviert wird, die zur korrekten Bearbeitung der Aufgabe unterdrückt werden muss. Damit ist nicht ausgeschlossen, dass der resultierende Effekt auf andere Relationen als die Assoziation der Zielkonzepte mit dem Attribut zurückgeht. Die Vermittlungshypothese behauptet lediglich eine monotone Abhängigkeit der Aufgabenwechsel- und Mischungskosten von der Stärke einer zugrunde liegenden Relation.

Unsystematische Kontaminierung

Die Hypothese unsystematischer Kontaminierung besagt, dass Aufgabenwechselkosten eine inhaltsunabhängige, aber unsystematische Kontaminierung des IAT-Effekts darstellen. Die triviale Konsequenz einer unsystematischen Kontaminierung wäre eine Vergrößerung des Anteils der Fehlervarianz an den Messungen. Dies hätte offenkundig eine Verringerung von Stabilität, interner Konsistenz und Validität der Messungen zur Folge.

Systematische Kontaminierung

Die Hypothese systematischer Kontaminierung behauptet, dass Aufgabenwechselkosten durch einen Fähigkeitsparameter bestimmt werden, der systematische und inhaltsunabhängige interindividuelle Unterschiede aufweist. Einen Hinweis auf die prinzipielle Möglichkeit einer solchen systematischen Kontaminierung liefern vor allem Arbeiten aus dem Bereich der Entwicklungspsychologie. Ein zentrales Ergebnis dieser Arbeiten ist, dass Aufgabenwechselkosten eine altersabhängige Komponente aufweisen, die mit vielen verschiedenen Materialien und Aufgabenstellungen auftritt (z. B. Cepeda et al., 2001; Kramer et al., 1999; Salthouse, Fristoe, McGuthry & Hambrick, 1998; Verhaeghen & Cerella, 2002). Im höheren Lebensalter ist die Leistung in Aufgabenblöcken, die den Wechsel zwischen verschiedenen Aufgaben erfordern, verringert. Eine Meta-Analyse von Verhaeghen und Cerella (2002) zeigt, dass allgemeine Mischungskosten eine spezifische Altersabhängigkeit aufweisen, die nicht durch eine globale kognitive Verlangsamung erklärt werden kann. Die spezifischen oder lokalen Aufgabenwechselkosten, die sich im Kontrast von Aufgabenwechsel- und –wiederholungsdurchgängen zeigen, sind ebenfalls altersabhängig, können jedoch durch unspezifische Veränderungen der kognitiven Leistungsfähigkeit erklärt werden (Verhaeghen & Cerella, 2002), wie sie am Zusammenhang zwischen fluider Intelligenzleistung und Aufgabenwechselkosten deutlich werden (Kray & Lindenberger, 2000; Salthouse et al., 1998).

Die systematische Kontaminierung des IAT unterscheidet sich von einer unsystematischen Kontaminierung vor allem dadurch, dass die entstehende methodenspezifische Varianz stabil und systematisch ausfällt. Unter anderem sind bedeutsame Korrelationen zwischen IATs zu erwarten, die von den spezifischen Inhalten unabhängig sind. Der Nachweis methodenspezifischer Varianz ist jedoch nicht immer direkt möglich. Wie weiter oben ausgeführt, hängt die *Richtung* der Effekte in verschiedenen IATs von der Richtung der zugrunde liegenden inhaltlichen Relationen ab. Wenn die Richtung der IAT-Effekte innerhalb

einer Person über die verschiedenen Inhalte hinweg korreliert, so führt methodenspezifische Varianz - gleichläufig mit der Variation in den Vorzeichen - zu einer Erhöhung der Korrelation zwischen den Maßen, das heißt, zu einer Überschätzung der Konstruktvalidität. Inhaltliche Varianz und methodenspezifische Varianz sind in diesem Fall konfundiert, und eine direkte Prüfung des Anteils methodenspezifischer Varianz an der Gesamtvarianz der IAT-Scores ist nicht ohne weiteres möglich.

Soll der Anteil methodenspezifischer Varianz an der Gesamtvarianz geschätzt werden, ist eine Dissoziation zwischen inhaltlicher und methodenspezifischer Varianz Voraussetzung. Dies ist immer dann der Fall, wenn ein inhaltlich begründeter Zusammenhang zwischen den IATs ausgeschlossen werden kann. Insbesondere ist ein Nachweis methodenspezifischer Varianz möglich, wenn die Richtung der IAT-Effekte über die Teilnehmer hinweg konstant ist, das heißt, wenn keine interindividuelle Variation in den Vorzeichen der Scores vorliegt. Diese Bedingung ist näherungsweise für den Blumen-Insekten Einstellungs-IAT (Greenwald et al., 1998) und IATs zur Einstellung gegenüber älteren Menschen (Nosek et al., 2002) erfüllt, bei denen sich eine relativ homogene Präferenz der Blumen beziehungsweise der jüngeren Menschen ergibt. Auch der in Experiment 2 verwendete IAT mit geometrischen Objekten genügt dieser Anforderung, da die Richtung der IAT-Effekte bei konstanter Kovariation ebenfalls konstant bleibt. In diesen Fällen kann der Anteil methodenspezifischer Varianz direkt durch Korrelationen bestimmt werden.

Auch wenn die Richtung der IAT-Effekte innerhalb der Teilnehmer zufällig variiert, kann der Anteil methodenspezifischer Varianz bestimmt werden, sofern ein inhaltlicher Zusammenhang zwischen den IATs ausgeschlossen werden kann. In diesem Fall sollte sich methodenspezifische Varianz in Korrelationen zwischen den *Absolutwerten* der IAT-Effekte äußern. Die Tatsache, dass methodenspezifische Varianz durch eine orthogonale Variation der Vorzeichen verdeckt werden kann, lässt sich leicht verdeutlichen. Angenommen, einige Teilnehmer bearbeiteten den IAT mit geometrischen Objekten aus Experiment 2 mehrfach hintereinander. Die Kovariation zwischen Größe und Farbe der geometrischen Objekte werde für jede Durchführung zufällig bestimmt. Eine direkte Prüfung des Anteils methodenspezifischer Varianz sollte in diesem Fall selbst dann zu einer Nullkorrelation führen, wenn die Stärke der Effekte vollkommen von der Ausprägung der Aufgabenwechselkosten abhängt. Erst wenn die Absolutwerte der IAT-Scores analysiert werden, zeigt sich der Zusammenhang. Dies sollte analog gelten, wenn inhaltlich nicht zusammenhängende IATs verwendet werden.

Die Korrelation der Absolutwerte liefert dann eine Schätzung des Anteils methodenspezifischer Varianz.

Konsequenzen methodenspezifischer Varianz

Eine Kontaminierung von IAT-Scores mit methodenspezifischer Varianz hat verschiedene problematische Konsequenzen für die Interpretation von Ergebnissen des IAT auf individueller Ebene, die im Folgenden kurz aufgezeigt werden sollen.

Wie weiter oben ausgeführt, sollte eine Kontaminierung mit methodenspezifischer Varianz immer dann zu einer Erhöhung von Korrelationen zwischen zwei IATs führen, wenn die Vorzeichen der IAT-Effekte in die gleiche Richtung weisen. Diese Bedingung ist offenbar bei Testwiederholungen erfüllt, sofern die Vorzeichen entweder über die Stichprobe hinweg konstant bleiben oder systematisch und stabil variieren. Im ersten Fall kann methodenspezifische Varianz zu einer vergleichsweise hohen Reliabilitätsschätzung führen, selbst wenn keine systematische inhaltliche Varianz erfasst wurde. Im zweiten Fall ist eine Erhöhung der Reliabilität über das Maß systematischer inhaltlich begründeter Varianz hinaus zu erwarten. Entsprechende Verzerrungen sollten sowohl die Schätzungen der Retest-Stabilität als auch der internen Konsistenz beeinflussen.

Korrelationen zwischen IATs und Verhaltensdaten beziehungsweise expliziten Maßen sollten durch eine Kontaminierung mit methodenspezifischer Varianz eher verringert sein. Zwar sind in diesen Fällen strukturell andersartige Quellen systematischer Verzerrungen, etwa eine Tendenz zu sozialer Erwünschtheit oder transsituationaler Konsistenz, zu erwarten. In Bezug auf die inhaltliche Varianz in diesen Verfahren erhöht die methodenspezifische Varianz im IAT jedoch lediglich den Anteil des Messfehlers. Ein problematischer Effekt methodenspezifischer Varianz ergibt sich jedoch, wenn die Richtung der Effekte über verschiedene IATs hinweg konstant ausfällt. Dies kann bei einem Einstellungs-IAT der Fall sein, wenn die Zielkonzepte über die Stichprobe hinweg homogen bewertet werden. In diesem Fall ginge methodenspezifische Varianz direkt in die Korrelationen zwischen verschiedenen IATs ein. Eine bedeutsame Korrelation zwischen verschiedenen IATs kann im Extremfall auch dann auftreten, wenn keine inhaltliche Überlappung vorliegt. Die Interpretation von theoretisch erwarteten Korrelationen zwischen IATs als Konstruktvalidität könnte insofern irreführend sein.

Zusammengenommen ist also zu erwarten, dass methodenspezifische Varianz zu einer Verringerung der Korrelationen von IATs mit expliziten und Verhaltensmaßen führt, während

Korrelationen bei wiederholter Testung ebenso wie Schätzungen der internen Konsistenz artifiziell erhöht sein sollten. Werden inhaltlich verschiedene IATs mit konstanten Vorzeichen erhoben, so ist ferner mit einer überhöhten Schätzung der Konstruktvalidität zu rechnen.

Das Ziel der im Folgenden geschilderten Serie von Experimenten ist es, methodenspezifische Varianz in verschiedenen Varianten des IAT nachzuweisen. In den Experimenten 9 und 10 wird zunächst methodenspezifische Varianz bei der Wiederholung des Blumen-Insekten Einstellungs-IAT und des IAT mit geometrischen Objekten geprüft. In den Experimenten 11 und 12 werden der Blumen-Insekten Einstellungs-IAT sowie ein Selbst-Andere Extraversions-IAT jeweils mit dem Geometrie-IAT gepaart. Bei diesen letzten beiden Experimenten kann ausgeschlossen werden, dass bedeutsame Korrelationen der Verfahren auf inhaltliche Überlappungen zurückgehen.

7.1 Experiment 9

In Experiment 9 sollte zunächst geprüft werden, ob und inwiefern der in Experiment 2 erprobte IAT mit geometrischen Objekten Hinweise auf systematische methodenspezifische Varianz aufweist. Da der Zusammenhang zwischen Größe und Farbe der Objekte in diesem IAT für alle Versuchsteilnehmer in gleicher Weise experimentell manipuliert wurde, sind keine systematischen interindividuellen Unterschiede in der Ausprägung des gemessenen Konstrukts zu erwarten. Signifikante Korrelationen zwischen verschiedenen Durchführungen dieses IAT können folglich nicht auf gemeinsame Varianz eines zugrunde liegenden Konstrukts zurückgehen, sondern müssen durch methodenspezifische Varianz vermittelt sein. Die Versuchsteilnehmer bearbeiteten diesen IAT zweimal in direkter Folge. Die zu prüfende Fragestellung ist, ob die IAT-Effekte des kontingenzbasierten IAT über die Messzeitpunkte korreliert sind, oder nicht. Eine bedeutsame Test-Retest-Korrelation wäre ein erster Hinweis auf eine systematische und stabile Kontamination des IAT mit methodenspezifischer Varianz.

Methode

Die Teilnehmer von Experiment 9 bearbeiteten den gleichen IAT mit geometrischen Objekten zweimal in direkter Folge. Wie in Experiment 2 waren die Zielkonzepte rot und blau gefärbte Objekte. Das zu beurteilende Attribut war die Größe der Objekte. Die Exemplare der Attributkategorien gehörten zu keinem der Zielkonzepte, das heißt, sie waren nie blau oder rot gefärbt, sondern wiesen immer eine von drei alternativen Farben auf. Wie in einem Einstellungs-IAT sind die Zielkonzepte disjunkte Kategorien, während das Attribut Größe

(ähnlich dem Attribut Valenz) eine kontinuierliche Variable ist. Im Unterschied zu einem herkömmlichen IAT bestehen jedoch keine präexperimentellen Assoziationen zwischen Attributausprägung und den Zielkonzepten – also der Größe und Färbung der Objekte - und aus diesem Grund auch keine systematischen interindividuellen Unterschiede bezüglich des Zusammenhangs.

Wie in Experiment 2 wurde der Zusammenhang zwischen Farbe und Größe der Objekte durch eine Kontingenz der Stimuluseigenschaften hergestellt. Über alle experimentellen Bedingungen hinweg waren alle Exemplare des Zielkonzepts „rot“ kleine Objekte und alle Exemplare des Zielkonzepts „blau“ große Objekte. Mit diesem Material ist eine Tastenzuordnung kompatibel, wenn rote und kleine Objekte einer Taste zugeordnet sind, und blaue und große der verbleibenden, da in diesem Fall alle großen Objekte einer Taste zugeordnet sind und alle kleinen Objekte der anderen. Die Kompatibilitätsreihenfolge wurde für die erste und zweite Durchführung des IAT gleichlaufend ausbalanciert. Zusätzlich wurde die komplette Tastenzuordnung für eine Hälfte der Versuchspersonen umgekehrt.

Stichprobe

An der Untersuchung nahmen 25 Personen teil. Acht Teilnehmer waren männlichen und 17 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 19 und 39 Jahren ($M = 25.76$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Als Material dienten die gleichen geometrischen Objekte wie in Experiment 2.

Prozedur

Der Ablauf des Experiments und die Darbietungsparameter der einzelnen Durchgänge entsprechen denen der Experimente 1 und 2. Allerdings bearbeitete jede Versuchsperson den gleichen IAT zweimal in direkter Folge. Die Wiederholung des IAT war in jeder Hinsicht identisch mit der ersten Durchführung. Lediglich die Instruktion wurde bei der zweiten Durchführung gekürzt, um Redundanzen zu vermeiden. Jede Versuchsperson bearbeitete in beiden IATs insgesamt 6 Übungsblöcke, in denen Attribut- und Zielkonzeptaufgabe getrennt eingeübt wurden und 8 Experimentalblöcke, in denen Attribut- und Zielkonzeptaufgabe

gemischt zu bearbeiten waren, je vier mit kompatibler und vier mit inkompatibler Tastenzuordnung. Das RSI betrug 800 ms.

Ergebnisse

Antwortlatenzen unter 100 ms wurden zu 100 ms (0.2 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.2 % der Durchgänge) rekodiert. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (6.3 % der Durchgänge) nicht in die Analyse aufgenommen. Die mittleren IAT-Scores lagen bei der ersten Durchführung des IAT bei 374 ms (SD = 155) und bei der zweiten bei 281 ms (SD = 75). Alle Versuchsteilnehmer hatten positive IAT-Effekte. Alle Effekte lagen zwischen 132 ms und 687 ms.

Psychometrische Analyse

Da im vorliegenden Experiment die psychometrischen Eigenschaften des kontingenzbasierten Geometrie-IAT von zentralem Interesse sind, wurde für die erste und zweite Durchführung des Tests die interne Konsistenz nach Cronbach bestimmt. Dazu wurden die Durchgänge der gemischten Phasen zufällig einem von vier Subtests zugewiesen. Für jeden der vier Subtests wurde ein IAT-Effekt berechnet, und anschließend Cronbachs α für den jeweiligen Messzeitpunkt. Die interne Konsistenz lag für den ersten Messzeitpunkt bei $\alpha = .95$, und bei $\alpha = .79$ für den zweiten. Ähnlich wie in der Untersuchung von Cunningham (2001) findet sich also deskriptiv eine graduelle Abnahme der internen Konsistenz über die Messzeitpunkte. Die Korrelation zwischen den IAT-Scores der beiden Messzeitpunkte ist, wie es die Hypothese einer systematischen Kontaminierung des IAT vorhersagt, bedeutsam größer als Null, $r = .66$, $p < .01$.

Varianzanalyse

Die mittleren Antwortlatenzen in den gemischten Phasen des IAT wurden mit einer 2 (Kompatibilitätsreihenfolge) x 2 (erster versus zweiter Messzeitpunkt) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf den letzten drei Faktoren ausgewertet. Ein Haupteffekt des Messzeitpunkts deutet an, dass die mittleren Antwortlatenzen am zweiten Messzeitpunkt kürzer ausfallen, als am ersten, $F(1, 23) = 54.79$, $p < .01$. Der IAT-Effekt ist über die Messzeitpunkte hinweg signifikant, wie der Haupteffekt des Faktors Kompatibilität zeigt, $F(1, 23) = 263.12$, $p < .01$. Eine Zweifach-

Interaktion von Kompatibilität und Messzeitpunkt geht zeigt jedoch, dass die mittleren IAT-Effekte am zweiten Messzeitpunkt signifikant kleiner ausfallen, $F(1, 23) = 17.85, p < .01$. Wie in den vorigen Experimenten sind die Latenzen bei Aufgabenwechseldurchgängen langsamer als bei Aufgabenwiederholung, $F(1, 23) = 60.96, p < .01$. Eine Zweifach-Interaktion zwischen Kompatibilität und Aufgabenwechsel zeigt jedoch, dass die Aufgabenwechselkosten in der inkompatiblen Phase des IAT größer ausfallen als in der kompatiblen, $F(1, 23) = 44.44, p < .01$. Eine Dreifach-Interaktion zwischen Kompatibilitätsreihenfolge, dem Kompatibilitätsfaktor und dem Messzeitpunkt verfehlt das Signifikanzniveau, $F(1, 23) = 3.12, n.s.$ Deskriptiv tritt eine Interaktion zwischen Kompatibilitätsreihenfolge und dem IAT-Effekt am ersten Messzeitpunkt tendenziell auf, $F(1, 23) = 4.05, p = .06$, nicht aber am zweiten $F(1, 23) = 1.57, n.s.$

Diskussion

Die Ergebnisse von Experiment 9 zeigen die erwartete Korrelation zwischen der ersten und zweiten Durchführung des IATs mit geometrischen Objekten. Ferner zeigen sich befriedigende interne Konsistenzen zu beiden Messzeitpunkten. Wie schon in der Einleitung diskutiert, kann dieser Befund nicht durch interindividuelle Unterschiede in einem gemessenen Konstrukt erklärt werden, da der vorliegende IAT den Effekt einer konstanten experimentellen Manipulation abbildet. Es muss also gefolgert werden, dass die aufgetretenen Zusammenhänge auf systematische und stabile methodenspezifische Varianz zurückgehen. Interessant ist in diesem Zusammenhang auch, dass die Korrelation zwischen den IATs ebenso wie die internen Konsistenzen in ihrer Höhe in etwa dem entsprechen, was in anderen Untersuchungen mit inhaltlich angereicherten, herkömmlichen Einstellungs-IATs gefunden wurde (z. B. Cunningham et al., 2001; Steffens & Buchner, 2003). Offenkundig scheinen systematische interindividuelle Unterschiede, wie sie in diesen Untersuchungen als Grundlage der Zusammenhänge vermutet wurden, keine notwendige Bedingung für das Erreichen vergleichbar guter psychometrischer Eigenschaften zu sein.

7.2 Experiment 10

In Experiment 10 sollten die Befunde von Experiment 9 mit einem inhaltlich angereicherten Material repliziert werden. Die Versuchsteilnehmer sollten dazu wieder einen IAT, für den keine bedeutsamen interindividuellen Unterschiede zu erwarten sind, zweimal in direkter Folge bearbeiten. In Anlehnung an Greenwald et al. (1998) wurde dabei ange-

nommen, dass sich bezüglich der Bewertung von Blumen und Insekten keine systematischen Unterschiede zwischen den Versuchsteilnehmern finden.

Methode

Die Versuchsteilnehmer bearbeiteten zweimal in direkter Folge einen IAT mit den Blumen und Insektenreizen, die schon in verschiedenen Vorgängerexperimenten verwendet wurden. Kompatibilitätsreihenfolge und Tastenzuordnung wurden wie in Experiment 9 ausbalanciert.

Stichprobe

An der Untersuchung nahmen 24 Personen teil. Zehn Teilnehmer waren männlichen und 14 waren weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 19 und 29 Jahren ($M = 22.75$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Als Material dienten die in den Experimenten 5, 6, 7 und 8 eingesetzten Wörter für Blumen und Insekten, sowie die dort verwendeten positiven und negativen Substantive.

Prozedur

Der Ablauf des Experiments und die Darbietungsparameter der einzelnen Durchgänge entsprechen denen von Experiment 9. Jede Versuchsperson bearbeitete pro IAT insgesamt 6 Übungsblöcke und 8 Experimentalblöcke, je vier mit kompatibler und vier mit inkompatibler Tastenzuordnung. Das RSI betrug 800 ms.

Ergebnisse

Antwortlatenzen unter 300 ms wurden zu 300 ms (0.6 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.3 % der Durchgänge) rekodiert. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (6.7 % der Durchgänge) nicht in die Analyse aufgenommen. Die mittleren IAT-Scores lagen bei der ersten Durchführung des IAT bei 187 ms ($SD = 116$) und bei der zweiten bei 102 ms ($SD = 70$). Alle Effekte lagen zwischen -44 ms und 389 ms.

Psychometrische Analyse

Die interne Konsistenz nach Cronbach wurde wie in Experiment 9 über vier zufällig zusammengestellte Subtests bestimmt. Sie lag für den ersten Messzeitpunkt bei $\alpha = .88$, und bei $\alpha = .82$ für den zweiten. Die Korrelation zwischen den IAT-Scores an den beiden Messzeitpunkten ist, wie es die Hypothese einer systematischen Kontaminierung des IAT vorhersagt, bedeutsam größer als Null, $r = .53$, $p < .01$.

Varianzanalyse

Die mittleren Antwortlatenzen in den gemischten Phasen des IAT wurden mit einer 2 (Kompatibilitätsreihenfolge) x 2 (erster versus zweiter Messzeitpunkt) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf den letzten drei Faktoren ausgewertet.

Der mittlere IAT-Effekt liegt über beide Messzeitpunkte hinweg bei 137 ms, und ist statistisch bedeutsam, $F(1, 22) = 80.97$, $p < .01$. Die Antwortlatenzen sind zum zweiten Messzeitpunkt insgesamt schneller als zum ersten, $F(1, 22) = 28.19$, $p < .01$. Wie im Vorgängerexperiment fällt auch der mittlere IAT-Effekt zum zweiten Messzeitpunkt geringer aus als am ersten, $F(1, 22) = 17.56$, $p < .01$. Die Latenzen sind bei Aufgabenwechselfurchgängen langsamer als bei Aufgabenwiederholung, $F(1, 22) = 145.98$, $p < .01$. Eine Zweifach-Interaktion zwischen Kompatibilität und Aufgabenwechsel zeigt, dass die Aufgabenwechselkosten in der inkompatiblen Phase des IAT erneut größer ausfallen als in der kompatiblen, $F(1, 22) = 40.23$, $p < .01$. Ferner fallen die Aufgabenwechselkosten am ersten Messzeitpunkt größer aus als am zweiten, $F(1, 22) = 8.36$, $p < .01$. Dieser Effekt geht vor allem auf eine größere Asymmetrie der Wechselkosten bei der ersten Durchführung des IAT zurück, wie eine Interaktion von Messzeitpunkt, Kompatibilität und Aufgabenwechsel zeigt, $F(1, 22) = 4.62$, $p < .05$. Wie zwei Kontraste verdeutlichen, ist die Asymmetrie der Aufgabenwechselkosten jedoch sowohl am ersten Messzeitpunkt signifikant, $F(1, 22) = 31.24$, $p < .01$, als auch am zweiten, $F(1, 22) = 10.32$, $p < .05$. Die Zweifach-Interaktion zwischen IAT-Effekt und Kompatibilitätsreihenfolge verfehlt das Signifikanzniveau, $F(1, 22) = 2.85$, $p = .11$, geht jedoch deskriptiv in die erwartete Richtung.

Diskussion

Die Ergebnisse mit dem Blumen-Insekten Einstellungs-IAT entsprechen weitgehend denen von Experiment 9. Auf Mittelwertsebene repliziert sich, wie in allen Vorgängerexperimenten, die Asymmetrie der Aufgabenwechselkosten. Im Gegensatz zum Effekt der Kompatibilitätsreihenfolge, der in den Experimenten 9 und 10 das Signifikanzniveau verfehlte, ist die Wechselkostenasymmetrie also ein äußerst stabiler Effekt. Ebenfalls stabil scheint der Befund zu sein, dass die IAT-Effekte bei Testwiederholung geringer ausfallen. Auf Basis der vorliegenden Befunde kann jedoch nicht geklärt werden, ob dieser Effekt für die wiederholte Durchführung desselben IAT spezifisch ist, oder ob er auch bei der seriellen Bearbeitung verschiedener IATs auftritt. Letzteres führt zu einer Reihe von Problemen bei der Interpretation von Daten aus verschiedenen seriell erhobenen IATs, wie sie etwa im Balanced Identity Design (Greenwald et al., 2002) üblich ist. Während solche Reihenfolgeeffekte auf Mittelwertsebene durch experimentelle Techniken - etwa der Balancierung der Testreihenfolge - kontrolliert werden können, stellen sie für die individuelle Interpretation von Ergebnissen des IAT – etwa im Rahmen differentialpsychologischer Fragestellungen - eine kaum lösbare Problematik dar.

Auch bezüglich der psychometrischen Eigenschaften des Blumen-Insekten IATs ergeben sich ähnliche Ergebnisse, wie für den IAT mit geometrischen Objekten in Experiment 9. Der Blumen-Insekten IAT weist eine befriedigende interne Konsistenz zwischen $\alpha = .88$, und $\alpha = .82$ auf und die Retest-Stabilität liegt mit $r = .53$ signifikant über Null. Die Ergebnisse liegen damit im Bereich dessen, was mit inhaltlich angereicherten IATs in anderen Anwendungsfeldern erzielt wird. Obwohl angenommen werden kann, dass den Ergebnissen keine systematischen interindividuellen Unterschiede in der Einstellung zu Insekten und Blumen zugrunde liegen, bleibt die These, dass die vorliegenden Ergebnisse systematische methodenspezifische Varianz widerspiegeln, angreifbar.

7.3 Experiment 11

In Experiment 11 soll methodenspezifische Varianz in dem Blumen-Insekten Einstellungs-IAT mit einem stärkeren Test nachgewiesen werden. Wie in der Einleitung zu diesem Abschnitt der Arbeit schon angesprochen, sollte eine systematische Kontaminierung von IAT-Effekten mit methodenspezifischer Varianz inhaltsunabhängig sein. Es ist also zu erwarten, dass auch IATs ohne jegliche inhaltliche Überlappung einen statistisch nachweisbaren Zusammenhang aufweisen. Um diese Vorhersage zu prüfen, bearbeiteten die Versuchs-

teilnehmer in direkter Folge zwei IATs, einerseits den Blumen-Insekten-IAT aus Experiment 10 und andererseits dem IAT mit geometrischen Objekten aus Experiment 9. Da der Geometrie-IAT offenkundig nicht die Präferenz von Blumen über Insekten erfasst, kann eine Korrelation zwischen den Maßen nicht durch inhaltsbedingte gemeinsame Varianz erklärt werden. Eine positive und signifikante Korrelation zwischen beiden IATs kann also als direkter Indikator für das Ausmaß methodenspezifischer Varianz interpretiert werden.

Neben dem Fehlen einer inhaltlichen Überlappung unterscheidet sich auch die Art der jeweils erfassten Relation zwischen Zielkonzepten und Attributen. Während der Blumen-Insekten-IAT auf einer aufgabenirrelevanten Assoziation von Blumen und Insekten mit positiver und negativer Valenz basiert, beruht der IAT mit geometrischen Objekten auf direkt sichtbaren aufgabenirrelevanten Eigenschaften der dargebotenen Objekte. Abgesehen von der identischen experimentellen Anordnung unterscheiden sich beide Maße also maximal. In dieser Hinsicht ist besonders interessant, ob die Verringerung der IAT-Effekte bei wiederholter Messung auch in diesem Fall auftritt. Es kann also zusätzlich geklärt werden, ob die in den Experimenten 9 und 10 aufgetretene Verringerung der Effekte am zweiten Messzeitpunkt auf die Bearbeitung eines zweiten IATs mit unterschiedlichem Inhalt generalisiert.

Methode

Die Versuchsteilnehmer von Experiment 11 bearbeiteten einen IAT mit geometrischen Objekten und einen herkömmlichen Blumen-Insekten Einstellungs-IAT. Prozedur und Design entsprechen den Experimenten 9 und 10. Die Reihenfolge, in der die IATs zu bearbeiten waren, wurde über die Teilnehmer hinweg ausbalanciert; eine Hälfte der Versuchsteilnehmer bearbeitete zuerst den Blumen-Insekten-IAT und dann den IAT mit geometrischen Objekten, die verbleibenden Versuchsteilnehmer bearbeiteten die Maße in umgekehrter Reihenfolge. Ferner wurden Kompatibilitätsreihenfolge und Tastenzuordnung wie in den Experimenten 9 und 10 ausbalanciert.

Stichprobe

An der Untersuchung nahmen 67 Personen teil. Zwanzig Teilnehmer waren männlichen und 47 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 19 und 44 Jahren ($M = 25.67$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Als Material dienten die in den Experimenten 5, 6, 7, 8 und 10 eingesetzten Wörter für Blumen und Insekten, sowie die dort verwendeten positiven und negativen Substantive. In dem IAT mit geometrischen Objekten wurden die gleichen Reize verwendet wie in den Experimenten 2 und 9. Die Kovariation von Größe und Farbe der Objekte blieb wie in Experiment 9 über alle Versuchsbedingungen hinweg gleich.

Ergebnisse

Für den IAT mit geometrischen Objekten wurden Antwortlatenzen unter 100 ms zu 100 ms (0.2 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.4 % der Durchgänge) rekodiert. Für den Blumen-Insekten-IAT wurden Antwortlatenzen unter 300 ms zu 300 ms (0.1 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (1.1 % der Durchgänge) rekodiert. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (insgesamt 7.7 % der Durchgänge) nicht in die Analyse aufgenommen. Die mittleren IAT-Scores lagen für den Blumen-Insekten-IAT bei 191 ms (SD = 121) und für den IAT mit geometrischen Objekten bei 366 ms (SD = 143).

Psychometrische Analyse

Die interne Konsistenz wurde wie in den vorigen Experimenten über vier zufällig zusammengestellte Subtests bestimmt. Sie lag für den Blumen-Insekten-IAT bei $\alpha = .88$, und bei $\alpha = .93$ für den IAT mit geometrischen Objekten. Ein statistisch bedeutsamer Zusammenhang zwischen diesen - auf substantieller Ebene vollkommen unabhängigen - IATs von $r = .39$, $p < .01$, bestätigt die Vorhersage der Hypothese einer systematischen Kontaminierung beider IATs mit methodenspezifischer Varianz.

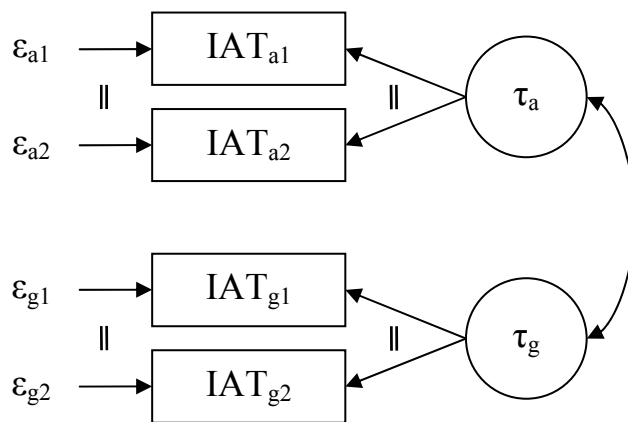


Abbildung 14: Strukturgleichungsmodell für die Analyse methodenspezifischer Varianz. IAT_{a1} und IAT_{a2} entsprechen den Testhälften des Blumen-Insekten-IAT, und IAT_{g1} und IAT_{g2} denen des IAT mit geometrischen Objekten. Die Residuen der manifesten Variablen sind mit ϵ_{a1} , ϵ_{a2} , ϵ_{g1} , ϵ_{g2} bezeichnet. τ_a entspricht der gemeinsamen Varianz beider Testhälften des Blumen-Insekten-IATs. τ_g entspricht der gemeinsamen Varianz der Testhälften des Geometrie-IAT. Die Faktorladungen und Residuen sind innerhalb der IATs gleichgesetzt.

Um die Ergebnisse dieser korrelativen Analyse abzusichern, wurde zusätzlich eine lineare Strukturgleichungsanalyse durchgeführt. Dazu wurden die einzelnen Durchgänge der beiden IATs zufällig in zwei Testhälften aufgeteilt. Für jede der Testhälften wurde der resultierende IAT-Effekt berechnet. Das Strukturgleichungsmodell ist in Abbildung 14 dargestellt. Die Kernannahme des Modells ist, dass der beobachteten Kovarianzmatrix zwei distinkte, aber korrelierte latente Variablen (τ_a und τ_g) zugrunde liegen. Die manifesten Variablen IAT_{a1} und IAT_{a2} beziehen sich auf die Testhälften des Blumen-Insekten-IAT. Die manifesten Variablen IAT_{g1} und IAT_{g2} entsprechen den Testhälften des IAT mit geometrischen Objekten. Residuen dieser Variablen sind mit ϵ_{a1} , ϵ_{a2} , ϵ_{g1} und ϵ_{g2} bezeichnet. Die Faktorladungen der Testhälften wurden jeweils innerhalb eines IATs gleichgesetzt. Die Anpassung dieses Modells an die Daten ergibt, dass das Modell gut mit den Daten vereinbar ist, $\chi^2(5) = 6.47$, $p = .26$. Wie aufgrund der hohen internen Konsistenzen zu erwarten, ist ein Großteil der Varianz innerhalb der Testhälften eines IATs geteilte Varianz. Die Korrelation der latenten Variablen liegt bei .42 mit einem Standardfehler der Schätzung von .11. Wird dieser Zusammenhang auf null gesetzt, ergibt sich ein signifikanter Verlust der Modell-

passung, $t(66) = 3.74, p < .01$. Die restlichen Parameterschätzer können Abbildung 15 entnommen werden. Die zugrunde liegende Kovarianzmatrix ist in Tabelle 3 aufgeführt.

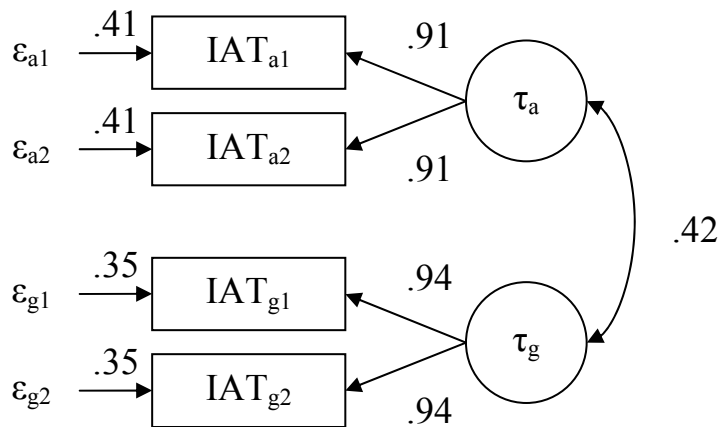


Abbildung 15: Ergebnisse der Strukturgleichungsanalyse von Experiment 11. Dabei entsprechen IAT_{a1} und IAT_{a2} den Testhälften des Blumen-Insekten-IAT, und IAT_{g1} und IAT_{g2} denen des IAT mit geometrischen Objekten. Die Residuen dieser manifesten Variablen sind mit ϵ_{a1} , ϵ_{a2} , ϵ_{g1} , ϵ_{g2} bezeichnet. τ_a entspricht der gemeinsamen Varianz beider Testhälften des Blumen-Insekten IATs. τ_g entspricht der gemeinsamen Varianz der Testhälften des Geometrie-IAT.

Varianzanalyse

Die Antwortlatenzen des Blumen-Insekten-IATs wurden mit einer 2 (Kompatibilitätsreihenfolge) x 2 (erster versus zweiter Messzeitpunkt) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf den letzten beiden Faktoren ausgewertet. Der IAT-Effekt ist statistisch bedeutsam, wie ein Haupteffekt des Faktors Kompatibilität zeigt, $F(1, 63) = 166.45, p < .01$. Wie in den Vorgängerexperimenten finden sich ferner signifikante Aufgabenwechselkosten, $F(1, 63) = 224.83, p < .01$, und eine Interaktion zwischen IAT-Effekt und Aufgabenwechselkosten, $F(1, 63) = 88.43, p < .01$. Die Antwortlatenzen sind insgesamt kürzer, wenn die inkompatible Phase zuerst bearbeitet wurde, $F(1, 63) = 5.28, p < .05$, wie ein Haupteffekt der Kompatibilitätsreihenfolge zeigt. Gleichzeitig fallen in dieser Bedingung die IAT-Effekte ebenso geringer aus, $F(1, 63) = 6.09, p < .05$, wie die Asymmetrie der Aufgabenwechselkosten, $F(1, 63) = 10.41, p < .01$. Die

Asymmetrie ist größer, wenn zuerst die kompatible Phase bearbeitet wird. Ob der Blumen-Insekten-IAT an erster oder zweiter Stelle im Experiment bearbeitet wird, hat weder Einfluss auf die Gesamtlatenz noch auf die Größe des IAT-Effekts, beide $F(1, 63) < 1$, n.s.

Tabelle 3: Varianzen (unterstrichen) und Kovarianzen der zwei zufällig zusammengestellten Testhälften des Geometrie-IAT (IAT_{g1} und IAT_{g2}) und des Blumen-Insekten-IAT (IAT_{a1} und IAT_{a2}) in Experiment 11

	IAT _{a1}	IAT _{a2}	IAT _{g1}	IAT _{g2}
IAT _{a1}	<u>17233.25</u>	13227.18	8523.04	6503.85
IAT _{a2}		<u>14797.33</u>	6859.61	4904.25
IAT _{g1}			<u>23716.19</u>	18878.98
IAT _{g2}				<u>19526.93</u>

Der IAT mit geometrischen Objekten wurde mit einer strukturell analogen Varianzanalyse ausgewertet. Der IAT-Effekt ist erneut statistisch bedeutsam, $F(1, 63) = 489.45$, $p < .01$, ebenso die Kosten bei Aufgabenwechselfurchgängen, $F(1, 63) = 225.93$, $p < .01$, und die Asymmetrie der Aufgabenwechselkosten über die Phasen des IAT, $F(1, 63) = 157.52$, $p < .01$. Die Interaktion zwischen dem IAT-Effekt und der Position des IAT im Experiment erreicht ebenfalls das Signifikanzniveau, $F(1, 63) = 5.19$, $p < .05$. Wie in den Vorgängerexperimenten ist der Effekt im zweiten IAT kleiner. Eine Interaktion von Kompatibilitätsreihenfolge und IAT-Effekt verfehlt das Signifikanzniveau knapp, $F(1, 63) = 3.37$, $p = .07$. Deskriptiv ist der IAT-Effekt kleiner, wenn die inkompatible Phase zuerst bearbeitet wurde.

Diskussion

Das entscheidende Ergebnis von Experiment 11 ist der signifikante Zusammenhang zweier IAT-Maße, für die eine inhaltliche Überlappung ausgeschlossen werden kann. Dieser zeigt sich in Form einer direkten Korrelation der beobachteten IAT-Scores, aber auch der latenten Variablen in der Strukturgleichungsanalyse. Dies Ergebnis deutet darauf hin, dass sowohl der Blumen-Insekten-IAT als auch der IAT mit geometrischen Objekten mit methodenspezifischer Varianz kontaminiert ist. Obwohl die Korrelation der Maße kleiner ausfällt als die Stabilitätsschätzungen in den Experimenten 9 und 10, liegt sie doch deutlich über null. Dies deutet darauf hin, dass die meist befriedigenden Schätzungen von Reliabilität und interner Konsistenz von IATs nur teilweise auf merkmalspezifische Varianz zurückgehen, während ein bedeutsamer weiterer Varianzanteil systematische methodenspezifische

Varianz widerspiegelt. In diesem Licht muss die üblicherweise ins Feld geführte Überlegenheit des IAT bezüglich dieser psychometrischen Kriterien neu bewertet werden: Teilweise geht sie offenkundig auf eine niedrige diskriminante Validität zurück.

Uneindeutig fallen die Befunde zur Testwiederholung aus. Während die serielle Position des IAT mit geometrischen Objekten einen deutlichen Effekt auf das Ausmaß des IAT-Effekts hat, findet sich kein analoger Effekt für den Blumen-Insekten-IAT. Dies Befundmuster könnte einerseits darauf zurückgehen, dass Nachwirkungen des Blumen-Insekten-IAT auf eine nachfolgende Bearbeitung anderer IATs generalisieren, während der IAT mit geometrischen Objekten lediglich spezifische Nachwirkungen hat. Ebenso denkbar ist jedoch, dass der Blumen-Insekten-IAT gegen Nachwirkungen immun ist, nicht jedoch der Geometrie-IAT. Zusammengenommen gilt über die drei Experimente hinweg, dass die mittleren Effekte des Blumen-Insekten-IAT an zweiter Position nur dann verringert waren, wenn vorher der gleiche IAT bearbeitet wurde, während die mittleren Effekte des IAT mit geometrischen Objekten immer dann verringert waren, wenn vorher ein IAT bearbeitet wurde, unabhängig von dessen Inhalt. Anders formuliert traten nach der Bearbeitung des Blumen-Insekten-IATs immer verringerte IAT-Effekte auf, während nach der Bearbeitung des IAT mit geometrischen Objekten lediglich dann verringerte Effekte auftraten, wenn der gleiche IAT wiederholt wurde.

7.4 Experiment 12

Ein entscheidender Unterschied zwischen dem Blumen-Insekten-IAT sowie dem IAT mit geometrischen Objekten auf der einen Seite, und den IATs, die in Anwendungskontexten verwendet werden, auf der anderen Seite ist, dass erstere nahezu keine interindividuellen Unterschiede in der Richtung der IAT-Effekte aufweisen. So zeigt der Blumen-Insekten-IAT aus Experiment 11 für fast alle Teilnehmer eine Präferenz von Blumen über Insekten. Es ist denkbar, dass der darin sichtbare Mangel an inhaltlich begründeter Varianz eine notwendige Bedingung für die systematische Kontaminierung von IAT-Scores mit methodenspezifischer Varianz darstellt. Experiment 12 sollte deshalb den zentralen Befund von Experiment 11 mit einem IAT replizieren, für den bedeutsame interindividuelle Unterschiede in der Richtung und Ausprägung der Effekte zu erwarten sind. Ferner sollte geprüft werden, welchen Effekt reliable methodenspezifische Varianz auf die Zusammenhänge zwischen dem IAT und expliziten Maßen hat.

Methode

Die Teilnehmer von Experiment 12 bearbeiteten einen Selbst-Andere Extraversions-IAT, der weiter unten genauer beschrieben wird, gefolgt von dem IAT mit geometrischen Objekten, wie er schon in den Vorgängeruntersuchungen verwendet wurde. Wie für den Blumen-Insekten-IAT aus Experiment 11 wird angenommen, dass der Extraversions-IAT mit reliabler methodenspezifischer Varianz kontaminiert ist, die sich in Form einer bedeutsamen Korrelation dieses Maßes mit dem Geometrie-IAT zeigen sollte. Auch in diesem Fall kann eine Vermittlung eventueller Zusammenhänge zwischen den IATs kaum durch ein gemeinsames inhaltliches Konstrukt erklärt werden, das von beiden IATs erfasst wird und inter-individuelle Unterschiede in den Scores bedingt.

Im Gegensatz zu Experiment 11 kann diese Hypothese jedoch nicht direkt an den IAT-Scores geprüft werden, denn die Vorhersage bezieht sich auf die Größe der Effekte, nicht auf ihre Richtung, für die ja systematische inhaltsabhängige Unterschiede erwartet werden. In den direkt gemessenen Effekten des Extraversions-IAT kann methodenspezifische Varianz durch eine Variation der Effektrichtung verdeckt werden, wie das folgende Beispiel nochmals verdeutlichen soll: Angenommen, mehrere Versuchsteilnehmer mit einer vergleichsweise geringen Performanz bei Aufgabenwechseldurchgängen bearbeiten beide IATs. Wenn dieser IAT systematisch mit Aufgabenwechselkosten kontaminiert ist, ist zu erwarten, dass die IAT-Scores beider Teilnehmer in dem IAT mit geometrischen Objekten stark positiv ausfallen. Sind die Scores in dem Extraversions-IAT ebenfalls reliabel kontaminiert, so sind ebenfalls größere Effekte zu erwarten, sie sollten jedoch für extravertierte und introvertierte Versuchsteilnehmer in unterschiedlicher Richtung ausfallen. Die unterschiedliche Richtung der Effekte könnte nun den Einfluss methodenspezifischer Varianz auf die Scores verdecken. Um den Effekt methodenspezifischer Varianz trotzdem nachweisen zu können, beziehen sich die entsprechenden Tests auf die absolute Größe, also den Betrag der IAT-Effekte.

Stichprobe

An der Untersuchung nahmen 81 Personen teil. 21 Teilnehmer waren männlichen und 60 weiblichen Geschlechts. Das Alter der Teilnehmer lag zwischen 19 und 60 Jahren ($M = 26.14$). Die Teilnehmer erhielten eine Aufwandsentschädigung in Höhe von DM 10,- (ca. € 5,-) oder eine Bescheinigung über eine Versuchspersonenstunde gemäß der Prüfungsordnung des Psychologischen Instituts der Universität Bonn.

Material

Auf Basis einer Wortnormierung von Ostendorf (1994) wurden zunächst diejenigen 130 Adjektive ausgewählt, die die höchste Typikalität für Extra- bzw. Introversion aufwiesen. Daraufhin wurden vier unabhängige Urteiler gebeten, die Valenz dieser Adjektive auf einer 5-stufigen Skala, die von sehr positiv bis sehr negativ reichte, zu bewerten. Das Valenzrating ergab, dass die für Extraversion typischen Adjektive im Durchschnitt positiver bewertet wurden als die für Introversion typischen. Es zeigte sich, dass die Herstellung zweier bezüglich der Evaluation vergleichbarer Stimuluslisten zu einem erheblichen Verlust an Typikalität führte. Da es wenig sinnvoll erschien, ein Konzept mit atypischen, jedoch evaluativ unkonfundierten Reizen zu repräsentieren, wurde der Typikalität der Reize für Extra- bzw. Introversion die höchste Priorität eingeräumt. Einige extrem positive und negative Adjektive konnten jedoch ohne einen Verlust an Typikalität aus dem Pool entfernt werden. Um Wortlängeneffekte zu vermeiden, wurden ferner solche Adjektive mit einer Länge von mehr als elf Zeichen aus dem Pool entfernt. Aus den verbleibenden Adjektiven wurden jeweils diejenigen 14 Adjektive ausgewählt, die für Intro- bzw. Extraversion die höchsten Typikalitätsscores erreicht hatten.

Die Attributdimension Selbst versus Andere wurde mit jeweils fünf Wörtern repräsentiert, die sich entweder auf das Selbst bezogen (ich, selbst, mich, mir, eigen) oder auf nicht weiter spezifizierte andere Personen (sie, ihnen, euer, ihr, andere). Diese Reize wurden der Untersuchung von Asendorpf et al. (2002) entnommen. Alle verwendeten Stimuluswörter sind in Anhang C aufgeführt.

Als explizites Maß der Extraversion wurde das Neo-Fünf-Faktoren Inventar (NEO-FFI, Borkenau & Ostendorf, 1993) verwendet, eine deutsche Übersetzung des NEO Five-Factor Inventory (Costa & McCrae, 1992; McCrae & Costa, 1989). Der NEO-FFI wurde in der Papier-und-Bleistift-Fassung angewendet.

Prozedur

Experiment 12 umfasste vier Phasen. In der ersten Phase bearbeiteten die Versuchsteilnehmer den NEO-FFI (Borkenau & Ostendorf, 1993) als explizites Maß der Extraversion. Die Versuchsteilnehmer bearbeiteten das gesamte Inventar, da die isolierte Bearbeitung der Extraversionsfragen zu einer Verfälschung der Ergebnisse führen könnte. Die Bearbeitung des NEO-FFI nahm etwa 10 Minuten in Anspruch.

In der zweiten Phase wurde der Extraversions-IAT bearbeitet. In diesem IAT bestand die Attributaufgabe darin, dargebotene Adjektive als entweder typisch extravertiert oder als typisch introvertiert zu klassifizieren. Diese Aufgabe wurde zunächst in einem Aufgabenblock mit 48 Durchgängen eingeübt. In der Instruktion zu dieser Aufgabe wurde den Teilnehmern eine Definition der Begriffe gegeben, die auf der Instruktion zu der Wortnormierungsstudie von Ostendorf (1994) beruhte. Die Zielkonzeptaufgabe erforderte die Klassifikation dargebotener Wörter mit den Antwortkategorien selbst und andere. In der dritten und fünften Phase des IAT wurden die Aufgaben in jeweils 3 Blöcken gemischt bearbeitet. In der vierten Phase wurde die Tastenzuordnung der Zielkonzeptaufgabe umgekehrt und in einem ungemischtem Block eingeübt. Wie in den Vorgängeruntersuchungen wurde die Reihenfolge der Aufgaben innerhalb eines Blocks, ebenso wie die Reihenfolge der einzelnen Reize zufällig bestimmt. Die Tastenzuordnung in den gemischten Phasen erfolgte so, dass entweder Extraversion und „Andere“ sowie Introversion und „Selbst“ jeweils derselben Taste zugeordnet waren, oder umgekehrt. Die Reihenfolge, in der diese Bedingungen im Experiment zu bearbeiten waren, wurde zwischen den Versuchsteilnehmern ausbalanciert. Das RSI betrug in beiden IATs 800 ms.

Im Anschluss an den Extraversions-IAT bearbeiteten die Teilnehmer den IAT mit geometrischen Objekten, der unter anderem in Experiment 11 verwendet wurde. Die Versuchsteilnehmer bearbeiteten jeweils 3 Aufgabenblöcke in den beiden Phasen mit gemischten Aufgaben und jeweils einen Aufgabenblock in den drei Übungsphasen. Abgesehen von dieser geringeren Zahl von Aufgabenblöcken entspricht die Prozedur des IAT mit geometrischen Objekten derjenigen von Experiment 11. Da für die Kompatibilitätsreihenfolge für den Extraversions-IAT von der Ausprägung des Merkmals Extraversion abhängt und folglich nicht experimentell kontrolliert werden kann, wurde die Kompatibilitätsreihenfolge für den Geometrie-IAT nicht variiert. Alle Versuchsteilnehmer bearbeiteten zunächst die kompatible Phase dieses IAT.

In der vierten und letzten Phase des Experiments wurden die Versuchsteilnehmer ausführlich über ihr persönliches Abschneiden im NEO-FFI unterrichtet und konnten sich über die Zielsetzung des Experiments informieren.

Ergebnisse

Für den IAT mit geometrischen Objekten wurden Antwortlatenzen unter 100 ms zu 100 ms (0.1 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.8 % der Durchgänge)

rekodiert. Für den Extraversions-IAT wurden Antwortlatenzen unter 300 ms zu 300 ms (0.1 % der Durchgänge) und Latenzen über 3000 ms zu 3000 ms (0.9 % der Durchgänge) rekodiert. Ferner wurden die ersten beiden Durchgänge eines Aufgabenblocks und Durchgänge mit Antwortfehlern (6.5 % der Durchgänge) nicht in die Analyse aufgenommen.

Zur Berechnung der Effekte im IAT mit geometrischen Objekten wurden die Antwortlatenzen in der kompatiblen Phase des IAT von den Latenzen in der inkompatiblen Phase subtrahiert. Die Scores des Extraversions-IAT wurden hingegen so berechnet, dass positive Scores als Hinweis auf eine stärkere Assoziation zwischen Selbst und Extraversion (bzw. Andere und Introversion) interpretiert werden können: Die Antwortlatenzen in der Bedingung, in der Extraversion und Selbst einer Taste zugeordnet waren, wurden von den Latenzen in der Bedingung mit Introversion und Selbst auf der gleichen Antworttaste abgezogen.

Für die Stichprobe von Experiment 12 lag der mittlere Extraversions-Score im NEO-FFI bei 2.47 (SD = .56), und ist damit vergleichbar mit dem Wert von 2.36 (SD = .57), der in der Normierung des NEO-FFI berichtet wird (Borkenau & Ostendorf, 1993). Die mittleren IAT-Effekte lagen bei 369 ms (SD = 165) für den IAT mit geometrischen Objekten, und bei -41 ms (SD = 205) für den Extraversions-IAT. Die Effekte des Extraversions-IAT lagen zwischen -768 ms und 446 ms, der Median bei -7 ms. Folgt man den Ergebnissen des IAT, bestehen also für eine Hälfte der Stichprobe stärkere Assoziationen zwischen Selbst und Extraversion beziehungsweise Andere und Introversion, während für die andere Hälfte der Stichprobe das Gegenteil zutrifft. Die Tatsache, dass der mittlere IAT-Effekt leicht im negativen Bereich liegt, weist, wie auch die Spannweite, darauf hin, dass letzteres Muster im Schnitt zu größeren Effekten führte.

Psychometrische Analyse

Cronbachs α wurde erneut über die IAT-Effekte von vier Subtests berechnet, die für jeden IAT zufällig zusammengestellt wurden. Für den IAT mit geometrischen Objekten ergab sich ein α von .93, und für den Extraversions-IAT lag die interne Konsistenz bei .94. Die konvergente Validität des Extraversions-IAT wurde über die Korrelation der vorzeichenbehafteten IAT-Effekte mit dem NEO-FFI bestimmt. Dabei ergibt sich eine kleine, aber statistisch bedeutsame Korrelation von $r = .24$, $p < .05$, die anzeigt, dass implizit und explizit gemessene Extraversion teilweise gemeinsame Varianz aufweisen. Wie erwartet war die

Korrelation zwischen der Extraversionsskala des NEO-FFI und den Effekten des IAT mit geometrischen Objekten praktisch null und lag bei $r = -.03$, n.s.

Methodenspezifische Varianz

Das Vorhandensein einer reliablen methodenspezifischen Varianzkomponente des Extraversions-IAT wurde mittels der Korrelation dieses IAT mit dem Geometrie-IAT geprüft. Wie oben erläutert kann dieser Test nicht direkt durchgeführt werden, da methodenspezifische Varianzanteile des Extraversions-IAT durch die Variation in der Richtung des IAT verdeckt werden, über die keine spezifische Hypothese formuliert wurde. Die im Folgenden berichteten Analysen basieren deswegen auf den Absolutwerten der Effekte in diesem IAT.

Die Korrelation zwischen den Absolutwerten des Extraversions-IAT und Effekten des IAT mit geometrischen Objekten ist mit $r = .39$, $p < .01$, signifikant größer null. Dieser Befund kann offenbar nicht durch Performanzunterschiede zwischen introvertierten und extravertierten Versuchsteilnehmern erklärt werden. Eine solche Erklärung wäre nur dann zutreffend, wenn der Analyse die vorzeichenbehafteten Effekte des Extraversions-IAT zugrunde lägen. Weiterhin sollte dann auch die Korrelation zwischen NEO-FFI und dem Geometrie-IAT signifikant sein. Da beides nicht der Fall ist, handelt es sich offenkundig um einen methodenspezifischen, aber reliablen Varianzanteil in beiden IATs. Versuchsteilnehmer, für die sich in dem IAT mit geometrischen Objekten große Effekte ergeben, produzieren also - unabhängig von der Richtung der Effekte – auch im Extraversions-IAT größere Effekte.

Tabelle 4: Varianzen (unterstrichen) und Kovarianzen der zwei zufällig zusammengestellten Testhälften des Geometrie-IAT (IAT_{g1} und IAT_{g2}) und des Extraversions-IAT (IAT_{e1} und IAT_{e2}) in Experiment 12

	IAT_{e1}	IAT_{e2}	IAT_{g1}	IAT_{g2}
IAT_{e1}	<u>21463.23</u>	20440.06	8289.94	9795.75
IAT_{e2}		<u>26971.34</u>	8964.79	11946.67
IAT_{g1}			<u>28657.55</u>	25187.76
IAT_{g2}				<u>30347.98</u>

Um zu prüfen, ob sich die Befunde auf der Ebene direkt beobachteter Variablen auch finden, wenn die Messfehlerkomponente beider IATs kontrolliert wird, wurde erneut eine Strukturgleichungsanalyse durchgeführt. Dazu wurden die IAT-Effekte für zwei zufällig

zusammengestellte Testhälften beider IATs berechnet. Für den Extraversions-IAT wurden der Analyse die Absolutwerte der Effekte zugrunde gelegt. Die resultierende Kovarianzmatrix zwischen den IAT-Scores kann Tabelle 4 entnommen werden. Zur Analyse dieser Kovarianzmatrix wurde erneut das Strukturgleichungsmodell aus Abbildung 14 verwendet. Der einzige Unterschied zu der Strukturgleichungsanalyse von Experiment 11 besteht darin, dass statt des Blumen-Insekten-IAT die absoluten Effekte des Extraversions-IAT verwendet wurden.

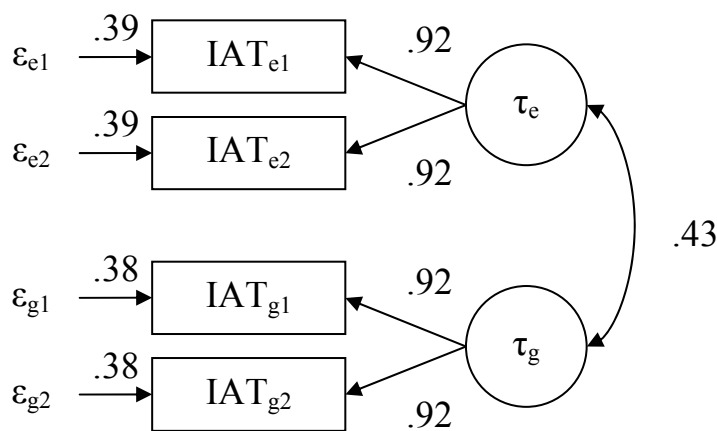


Abbildung 16: Ergebnisse der Strukturgleichungsanalyse von Experiment 12. Dabei entsprechen IAT_{e1} und IAT_{e2} den Testhälften des Extraversions-IAT, und IAT_{g1} und IAT_{g2} denen des IAT mit geometrischen Objekten. Die Residuen dieser manifesten Variablen sind mit ϵ_{e1} , ϵ_{e2} , ϵ_{g1} , ϵ_{g2} bezeichnet. τ_e entspricht der gemeinsamen Varianz beider Testhälften des Extraversions-IAT. τ_g entspricht der gemeinsamen Varianz der Testhälften des Geometrie-IAT.

Eine Anpassung dieses Modells an die Daten ergibt, dass das Modell die Daten zufrieden stellend abbildet, $\chi^2(5) = 7.75$, $p = .17$. Die Korrelation zwischen den latenten Variablen liegt bei $.43$ ($SE = .10$), und ist auch in dieser Analyse signifikant von null verschieden, $t(80) = 4.23$, $p < .01$. Die standardisierten Parameterschätzer können Abbildung 16 entnommen werden.

Varianzanalyse

Die Antwortlatenzen des Extraversions-IAT wurden mit einer 2 (Kompatibilitätsreihenfolge) x 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs.

Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf den letzten beiden Faktoren ausgewertet. Um eine Vergleichbarkeit mit den vorherigen Analysen herzustellen, wurden die Variablen post hoc so kodiert, dass die langsamere von beiden gemischten Phasen als inkompatibel kodiert wurde und die schnellere als kompatibel. Der Faktor Kompatibilitätsreihenfolge wurde ebenfalls entsprechend umkodiert. Erwartungsgemäß ist der absolute IAT-Effekt von 136 ms in dieser Analyse statistisch bedeutsam, $F(1, 79) = 54.23, p < .01$. Der mittlere IAT-Effekt ist also, unabhängig von der Richtung der Effekte, von null verschieden. Wie in den Vorgängeruntersuchungen fallen über die Bedingungen hinweg bedeutsame Kosten für den Wechsel zwischen Attribut- und Zielkonzeptaufgabe an, $F(1, 79) = 150.78, p < .01$. Diese Kosten fallen jedoch, wie schon in den bisherigen Untersuchungen, in der langsameren inkompatiblen Phase des Extraversions-IAT deutlich größer aus als in der kompatiblen, $F(1, 79) = 30.00, p < .01$. Ferner findet sich eine Interaktion zwischen Kompatibilitätsreihenfolge, der Größe des IAT-Effekts und dem Aufgabenwechselfaktor, $F(1, 79) = 7.50, p < .01$. Wie schon in Experiment 11 ist die Asymmetrie der Aufgabenwechselkosten größer, wenn die kompatible, also die schnellere beider Phasen, zuerst bearbeitet wurde.

Die Antwortlatenzen des IAT mit geometrischen Objekten wurden mit einer 2 (kompatible vs. inkompatible Tastenzuordnung) x 2 (Wechsel vs. Wiederholung der Aufgabe zwischen Durchgang n und Durchgang n-1) faktoriellen Varianzanalyse mit Messwiederholung auf beiden Faktoren ausgewertet, da die Kompatibilitätsreihenfolge für diesen IAT nicht manipuliert wurde. Dabei ergaben sich wie erwartet ein signifikanter Haupteffekt der Kompatibilität, $F(1, 79) = 409.33, p < .01$, und ein Haupteffekt des Aufgabenwechselfaktors, $F(1, 79) = 242.25, p < .01$. Auch die Zweifach-Interaktion beider Faktoren wurde wie in den Vorgängeruntersuchungen bedeutsam, $F(1, 79) = 179.27, p < .01$. Die Aufgabenwechselkosten fielen in der inkompatiblen Phase des IAT deutlich höher aus.

Diskussion

Die Hauptergebnisse von Experiment 12 sind ein kleiner, aber statistisch bedeutsamer Anteil gemeinsamer Varianz zwischen dem expliziten und dem impliziten Extraversionsmaß, sowie eine bedeutsame Korrelation der Effektgrößen des Extraversions-IAT und des IAT mit geometrischen Objekten. Dieses Befundmuster entspricht dem, was auf der Basis einer reliablen Kontaminierung der IAT-Scores zu erwarten wäre. Dieses Muster weist sowohl auf eine niedrige konvergente als auch eine niedrige diskriminante Validität des IAT hin: Der

Anteil merkmalspezifischer Varianz ist gering, während der Anteil systematischer methodenspezifischer Varianz im mittleren Bereich liegt. In der Terminologie der Multi-Trait-Multi-Method-Analyse ausgedrückt, ist die Mono-Trait-Hetero-Method Korrelation zwischen impliziten und expliziten Maßen der Extraversion unbefriedigend gering, während die Hetero-Trait-Mono-Method Korrelation zwischen dem Extraversions-IAT und dem IAT mit geometrischen Objekten, wie in Experiment 11, im Bereich von .40 liegt.

In diesem Zusammenhang muss jedoch beachtet werden, dass neben interindividuellen Unterschieden in den Aufgabenwechselkosten eine Reihe anderer Faktoren zu der geringen Korrelation zwischen impliziter und expliziter Messung beigetragen haben könnten, etwa die Balancierung der Kompatibilitätsreihenfolge (Asendorpf et al., 2002) oder Unterschiede in den durch implizite und explizite Maße erfassten Konstrukten (Bosson et al., 2000; Wilson et al., 2000). Wie weiter oben schon ausgeführt ist jedoch mehr als zweifelhaft, ob eine Konstanthaltung der Tastenzuordnung zu valideren Ergebnisse führt. Allenthalben könnte die dadurch erzeugte Konfundierung von Merkmalsausprägung und Kompatibilitätsreihenfolge eine systematische Asymmetrie der Ergebnisse hervorrufen, da die Versuchsteilnehmer abhängig von der Ausprägung der erfassten Assoziationen unterschiedliche Kompatibilitätsreihenfolgen bearbeiten.

Nicht ausgeschlossen werden können ferner Effekte prozeduraler Variablen, etwa des vergleichsweise langen RSIs von 800 ms, oder der Bearbeitung von Aufgabenwiederholungsdurchgängen. Angesichts der bisher geschilderten Befunde scheint jedoch unwahrscheinlich, dass diese Variationen des ursprünglichen Designs den geringen Zusammenhang erklären können. Experiment 1 zeigte in diesem Zusammenhang, dass die durch Aufgabenwechsel bedingte Varianz in den IAT-Effekten bei einem RSI von 1000 ms bedeutsam geringer ausfällt als bei einem RSI von 100 ms. Insofern diese Varianz merkmalsunabhängig ist, sollte eine Verlängerung des RSI also zu einer Verringerung des Messfehlers und damit zu einer besseren Validität führen.

Abgesehen davon sollten die IAT-Effekte generell stärker von Aufgabenwechselkosten beeinflusst sein, wenn keine Aufgabenwiederholungsdurchgänge zu bearbeiten sind. Aus dieser Sicht ist zu erwarten, dass der Effekt von Aufgabenwechselkosten auf den IAT stärker ausfällt, wenn die Aufgaben streng alternierend bearbeitet werden, wie es in der Standardprozedur vorgesehen ist (Greenwald et al., 1998). In diesem Zusammenhang sei nochmals betont, dass die alternierende Bearbeitung der Aufgaben keineswegs dazu führt, dass Aufgabenwechselkosten reduziert oder kontrolliert werden. Zwar können die spezifischen

Wechselkosten unter dieser Bedingung nicht bestimmt werden, dies bedeutet jedoch keineswegs, dass sie nicht auftreten. Sofern die durch Aufgabenwechsel bedingte Varianz merkmalsunabhängig ist, sollte eine strikt alternierende Bearbeitung der Aufgaben folglich zu einer Vergrößerung des Messfehlers und damit zu einer geringeren Validität führen.

7.5 Techniken zur Reduktion methodenspezifischer Varianz

Wie bereits erwähnt, hat die reliable Kontamination des IAT mit methodenspezifischer Varianz einige unangenehme Konsequenzen für die psychometrischen Eigenschaften des Verfahrens: Sie erhöht erstens Schätzungen der internen Konsistenz und der Retest-Reliabilität, sie erhöht zweitens die Korrelationen zwischen den (absoluten) Effekten zweier beliebiger IATs, selbst wenn diese keinerlei inhaltlichen Bezug haben, und sie reduziert drittens die konvergente Validität eines IAT. In dieser Hinsicht kann methodenspezifische Varianz zur Erklärung der bekannten Diskrepanz zwischen hoher interner Konsistenz (Steffens & Buchner, 2003) und der eher moderaten Validität vieler IAT-Maße beitragen. Das Auftreten methodenspezifischer Varianz ist durchaus mit der vorliegenden Evidenz für die prädiktive Validität des IAT in Bezug auf Selbstberichte und Verhaltensmaße vereinbar. Es ist jedoch anzunehmen, dass solche Zusammenhänge durch eine Reduzierung methodenspezifischer Varianz gesteigert werden können, wenn auch in geringem Ausmaß.

Statistische Bereinigung methodenspezifischer Varianz

Dass der Effekt jedoch vergleichsweise gering sein sollte, lässt sich leicht zeigen, wenn man betrachtet, wie Korrelation und Kovarianz zwischen implizit und explizit gemessener Extraversion in Experiment 12 beeinflusst werden sollten, wenn ein Indikator für methodenspezifische Varianz, etwa der Geometrie-IAT, aus den Daten des Extraversions-IAT herauspartialisiert wird. Setzt man voraus, dass die Korrelation zwischen dem expliziten Extraversionsmaß und dem Geometrie-IAT null ist, so sollte die *Kovarianz* zwischen dem impliziten und dem expliziten Maß durch das Auspartialisieren methodenspezifischer Varianz nicht verändert werden. Die *Korrelation* zwischen den beiden Extraversionsmaßen sollte im Gegensatz dazu steigen, da das korrigierte implizite Maß einen geringeren Anteil an Fehlervarianz aufweist. Unter den genannten Voraussetzungen kann die Partialkorrelation durch folgende einfache Gleichung geschätzt werden:

$$r_{IE.G} = \frac{\text{cov}_{IE}}{\sqrt{(1 - r_{IG}^2) \text{var}_I \text{var}_E}} = \frac{1}{\sqrt{1 - r_{IG}^2}} r_{IE},$$

wobei I der Score des Extraversions-IAT ist, G der Geometrie-IAT und E der Score des expliziten Extraversionsmaßes. Auf dieser Basis kann ein Faktor berechnet werden, der den Anstieg der Korrelation zwischen dem korrigierten IAT-Score und dem expliziten Maß widerspiegelt. Da der Extraversions-IAT jedoch beträchtliche Varianz in der Richtung des Effekts aufweist, wird der Anteil methodenspezifischer Varianz durch die Korrelation r_{IG} systematisch unterschätzt. Aus den oben ausgeführten Gründen kann jedoch die Korrelation zwischen den Absolutwerten des Extraversions-IAT und dem Geometrie-IAT zugrunde gelegt werden. Wird der in Experiment 12 aufgetretene Wert dieser Korrelation in die Gleichung eingesetzt, ergibt sich eine moderate Steigerung der Korrelation zwischen den bereinigten Scores des Extraversions-IAT und dem expliziten Maß um den Faktor 1.09.

Um diese Analyse empirisch nachzuvollziehen, wurden die Absolutwerte der Scores des Extraversions-IAT auf die Effekte des Geometrie-IAT regrediert. In den resultierenden Regressionsresiduen sollte der Anteil methodenspezifischer Varianz null sein. Um die ursprüngliche Vorzeicheninformation der Scores wiederherzustellen, wurden die Residuen durch Addition einer Konstanten so verschoben, dass das kleinste Residuum bei null lag, anschließend wurden die Werte mit plus oder minus eins multipliziert, abhängig vom Vorzeichen der ursprünglichen Extraversionseffekte². Die Korrelation zwischen den so bereinigten Scores des Extraversions-IAT und der Extraversionsskala des NEO-FFI liegt etwas höher als die Korrelation mit unbereinigten IAT-Scores, nämlich bei $r = .27, p < .01$. Die Entfernung gemeinsamer Varianz des Extraversions- und des Geometrie-IAT führt also wie erwartet nicht zu einer Verringerung der konvergenten Validität; der beobachtete Anstieg ist jedoch erwartungsgemäß auch nicht statistisch bedeutsam.

Kontrolle methodenspezifischer Varianz durch Kodierung und Selektion von Durchgängen

Die kürzlich vorgestellten Optimierungen der Berechnungsvorschrift des IAT-Effekts (Greenwald et al., 2003) stellen eine weitere Methode zur Verringerung methodenspezifischer Varianz dar. Wie schon angesprochen wurde, verringern diese Optimierungen eine Reihe

² Diese Verschiebung der Residuen stellt sicher, dass die ursprüngliche Vorzeicheninformation wieder hergestellt wird. Verschiebe man die Werte nicht, so hinge das Vorzeichen der bereinigten Scores davon ab, ob die auf Basis des Geometrie-IAT vorhergesagten Effekte niedriger oder höher ausfallen als die empirisch aufgetreten, was offenkundig zu unsinnigen Ergebnissen führen würde.

bekannter Verzerrungen des IAT-Effekts, etwa die Korrelation des IAT-Effekts mit der mittleren Antwortlatenz, und sollen zudem höhere Korrelationen zwischen expliziten und IAT-basierten Maßen zur Folge haben. Ein Vorteil dieser Herangehensweise ist, dass kein spezifischer Indikator methodenspezifischer Varianz benötigt wird. Im Folgenden sollen die Ergebnisse dieses optimierten Algorithmus mit zwei konventionellen Verfahren zur Berechnung des IAT-Effekts verglichen werden.

Der erste konventionelle Algorithmus beruht auf der Berechnung der Differenz mittlerer untransformierter Antwortlatenzen, wobei jeweils die ersten zwei Durchgänge eines Aufgabenblocks von der Auswertung ausgenommen werden, sowie Latenzen unter 300 ms zu 300 rekodiert werden und Latenzen über 3000 ms zu 3000 (IAT_{RT}). Der zweite konventionelle Algorithmus ist der von Greenwald et al. (1998) verwendete, der zusätzlich eine Logarithmierung der Antwortlatenzen vor der Mittelbildung vorsieht (IAT_{LN}). Als dritter Algorithmus wurde der von Greenwald et al. (2003) favorisierte optimierte Algorithmus angewendet, in welchem die IAT-Effekte für untransformierte Antwortlatenzen berechnet und an den individuellen Standardabweichungen relativiert werden. Da die Ergebnisse mit den verschiedenen Variationen dieses Maßes nahezu identisch ausfallen, werden nur die Ergebnisse der einfachsten Variation berichtet, die weder eine spezifische Behandlung von Durchgängen mit Antwortfehlern noch eine Rekodierung der Latenzen am unteren Ende der Verteilung beinhaltet (IAT_{DI}).

Ein weiterer theoretisch motivierter Ansatz zu Reduktion methodenspezifischer Varianz kann aus dem Aufgabenwechselmodell abgeleitet werden. Folgt man der Vorhersage dieses Modells, so sollten Aufgabenwiederholungsdurchgänge in geringerem Maße von systematischen interindividuellen Unterschieden in der Aufgabenwechselperformanz beeinflusst sein als Aufgabenwechseldurchgänge. Zwar lässt der Trägheitsmechanismus erwarten, dass sich Aufgabenwechselkosten, wie es in den Experimenten 5 und 6 demonstriert wurde, über mehrere aufeinander folgende Durchgänge hinweg auswirken. Die signifikante Asymmetrie der spezifischen Aufgabenwechselkosten zeigt jedoch, dass diese Kosten schon nach einmaliger Wiederholung der Aufgabe reduziert sind. Dies legt nahe, dass IAT-Effekte, die ausschließlich auf Basis der Aufgabenwiederholungsdurchgänge berechnet wurden, ein geringeres Ausmaß methodenspezifischer Varianz aufweisen sollten als IAT-Effekte, die auf Basis aller Durchgänge berechnet werden, oder als IAT-Effekte, die auf Basis der Aufgabenwechseldurchgänge berechnet werden. Gleichzeitig sollten die IAT-Effekte mit Aufgaben-

wiederholungsdurchgängen eine höhere konvergente Validität aufweisen, sofern diese Selektion zu einer Verringerung der Fehlervarianz führt.

Um diese Vorhersagen, ebenso wie den Effekt der verschiedenen Scoring-Algorithmen zu prüfen, wurden die Scores des Blumen-Insekten-IAT aus Experiment 11 und die des Extraversions-IAT aus Experiment 12 mit den genannten Algorithmen berechnet und anschließend mit dem Geometrie-IAT als Indikator methodenspezifischer Varianz korreliert. Weil der Anteil methodenspezifischer Varianz in diesem Indikator maximal sein sollte, wurden nur die Aufgabenwechseldurchgänge des Geometrie-IAT in die Auswertung aufgenommen. Die Effekte dieses IAT wurden auf Basis untransformierter Antwortlatenzen berechnet, wie im Ergebnisteil der Experimente 11 und 12 geschildert. Die konvergente Validität wurde für die Daten von Experiment 12 durch die Korrelation des Extraversions-IAT mit der entsprechenden Subskala des NEO-FFI (Borkenau & Ostendorf, 1993) geschätzt.

Die Ergebnisse dieser Analysen sind in Tabelle 5 aufgeführt. Das hervorstechende Ergebnis der Analysen ist, dass der Anteil reliabler methodenspezifischer Varianz in den IAT-Scores durch die Verwendung des optimierten Algorithmus (Greenwald et al., 2003) deutlich reduziert wird. Die methodenspezifische Varianz des Blumen-Insekten-IAT ist nahezu vollständig aus den optimierten Scores entwichen, die methodenspezifische Varianz des Extraversions-IAT ist deutlich reduziert. Weiterhin ergibt sich kein Unterschied im Ausmaß methodenspezifischer Varianz zwischen den Scores, die auf Basis der Aufgabenwechseldurchgänge, und denen, die auf Basis der Aufgabenwiederholungsdurchgänge, berechnet sind. Letzteres spricht zusätzlich dafür, dass der optimierte Algorithmus zu einer Reduktion methodenspezifischer Varianz in den Ergebnissen führt.

Tabelle 5: Korrelationen der unterschiedlich berechneten IAT-Scores des Blumen-Insekten-IAT (Exp. 11) und des Extraversions-IAT (Exp. 12) mit dem Geometrie-IAT als Indikator methodenspezifischer Varianz bzw. mit der Extraversions-Skala des NEO-FFI (konvergente Validität).

		Methodenspezifische Varianz		Konvergente Validität
		Exp. 11	Exp. 12	Exp. 12
Aufgabenwechsel- durchgänge	IAT _{RT}	.40**	.37**	.21
	IAT _{LN}	.31*	.29**	.26*
	IAT _{D1}	.08	.12	.25*
Alle Durchgänge	IAT _{RT}	.33**	.34**	.26*
	IAT _{LN}	.21	.26*	.30**
	IAT _{D1}	-.02	.16	.28*
Aufgabenwiederholungs- durchgänge	IAT _{RT}	.13	.23*	.32**
	IAT _{LN}	.00	.18	.34**
	IAT _{D1}	-.12	.06	.28*

Anmerkung. IAT_{RT} bezieht sich auf IAT-Scores mit untransformierten Antwortlatenzen. IAT_{LN} bezieht sich auf die Scores, die sich mit log-transformierten Antwortlatenzen ergeben. IAT_{D1} bezeichnet die erste Variation der D-Maße von Greenwald et al. (2003). Scores, die ausschließlich auf Basis von Aufgabenwechseldurchgängen berechnet wurden, liefern die beste Approximation an die von Greenwald et al. (1998) verwendeten Prozeduren.

* $p < .05$. ** $p < .01$.

Werden nur die Latenzen bei Aufgabenwechseldurchgängen berücksichtigt, so ist die Korrelation der inhaltlich gesättigten IATs mit dem Indikator methodenspezifischer Varianz sowohl für die IAT-Effekte auf Basis log-transformierter Antwortlatenzen (IAT_{LN}) als auch für die IAT-Effekte mit untransformierten Latenzen (IAT_{RT}) statistisch bedeutsam. Wie erwartet fällt der Anteil methodenspezifischer Varianz geringer aus, wenn nur Aufgabenwiederholungsdurchgänge in die Berechnung der Scores einfließen. Für die Daten von Experiment 11 fällt der Unterschied zwischen den Korrelationen der IAT-Effekte auf Basis der Aufgabenwechseldurchgänge und denen auf Basis der Wiederholungsdurchgänge

signifikant aus ($z = 1.66$, $p < .05$ für IAT_{RT} und $z = 1.81$, $p < .05$ für IAT_{LN}), nicht jedoch für die Daten von Experiment 12, für die der Effekt lediglich deskriptiv in die erwartete Richtung geht.

Wie schon angesprochen liefern die Effekte, die auf Basis der Aufgabenwechselfurchgänge berechnet werden, die beste Annäherung an die von Greenwald et al. (1998) verwendete Prozedur des IAT, die eine strikte Alternierung von Attribut- und Zielkonzeptaufgabe vorsieht, sodass Aufgabenwechsel bei jedem Durchgang der gemischten Phasen auftreten. Auf Grundlage der vorliegenden Analysen scheint es, dass die gegenwärtig üblichen Algorithmen in diesem Fall kaum zu einer Verringerung methodenspezifischer Varianz führen. Legt man die IAT-Effekte zugrunde, die ausschließlich auf Durchgängen mit Aufgabenwechseln beruhen, so zeigt sich jedoch ein signifikanter Unterschied des Anteils methodenspezifischer Varianz zwischen den IAT-Effekten des optimierten Algorithmus und den IAT-Effekten auf Basis der untransformierten Latenzen ($z = 1.94$, $p < .05$ für Experiment 11 und $z = 1.67$, $p < .05$ für Experiment 12). Dies legt nahe, dass der optimierte Algorithmus in der Tat den Anteil methodenspezifischer Varianz reduziert.

Die konvergente Validität fällt für alle verwendeten Algorithmen in etwa gleich aus. Die entsprechenden Korrelationen liegen jedoch deskriptiv höher, wenn nur Aufgabenwiederholungsdurchgänge in die Berechnung aufgenommen werden. Die höchste konvergente Validität von $r = .34$ ergibt sich für die IAT-Effekte auf Basis log-transformierter Latenzen (IAT_{LN}) bei Aufgabenwiederholungsdurchgängen. Dieses vergleichsweise bessere Abschneiden der IAT-Effekte für Aufgabenwiederholungsdurchgänge tritt auf, obwohl in diese Analyse nur etwa die Hälfte der insgesamt verfügbaren Durchgänge eingeht, und obwohl die mittleren IAT-Effekte für Aufgabenwiederholungsdurchgänge in allen bisher berichteten Experimenten geringer ausfallen. Dieses Muster passt zu den Vorhersagen des Aufgabenwechselmodells, das die größeren Effekte für Aufgabenwechselfurchgänge durch eine merkmalsunabhängige Kontaminierung erklärt, die eher zu einer Verringerung der konvergenten Validität führen sollte.

8 Gesamtdiskussion

Erlauben implizite Verfahren einen unmittelbaren Blick in die „Black-Box“? Diese Frage wurde am Beginn dieser Arbeit aufgeworfen. Eine positive Antwort auf diese Frage setzt zunächst eine Reihe von Annahmen über die *Struktur* dessen voraus, was man in der Black-Box vorzufinden hofft. Im Falle des IAT und einiger anderer impliziter Verfahren ist eine Kernannahme, dass assoziative Verknüpfungen das entscheidende Organisationsprinzip mentaler Repräsentation darstellen. Ob und inwiefern sich die makroskopischen Phänomene und Konstrukte der differentiellen, klinischen oder Sozialpsychologie, die mit impliziten Verfahren gemessen wurden, ohne substantiellen Bedeutungsverlust durch die Stärke assoziativer Verknüpfungen abbilden lassen, ist eine inhaltliche, weitgehend verfahrens-unabhängige Frage. Sie lässt sich wohl nicht auf abstrakter Ebene beantworten; ihre Klärung bleibt vielmehr den jeweiligen Teildisziplinen der Psychologie vorbehalten, durch konkrete Validierungsstudien.

Die Frage, ob eine Neubestimmung klassischer psychologischer Konzepte durch assoziative Verknüpfungen nun im Einzelfall substantiell gerechtfertigt ist oder nicht, ist jedoch nur eine von mehreren notwendigen Bedingungen, die die praktische Nützlichkeit des IAT- und anderer impliziter Verfahren - in Anwendungskontexten zur Voraussetzung hat. Unabhängig von dieser Frage kann etwa geklärt werden, inwiefern der IAT zur Messung assoziativer Verknüpfungen prinzipiell geeignet ist, inwiefern diese Messung direkt oder nur mittelbar erfolgt, und inwiefern bedeutsame IAT-Effekte die Folgerung zulassen, dass gemessen wurde, was gemessen werden sollte. Diese Fragen stehen im Zentrum der vorliegenden Arbeit. In Bezug auf die interne Validität wurde eine Erklärung von IAT-Effekten durch einen Trägheitsmechanismus vorgeschlagen und empirisch geprüft, die eine Reihe von Antworten auf die angesprochenen Fragen erlaubt. Im Folgenden sollen zunächst die Befunde in Bezug auf diesen Mechanismus zusammenfassend eingeordnet werden. Daraus ergibt sich ein in vieler Hinsicht konkretes Bild der Möglichkeiten und Grenzen des IAT. Im Anschluss daran, sollen die Konsequenzen der Befunde für andere Prozessmodelle wie das Figur-Grund Modell (Rothermund & Wentura, 2001), das Random-Walk Modell (Brendl et al., 2001) und das Stimulus-Response Kompatibilitätsmodell des IAT (De Houwer, 2001b) diskutiert werden. Den Abschluss der Arbeit bildet eine zusammenfassende Diskussion der praktischen Konsequenzen, die sich aus den Befunden ergeben.

8.1 Bewertung der Befunde in Bezug auf das Aufgabenwechselmodell

Insgesamt sprechen die in der vorliegenden Arbeit geschilderten Experimente für eine Erklärung von IAT-Effekten durch den Trägheitsmechanismus, der sich an den Arbeiten von Allport und Kollegen (Allport et al., 1994; Wylie & Allport, 2000) orientiert. Die zentrale Eigenschaft des Mechanismus ist, dass aufgabenirrelevante Attribute der präsentierten Reize - beziehungsweise des auf die Reize zurückgehenden Perzepts -, dann zu Interferenzen führen, wenn diese Attribute in einem kürzlich aktivierten Aufgabenset relevant waren. Maximale Interferenz ist folglich dann zu erwarten, wenn das irrelevante Attribut im direkt vorhergehenden Durchgang zu bearbeiten war, wie es bei Aufgabenwechseldurchgängen der Fall ist. Die entsprechende Interaktion zwischen IAT-Effekt und lokalen Aufgabenwechseln findet sich stabil in allen geschilderten Experimenten.

Dass dieses Muster nicht auf einem trivialen Effekt der Wiederholung spezifischer S-R Verknüpfungen beruht, zeigt die entsprechende Analyse von Experiment 1. Diese demonstriert, dass Aufgabenwechselkosten auch dann bedeutsam sind, wenn Durchgänge mit einer Wiederholung der spezifischen S-R Verknüpfung des vorherigen Durchgangs bzw. der wiederholten Klassifikation von Reizen aus der gleichen Kategorie von der Auswertung ausgeschlossen werden. Aufgabenwechsel zwischen Attribut- und Zielkonzeptaufgabe sind im Vergleich zu einem Wechsel der S-R Verbindung innerhalb der jeweils anderen Aufgabe mit bedeutsam höheren Kosten verbunden. Die Aufgabenwechselkosten können also weder durch einen trivialen Wiederholungseffekt der motorischen Reaktion, der Wiederverwendung der gleichen S-R Regel, noch der erneuten Klassifikation eines Stimulus der gleichen Kategorie erschöpfend erklärt werden. Gleichwohl sind dies sicher Teilaspekte von Aufgabensets, sodass auf Basis des Trägheitsmechanismus also durchaus Erleichterungen infolge spezifischer Wiederholungen zu erwarten sind. Der entscheidende Punkt ist, dass es sich nur um Teilaspekte handelt, die die Aufgabenwechselkosten keineswegs vollständig erklären können.

Hemmung und negatives Priming von Aufgabensets

Der Trägheitsmechanismus erklärt das Auftreten von Kompatibilitätseffekten bei der Attributaufgabe - wie sie in den Experimenten 1 und 3 geprüft und nachgewiesen wurden - durch einen Carryover von Hemmung, also durch einen negativen Primingeffekt. Hemmung ist aus Sicht des Mechanismus immer dann notwendig, wenn im vorigen Durchgang ein inkompatibler Zielkonzeptreiz zu bearbeiten war, der eine Unterdrückung des Attribut-

aufgabensets erfordert. Dabei ist zunächst nicht relevant, ob das Produkt dieses Hemmungsprozesses lediglich in der Reduktion des Aktivierungsgrades auf ein Minimum besteht oder ob eine „aktive“ Unterdrückung, im Sinne einer negativen Aktivierung resultiert. In jedem Fall sind die beobachteten Effekte weniger das Produkt von Interferenzen im Sinne gegenläufiger Evidenz, wie es das Random-Walk Modell von Brendl et al. (2001) postuliert, sondern vielmehr das Resultat von Kontrollprozessen, die an der Hemmung der irrelevanten Information beteiligt sind. Daraus leitet sich unter anderem ab, dass die Latenzen bei der Bearbeitung der Attributaufgabe eine Funktion der Kompatibilität des Vorgängerdurchgangs sein müssen: Wenn im Vorgängerdurchgang ein inkompatibler Zielkonzeptreiz zu bearbeiten war, sollte eine Übertragung (Carryover) der Hemmung die nachfolgende Anwendung der Attributaufgabe erschweren. Wenn im Vorgängerdurchgang ein kompatibler Zielkonzeptreiz zu bearbeiten war, sollte hingegen keine Verlangsamung der Attributaufgabe auftreten. Dieses Muster zeigte sich in Experiment 3, in dem kompatible und inkompatible Zielkonzeptreize innerhalb eines Blocks gemischt zu bearbeiten waren: Die Bearbeitung der Attributaufgabe war nach inkompatiblen Vorgängerdurchgängen stärker verlangsamt als nach kompatiblen. Wie schon in der Diskussion zu Experiment 3 ausgeführt, kann dieses spezifische Muster nicht durch eine Verschiebung des Antwortkriteriums erklärt werden, wie sie das Modell von Brendl et al. (2001) postuliert, da die Attributaufgabe bei Aufgabenwiederholungsdurchgängen ungehindert schnell bearbeitet wurde.

Eine weitere zentrale Vorhersage des Trägheitsmechanismus ist, dass die Kosten eines Aufgabenwechsels geringer ausfallen sollten, je länger das Intervall zwischen den Aufgabebearbeitungen ausfällt, da sowohl residuale Aktivierung als auch der Carryover von Hemmung im zeitlichen Verlauf passiv abklingen (Allport et al., 1994). Die Interaktion der Aufgabenwechselkosten mit dem Response-Stimulus Intervall in Experiment 1 bestätigt diese Annahme. Wie aufgrund des Trägheitsmechanismus zu erwarten, tritt diese Interaktion vornehmlich in der inkompatiblen Phase des IAT auf. Auf die kompatible Phase des IAT wirkt sich das RSI hingegen in bedeutsam geringerem Maße aus, vermutlich, weil die Trägheit von Aufgabensets in diesem Fall keine störenden Interferenzen zur Folge hat, allenfalls erleichternde Effekte. Infolgedessen sind weder eine Hemmung der Attributaufgabe noch ein Carryover der Hemmung zu erwarten, der mit zunehmenden RSI geringer ausfallen könnte, was die empirischen Befunde auch bestätigen.

In analoger Weise lassen sich die Effekte der Länge von Aufgabenwiederholungssequenzen in Experiment 8 erklären. In diesem Experiment ergab sich eine stärkere

Asymmetrie der Aufgabenwechselkosten, wenn die Aufgabenwechsel auf Sequenzen mit durchschnittlich vier Aufgabenwiederholungen folgten, als wenn die Sequenzen im Mittel nur zwei Aufgabenwiederholungen umfassten. Die Wechselkosten sind offenbar eine Funktion des Grades, zu dem die vorher bearbeitete Aufgabe eingeübt wurde. Dies entspricht dem Befund von Allport et al. (1994), dass Wechsel von einer gut eingeübten Aufgabe ausgehend mit größeren Kosten verbunden sind als Wechsel von einer weniger gut eingeübten Aufgabe aus. Dieser Mechanismus ist analog zu dem Mechanismus, durch den der umgekehrte Stroop-Effekt erklärt wird - also das Auftreten stärkerer Stroop-Effekte beim Vorlesen der Farbwörter, wenn vorher die Farbnennung eingeübt wurde (Allport et al., 1994).

Globale und längerfristige Effekte

Neben den hier angesprochenen lokalen Sequenzeffekten liefern die Experimente 4 bis 6 eine Reihe von Hinweisen auf längerfristige und globale Effekte der Bearbeitung eines IAT. Die längerfristigen Trägheitseffekte auf die verbale Bewertung der Zielkonzeptreize in Experiment 5 und die evaluativen Ratings der Zielkonzeptreize in Experiment 6 belegen, dass ein globales Analogon der lokalen Sequenzeffekte über mehrere Durchgänge hinweg nachwirkt. Nachwirkungen traten nur nach inkompatiblen IAT-Blöcken auf, und auch nur dann, wenn diese Aufgabenwechseldurchgänge umfassten. Weder nach kompatiblen IAT-Blöcken noch nach Blöcken, in denen Attribut- und Zielkonzeptaufgabe entmischt waren, konnten vergleichbare Nachwirkungen nachgewiesen werden. Gemäß den Vorhersagen des Trägheitsmechanismus waren diese längerfristigen Nachwirkungen spezifisch für die Bewertung der Zielkonzeptreize. Eine analog konstruierte Kontrollaufgabe, die die Nennung (Experiment 5) bzw. das Einstufen (Experiment 6) der Farbe von Kontrollreizen umfasste, zeigte sich von der Vorgängerphase unbeeinflusst, obwohl Farbnennungs- und Bewertungsaufgabe innerhalb der gleichen Testblöcke gemischt zu bearbeiten waren.

Die längerfristigen Nachwirkungen zeigen sich überraschend analog zu den kurzfristigen und lokalen Trägheitseffekten, die im vorigen Abschnitt diskutiert wurden. Trotz einer nachweisbaren Verringerung von Hemmungseffekten innerhalb des kurzen zeitlichen Intervalls von 1000 ms in Experiment 1, und trotz des nachweisbaren Abklingens von Trägheitseffekten über eine vergleichsweise geringe Zahl von Durchgängen, wie sie in der generellen Asymmetrie der Aufgabenwechselkosten sichtbar ist, scheinen residuale Trägheitseffekte also über viele Durchgänge hinweg anzuhalten. Dies zeigen auch die Ergebnisse von Experiment 4, in dem Nachwirkungen der Bearbeitung von inkompatiblen IAT-Blöcken

sowohl unmittelbar als auch längerfristig auftraten. Die Bearbeitung kompatibler Blöcke war nach inkompatiblen signifikant verlangsamt, obwohl die Tastenzuordnung in diesem Experiment über den Wechsel der Kompatibilität hinweg unverändert blieb. Tatsächlich wurden die kompatiblen Reize zunächst kaum schneller bearbeitet als die direkt vorhergehenden inkompatiblen Reize. Dieser Effekt entspricht exakt dem, was bei einem Abklingen der Hemmung des Attributaufgabensets im Verlauf der Bearbeitung zu erwarten wäre. Die zunächst vorgeschlagene Erklärung dieses globalen Effekts über den Vereinfachungsmechanismus scheint in diesem Sinne wenig überzeugend. Auch die globalen Effekte können sparsam über eine allgemeine Trägheit von Aufgabensets erklärt werden.

Die Rolle exekutiver Kontrollprozesse

Ein zentrales Ergebnis der Experimente zur Rolle exekutiver Kontrollprozesse ist, dass die lokalen Aufgabenwechselkosten weder durch symbolische Aufgabenhinweise (Koch, 2003; Meiran, 1996) in Experiment 7 noch durch vorhersagbare Aufgabensequenzen oder eine nicht-symbolische Vorankündigung der Aufgabe in Experiment 8 reduziert werden konnten. Gleichzeitig führten alle drei Typen von Aufgabenhinweisen zu einer signifikanten Reduktion der IAT-Effekte. Dieses zunächst überraschende Ergebnis wirft eine Reihe von Fragen auf. Einerseits ist dieser Befund gut mit den Annahmen des Trägheitsmechanismus vereinbar, der Aufgabenwechselkosten nicht durch die Dauer eines exekutiven Kontrollprozesses erklärt, welcher durch eine Vorbereitung innerhalb des RSI erleichtert werden könnte. Andererseits stellt sich die Frage, weshalb IAT-Effekte überhaupt eine Funktion der Vorbereitung auf die zu bearbeitende Aufgabe sein sollen, wie es der Befund ihrer bedeutsamen Reduktion durch die Aufgabenhinweise nahe legt.

Aufgabenhinweise könnten die Performanz insofern beschleunigt haben, als sie die Menge der danach wahrscheinlichen Reize verkleinern. Die Erwartung darüber, welche Reize im folgenden Durchgang wahrscheinlich sind, könnte dann zu einer allgemeinen Beschleunigung der Performanz führen. Damit ist jedoch keineswegs erklärt, warum der Effekt für kompatible und inkompatible Aufgabenblöcke unterschiedlich ausfallen sollte. Vielmehr wäre zu erwarten, dass er von der Tastenzuordnung unabhängig ist, da er eine rein perzeptuelle Erleichterung darstellt. Es scheint also nahe liegend, dass der Effekt der Aufgabenhinweise tatsächlich mit der Identifizierung, Auswahl und Implementierung von Aufgabensets zu tun hat. Darüber, welche Prozesse dies genau sind, kann nur spekuliert werden. Das Ausbleiben einer Interaktion zwischen Aufgabenhinweisen und dem Aufgaben-

wechselfaktor zeigt jedoch, dass die Vorbereitung sowohl bei Aufgabenwechsel- als auch bei Aufgabenwiederholungsdurchgängen in der inkompatiblen Phase hilfreich ist.

Eine plausible Erklärung wäre, dass der Einfluss der Aufgabenvorbereitung vor allem auf eine Erleichterung der *Identifizierung* der anstehenden Aufgabe zurückgeht. Eine Lokation der Vorbereitungseffekte in der Phase der Aufgabenidentifizierung kann das beobachtete Muster ohne weitere Vorannahmen erklären. Einerseits sollte die Identifizierung der zu bearbeitenden Aufgabe für Aufgabenwiederholungs- und Aufgabenwechseldurchgänge in der inkompatiblen Bedingung gleich entscheidend sein, da in jedem Fall erkannt werden muss, ob sich die Aufgabe wiederholt oder nicht. Andererseits sollte die Identifizierung der Aufgabe in der kompatiblen Bedingung weniger bedeutend sein, da auch die Verarbeitung der irrelevanten Attributinformation zu korrekten Antworten führt. Letzteres erklärt das vollständige Ausbleiben von Effekten in der kompatiblen Bedingung des IAT, ersteres das Auftreten von unspezifischen Effekten mit inkompatibler Tastenzuordnung³. Da es sich dabei jedoch um eine post hoc Erklärung handelt, sind weitere Studien nötig, um diese Frage abschließend zu beantworten.

Materialeigenschaften und Stimuluseffekte

Was der IAT eigentlich misst, ist nach wie vor ungeklärt. Obwohl diese Frage nicht im Zentrum der vorliegenden Arbeit steht, soll sie im Folgenden kurz aufgegriffen werden. Aus Sicht des hier vorgeschlagenen Trägheitsmechanismus ist die Menge der Relationen zwischen Zielkonzepten und Attributen, mit denen IAT-Effekte auftreten können, zunächst offen. Die einzige Voraussetzung, die der Mechanismus in dieser Hinsicht macht, ist, dass das Reizmaterial ein Merkmal aufweisen muss, das nur in einer der Aufgaben relevant ist, aber

³ Eine zentrale Rolle bei der erleichterten Identifizierung von Aufgabenwechsel- und Aufgabenwiederholungsdurchgängen scheint das Gleichbleiben bzw. Wechseln der Aufgabenhinweise selbst zu spielen. Wie Logan und Bundesen (2003) zeigen, sind die Antwortlatenzen bei Wiederholung des gleichen Aufgabenhinweises signifikant schneller als bei Wiederholung der gleichen Aufgabe: Die Latenzen bei Aufgabenwiederholungen mit wechselndem Hinweisreiz sind nur unwesentlich schneller als die Latenzen bei Aufgabenwechseln. Das Gleichbleiben und Wechseln der Aufgabenhinweise in den Experiment 7 und 8 könnte also eine Bedingung für die schnelle Identifikation von Aufgabenwiederholungs- und Wechseldurchgängen sein.

bei den Exemplaren beider Aufgaben prinzipiell vorliegt. Im Falle eines Einstellungs-IAT trifft dies üblicherweise auf die Bewertung der Zielkonzeptreize zu, im Fall des IAT mit geometrischen Objekten auf die Größe der Zielkonzeptreize. Aus dieser Perspektive lässt sich der Effekt zusätzlicher „Crosscategory Associations“ erklären (De Houwer, 2003; Steffens et al., 2004; Steffens & Plewe, 2001), etwa die Verschiebung der IAT-Effekte durch zusätzliche stereotype Assoziationen der Attributreize. Ebenfalls mit dem Mechanismus kompatibel sind Effekte von Salienzasymmetrien, wie sie von Rothermund und Wentura (2001) gefunden wurden. Sofern die Versuchsteilnehmer ihre Urteile tatsächlich auf Basis unterschiedlich ausgeprägter Salienz der Reize oder Reizkategorien treffen, ist die oben genannte Voraussetzung erfüllt.

Gleichzeitig zeigen jedoch die Befunde mit dem Geometrie-IAT, insbesondere die Ergebnisse von Experiment 2, dass Salienzunterschiede keine notwendige Bedingung für das Auftreten von IAT-Effekten sind, sondern allenfalls unter bestimmten Bedingungen hinreichend. Die Kritik des Figur-Grund Modells muss also in dieser Hinsicht relativiert werden. Wenn nicht gezeigt werden kann, dass Figur-Grund Asymmetrien eine notwendige Voraussetzung für IAT-Effekte sind, ist fraglich, inwiefern die Erklärung inhaltlicher Effekte über diesen Mechanismus sinnvoll ist, denn in vielen Fällen scheint eine Erklärung über Assoziationsmuster weitaus nahe liegender und sparsamer.

Problematisch für den Trägheitsmechanismus ist jedoch, dass die Verwendung atypisch bewerteter prominenter In- und Ausländer als Exemplare der Zielkonzepte in der Untersuchung von De Houwer (2001b) keinen nachweisbaren Einfluss auf die IAT-Effekte hatte. Ebenso schwer zu erklären scheint der Befund von Blümke und Friese (2004), dass IAT-Effekte auch dann auftreten, wenn die Bewertung der Zielkonzeptexemplare in einem Einstellungs-IAT neutral gehalten wird. Diese Befunde mit atypischen Reizen sprechen für eine entscheidende Rolle der übergeordneten Konzepte. Gleichzeitig zeigen die hier vorgelegten Ergebnisse mit dem Geometrie-IAT, dass eine Beziehung auf der Ebene der Konzepte keineswegs notwendig für das Auftreten von IAT-Effekten ist; diese basieren ausschließlich auf der Ausprägung spezifischer Merkmale der Exemplare.

In diesem Zusammenhang sei darauf hingewiesen, dass die dargestellten Exemplareffekte keineswegs spezifisch für den IAT sein müssen. Dass die Wahrnehmung und Verarbeitung von Informationen über Exemplare einer Kategorie im sozialen Kontext mehr oder weniger automatisch in Termini kategorialer Zugehörigkeiten erfolgen kann, ist ein wohlbekannter Effekt (Gawronski et al., 2003; Klauer & Wegener, 1998; Taylor et al., 1978).

Soziale Kategorien unterscheiden sich systematisch bezüglich ihres Einflusses auf die Wahrnehmung von Information über spezifische Exemplare (z. B. Stangor et al., 1992). Entsprechend sollten Exemplareffekte auch im IAT vor allem bei schwachen Kategorien auftreten. Bei starken sozialen Kategorien wäre dagegen zu erwarten, dass individuierende und spezifische Information durch die Eigenschaften der übergeordneten Kategorie überschrieben oder verschoben wird. Das Ausbleiben von Exemplareffekten im IAT könnte aus dieser Perspektive als Stereotypisierungsprozess verstanden werden. Die Befundlage zur Rolle der Exemplare im IAT ist jedoch insgesamt widersprüchlich. Eine abschließende Bewertung kann nicht ohne weitergehende experimentelle Evidenz gelingen und soll deshalb an dieser Stelle auch nicht versucht werden.

Methodenspezifische Varianz

Die Befunde zu methodenspezifischer Varianz im IAT zeigen, dass die Rolle von Aufgabenwechselkosten im IAT über die eines vermittelnden Prozesses hinausgeht. Wäre die Rolle der Aufgabenwechselprozesse rein vermittelnder Natur, so wäre zu erwarten, dass Aufgabenwechselkosten ein Gemisch aus merkmalsabhängiger Varianz und einem Anteil von unsystematischem Messfehler darstellen. Unter dieser Bedingung wären Zusammenhänge zwischen IATs offenbar nur dann zu erwarten, wenn auch die gemessenen Merkmale zusammenhängen. Die bedeutsamen Korrelationen des Geometrie-IAT mit dem Blumen-Insekten-IAT und dem Extraversions-IAT zeigen jedoch, dass dies nicht der Fall ist. Ob diese Zusammenhänge tatsächlich auf die spezifischen exekutiven Kontrollerfordernisse bei Aufgabenwechseln im IAT zurückgehen, oder auf ein breiteres Fähigkeitskonstrukt, etwa die allgemeine kognitive Geschwindigkeit oder die Arbeitsgedächtniskapazität, kann auf Basis der vorliegenden Daten nicht eindeutig geklärt werden. Sicher folgern kann man jedoch, dass eine systematische merkmalsunabhängige Varianzkomponente - die unterschiedlichen IATs gemeinsam und über verschiedene Inhalte hinweg stabil ist - Korrelationen zwischen IATs erklären kann. Dies gilt nicht minder, wenn eine inhaltliche Ursache der Korrelationen plausibel erscheint, weshalb bei der Interpretation von Korrelationen zwischen IATs als Indikator von Merkmalszusammenhängen äußerste Vorsicht geboten ist.

Dass Prozesse, die an der Lösung exekutiver Kontrollerfordernisse beteiligt sind, die Ursache methodenspezifischer Varianz darstellen, ist gegenwärtig die plausibelste Erklärung. Zum einen steht diese Erklärung im Einklang mit Befunden der Altersforschung (Kray et al., 2002; Kray & Lindenberger, 2000; Salthouse et al., 1998; Verhaeghen & Cerella, 2002), zum

anderen kann keiner der in der Literatur vorgeschlagenen Mechanismen (Brendl et al., 2001; De Houwer, 2001b; Karpinski & Hilton, 2001; Rothermund & Wentura, 2001) die Effekte erklären. Abgesehen davon sprechen die korrelativen Befunde, die im Abschnitt zur Reduktion methodenspezifischer Varianz berichtet sind, dafür, dass Aufgabenwechselkosten mindestens eine Komponente methodenspezifischer Varianz darstellen. Der Anteil methodenspezifischer Varianz fiel in diesen Analysen niedriger aus, wenn nur Aufgabenwiederholungsdurchgänge ausgewertet wurden, und höher, wenn nur Aufgabenwechseldurchgänge in die IAT-Effekte eingingen. Gleichzeitig fiel die konvergente Validität für Aufgabenwiederholungsdurchgänge höher aus als für Aufgabenwechseldurchgänge. Die entsprechenden Korrelationsunterschiede sind statistisch nur schlecht nachweisbar. Der Vergleich ist aber aus theoretischer Sicht ohnehin nicht befriedigend, da auch Aufgabenwiederholungsdurchgänge nicht frei von Nachwirkungseffekten sein dürften, wie die Experimente 4 bis 6 nahe legen. Eine endgültige Klärung der Ursache methodenspezifischer Varianz dürfte also erst mit Hilfe weiterer Untersuchungen möglich werden.

8.2 Bewertung der Befunde in Bezug auf alternative Prozessmodelle des IAT

In den bisherigen Abschnitten wurde diskutiert, inwiefern die Befunde mit dem Trägheits- und dem Vereinfachungsmechanismus des IAT im Einklang stehen. Der Trägheitsmechanismus kann auf Basis dieser Betrachtung als empirisch bewährt gelten, da der größte Teil der Befunde unmittelbar und ohne Zusatzannahmen durch diesen Mechanismus erklärt werden kann. Der Vereinfachungsmechanismus ist hingegen mit einigen Befunden nicht vereinbar. Welche Folgerungen lassen sich jedoch in Bezug auf die anderen Prozessmodelle des IAT ableiten? Diese Frage soll im Folgenden geklärt werden.

Aktivierungsausbreitung

Wie schon im Theorieteil der Arbeit erläutert, wirft die Erklärung von IAT-Effekten durch Aktivierungsausbreitung eine Reihe konzeptueller Probleme auf, die darauf zurückgehen, dass die kompatible und die inkompatible Phase eines IAT stimulusseitig identisch sind. Aktivierungsausbreitung als rein stimulusgetriebener Prozess sagt auf dieser Grundlage zunächst keine Performanzunterschiede vorher. Trotzdem wird Aktivierungsausbreitung teilweise als vermittelnder Mechanismus vermutet. Die vorliegende Arbeit liefert neben den konzeptuellen Argumenten gegen diesen Ansatz eine große Zahl empirischer hinzu: Weder

die Asymmetrie der Aufgabenwechselkosten, noch die Kompatibilitätseffekte bei der Attributaufgabe, noch die längerfristigen Trägheitseffekte oder die Präsenz methodenspezifischer Varianz sind durch reine Aktivierungsausbreitung erklärbar. Ferner zeigen die Experimente mit dem Geometrie-IAT, dass assoziative Verknüpfungen keineswegs notwendige Bedingung für das Auftreten von IAT-Effekten sind. Aktivierungsausbreitung hat sich als Erklärung von IAT-Effekten somit nicht bewährt.

Das Random-Walk Modell

Das Random-Walk Modell (Brendl et al., 2001) liefert eine einfache und nahe liegende Erklärung von IAT-Effekten durch Unterschiede in der Netto-Akkumulationsrate antwortrelevanter Information und der Verschiebung des Antwortkriteriums in der inkompatiblen Phase eines IAT. Einige der vorliegenden Befunde sind mit diesem Modell gut vereinbar, etwa das Auftreten von Effekten im Geometrie-IAT, oder die Tatsache, dass Kompatibilitätseffekte auch bei der Attributaufgabe auftreten. Das Auftreten von Aufgabenwechselkosten steht jedoch mit einer Grundannahme des Modells nicht in Einklang, nämlich der Annahme konstanter Akkumulationsraten für eingehende Information. Aufgabenwechseleffekte lassen sich prinzipiell jedoch auch in einem Random-Walk Modell modellieren, etwa wenn man annimmt, dass exekutive Kontrollprozesse die Gewichtung eingehender Information aufgabenabhängig verschieben. Auch wenn lokale Sequenzeffekte somit nicht prinzipiell inkompatibel mit dem vorgeschlagenen Mechanismus sind, lässt sich sagen, dass diese durch das Modell zunächst nicht vorhergesagt werden.

Längerfristige Reihenfolge- und Testwiederholungseffekte (Experimente 4 bis 6) lassen sich durch das Random-Walk Modell mit einer einfachen Zusatzannahme erklären. Wie weiter oben erläutert ist eine Kernannahme des Modells, dass die Bearbeitung einer inkompatiblen Phase eines IAT zu einer Verschiebung des Antwortkriteriums führen sollte. Nimmt man nun zusätzlich an, dass ein konservativeres Antwortkriterium auch auf die nachfolgenden Aufgabenbearbeitungen generalisiert, lassen sich längerfristige Nachwirkungen durch eine damit einhergehende aufgabenunspezifische Verlangsamung erklären. Dass der Effekt aus Sicht des Random-Walk Modells aufgabenunspezifisch sein muss, liegt auf der Hand: Die Kriteriumsverschiebung wird zur Erklärung der verlangsamten Bearbeitung der Attributaufgabe ins Feld geführt, also eines unspezifischen Effekts. In den Experimenten 5 und 6 traten die Trägheitseffekte entgegen dieser Vorhersage jedoch nicht unspezifisch auf.

Eine relative Verlangsamung der Bearbeitung nach inkompatiblen Blöcken zeigte sich äußerst spezifisch nur für die Bewertungsaufgabe, nicht aber für eine analoge Farbnennungsaufgabe.

Ein weiteres Argument gegen das Random-Walk Modell betrifft eine zentrale Eigenschaft des vorgeschlagenen Mechanismus. Nimmt man mit Brendl und Kollegen an, dass IAT-Effekte tatsächlich die Gegen- bzw. Gleichläufigkeit irrelevanter Information für die Zielkonzeptaufgabe widerspiegeln, so sollte eine Verschiebung des Antwortkriteriums zu einer stärkeren Vergrößerung der IAT-Effekte für die Zielkonzeptaufgabe im Vergleich zur Attributaufgabe führen. Zwar sollte die unspezifische Verschiebung des Antwortkriteriums auch die Performanz bei der Attributaufgabe mit inkompatibler Tastenzuordnung betreffen, die Kompatibilitätseffekte sollten jedoch für die Zielkonzeptaufgabe stärker betroffen sein, da hier das konservativere Antwortkriterium *und* die unterschiedliche Nettoakkumulationsrate zusammenfallen. Die Ergebnisse von Experiment 1 zeigen jedoch, dass genau das Gegenteil der Fall ist. Die IAT-Effekte sind in Bezug auf die Attributaufgabe größer als in Bezug auf die Zielkonzeptaufgabe, stehen also in starkem Kontrast zu der Kernannahme des Random-Walk Modells. Aus Sicht des Trägheitsmechanismus ist dieser Effekt leicht zu erklären: Die erschwerte Bearbeitung der Attributaufgabe in der inkompatiblen Bedingung geht auf negatives Priming zurück. Nur wenn die Verarbeitung des Attributs in einem vorhergehenden Durchgang interferiert, muss diese unterdrückt werden. Experiment 3 zeigt genau das vorhergesagte Muster: Die Attributaufgabe wird nach inkompatiblen Zielkonzeptreizen signifikant langsamer bearbeitet als nach kompatiblen, auch wenn diese innerhalb des gleichen Aufgabenblocks gemischt sind. Das Random-Walk Modell ist mit diesen Ergebnissen nicht vereinbar.

Stimulus-Response-Kompatibilität

Das Modell von De Houwer (2001b) erklärt den IAT-Effekt durch eine erworbene Kompatibilität zwischen Reizen und Reaktionen. Der Mechanismus setzt voraus, dass die einer Antworttaste zugeordneten Antwortkategorien konsistent bewertet sind. Dies ist eine hinreichende, aber auch notwendige Bedingung dafür, dass die Antworttasten eine evaluative Konnotation erwerben. Zur Erklärung von IAT-Effekten nimmt De Houwer an, dass diese Bedingung in der kompatiblen Phase erfüllt ist, wenn etwa (positive) Blumen und positive Adjektive einer Taste zugeordnet sind, nicht jedoch in der inkompatiblen Phase, wenn etwa (negative) Insekten und negative Adjektive einer Taste zugeordnet sind. Obwohl De Houwer (2001b) vornehmlich auf Einstellungs-IATs Bezug nimmt, kann dieser Mechanismus im

Prinzip auf beliebige Inhalte verallgemeinert werden. Die Effekte des IAT mit geometrischen Objekten sind also mit dem Modell vereinbar. Das gilt auch für andere Ergebnisse der geschilderten Experimente, etwa die asymmetrischen Aufgabenwechselkosten, oder das Auftreten methodenspezifischer Varianz. Einerseits werden die Effekte durch das Modell nicht vorhergesagt, andererseits sind sie nicht streng unvereinbar mit dessen Kernannahmen.

Problematisch für das Stimulus-Response Kompatibilitätsmodell ist jedoch, dass in Experiment 3 auch dann Kompatibilitätseffekte auftraten, wenn kompatible und inkompatible Durchgänge innerhalb des gleichen Aufgabenblocks *gemischt* bearbeitet werden. Die in diesem Experiment realisierte Prozedur erfüllt offenkundig nicht die Voraussetzung einer konsistenten Attributausprägung der einer Antworttaste zugeordneten Antwortkategorien. Es ist kaum plausibel erklärbar, wie unter dieser Bedingung eine Konnotation der Antworttasten erworben werden soll. Ebenso scheinen die längerfristigen modalitätsunabhängigen Nachwirkungseffekte in den Experimenten 5 und 6 nur schwer mit dem Modell von De Houwer vereinbar. Wenn der IAT-Effekt auf einer erworbenen Konnotation der Antworttasten beruht, sollten Nachwirkungen – sofern man sie mit diesem Modell erklären möchte - für die Antwortmodalität spezifisch sein. Insgesamt steht eine eingehende empirische Prüfung des Modells gegenwärtig noch aus. Es gibt weder klare Belege gegen das Modell, noch Befunde, die für den vorgeschlagenen Mechanismus sprechen.

Figur-Grund Modell

Das Figur-Grund Modell (Rothermund & Wentura, 2001) ist nur in zweiter Linie ein Prozessmodell des IAT. Auf Prozessebene entspricht es weitgehend dem Vereinfachungsmechanismus des IAT. Damit treffen die weiter oben geschilderten Probleme des Vereinfachungsmechanismus auch auf das Figur-Grund Modell zu. Unter anderem können die Effekte des RSI in Experiment 1 und die Effekte der Länge von Aufgabenwiederholungssequenzen in Experiment 8 durch das Modell nicht erklärt werden.

Die Kernannahmen des Figur-Grund Modells beziehen sich jedoch auf die Art der gemessenen Inhalte. Rothermund und Wentura gehen davon aus, dass in vielen angewandten IATs nicht die intendierten assoziativen Verknüpfungen erfasst werden, sondern Salienzunterschiede, die mit den Inhalten konfundiert sind. Das Modell nimmt an, dass die Reize nicht wie instruiert anhand der relevanten Merkmale klassifiziert werden, sondern vielmehr nach ihrer Salienz. Dieser Mechanismus lässt offenkundig nur dann IAT-Effekte erwarten, wenn auch Salienzasymmetrien vorliegen. Solche Asymmetrien können für den IAT mit

geometrischen Objekten relativ sicher ausgeschlossen werden. Selbst wenn man sich darauf festlegt, dass eine der Attributausprägungen (groß oder klein) und eines der Zielkonzepte (rot oder blau) salienter ist als der jeweilige Gegenpol, bleibt offen, warum sich die Effekte in Experiment 2 exakt umkehren, wenn die Kontingenz zwischen Farbe und Größe der Reize umgekehrt wird. Einen Effekt dieser Umkehrung auf die Salienz kann man ausschließen.

Effekte von Figur-Grund Asymmetrien sind prinzipiell sowohl mit dem Trägheits- als auch mit dem Vereinfachungsmechanismus vereinbar. Wie das Figur-Grund Modell selbst setzt diese Erklärung jedoch voraus, dass nicht die instruierten Klassifikationsaufgaben bearbeitet werden, was für verschiedene Materialien mehr oder weniger plausibel sein kann. Offenkundig ist jedoch, dass die Relevanz eines Merkmals aus Sicht des Trägheitsmechanismus davon abhängt, welche Aufgabe die Versuchsteilnehmer bearbeiten, nicht jedoch davon, welche Aufgabe instruiert wurde. Aus dieser Perspektive erscheinen Figur-Grund Asymmetrien als eine Ursache von IAT-Effekten, die *hinreichend*, nicht jedoch *notwendig* für deren Auftreten ist. Im Zusammenhang mit den oben genannten Ergebnissen, die nicht mit dem Modell vereinbar sind, scheint dieses Fazit die Befundlage gut zusammenzufassen.

8.3 Praktische Konsequenzen

Neben den dargestellten theoretischen Konsequenzen haben die Befunde dieser Arbeit auch Folgen für die große Zahl (mehr oder weniger unkritischer) Anwendungen des IAT auf praktische Fragestellungen. Dabei lassen sich zwei Problemfelder unterscheiden, nämlich solche Probleme, die auf Nachwirkungen der Bearbeitung eines IAT zurückgehen, wie sie in den Experimenten 4 bis 6 untersucht wurden, und solche, die sich aus der reliablen Kontamination des IAT mit methodenspezifischer Varianz ergeben. Zum Abschluss soll eine allgemeine Bewertung der Tauglichkeit des IAT als Maß interindividueller Unterschiede versucht werden.

Längerfristige Nachwirkungen

Eine Reihe von Studien legt nahe, dass die Bearbeitung von IAT-Blöcken die Bearbeitung nachfolgender IAT-Blöcke beeinflussen kann. Dies zeigt der schon angesprochene Effekt der Kompatibilitätsreihenfolge ebenso wie der Befund, dass die Retest-Reliabilität meist geringer ausfällt als die interne Konsistenz (Asendorpf et al., 2002; Bosson et al., 2000; Egloff & Schmukle, 2002; Greenwald & Farnham, 2000; Greenwald et al., 1998;

Nosek et al., 2003; Steffens & Buchner, 2003). Diese beiden Phänomene spiegeln insofern eine Reaktivität des IAT wider, als sie auf die vorhergehende Bearbeitung einzelner IAT-Blöcke bzw. auf eine komplette Testwiederholung zurückgehen müssen. Vor allem für die individuelle Interpretation von Testergebnissen ist es jedoch unangenehm, wenn die Effekte prozeduraler Variationen, wie der Test- oder Kompatibilitätsreihenfolge, mit individuellen Merkmalsausprägungen konfundiert sind.

Der Effekt der Kompatibilitätsreihenfolge (z. B. Greenwald et al., 1998; Greenwald et al., 2003; Nosek et al., 2003) trat in den hier berichteten Experimenten häufig an der Schwelle statistischer Bedeutsamkeit auf. Die Ergebnisse weisen jedoch deskriptiv ein spezifisches Muster auf, das über die Untersuchungen hinweg konsistent und stabil auftrat: Die IAT-Effekte waren stärker ausgeprägt, wenn die kompatible Phase der inkompatiblen voranging. Experiment 4 zeigt in diesem Zusammenhang, dass die Effekte offenbar eine Nachwirkung der inkompatiblen Phase darstellen, also einen negativen Transfer (Nosek et al., 2002). In den Experimenten 5 und 6 wurde einer Erklärung solcher Nachwirkungseffekte durch eine längerfristige Trägheit von Aufgabensets nachgegangen. Dabei zeigte sich, dass Nachwirkungen der Bearbeitung von IAT-Blöcken signifikant stärker ausfielen, wenn Attribut- und Zielkonzeptaufgabe gemischt bearbeitet wurden. Wurden hingegen beide Aufgaben innerhalb der IAT-Blöcke entmischt, hatte die Kompatibilität der Tastenzuordnung keinen bedeutsamen Effekt auf die folgenden Aufgabenblöcke. Die Nachwirkungen scheinen also auf Aufgabenwechselprozesse zurückzugehen.

Interessanterweise traten die Nachwirkungseffekte in diesen Experimenten auf, obwohl die Antwortmodalität vom Tastendrücken zum Aussprechen der Antwortkategorie beziehungsweise zu einer Ratingaufgabe wechselte. Dies spricht dafür, dass die Effekte relativ unabhängig von der spezifischen zu lösenden Aufgabe sind, sofern diese eine evaluative Bewertung umfasst. Vielmehr scheint die Bearbeitung der IAT-Blöcke in diesen Experimenten den Abruf der eigenen Einstellungen und Bewertungen insgesamt zu erschweren. Könnten solche Effekte also auch explizite und Verhaltensmaße beeinflussen? Die vorliegenden Daten legen nahe, dass dies der Fall ist. Eine empirische Untermauerung dieser Hypothese steht jedoch gegenwärtig aus. Träfe sie zu, ergäben sich jedoch offenkundig weitreichende Konsequenzen für das Design von Untersuchungen mit dem IAT. Nicht geklärt ist, ob die Effekte auch bei neuen Reizen auftreten, also solchen, die nicht im Kontext der IAT-Blöcke bearbeitet wurden. Aus Sicht des Trägheitsmechanismus scheint plausibel, dass nur die Bewertung derjenigen Zielkonzeptreize betroffen ist, die in den IAT-Blöcken

bearbeitet wurden. Wäre dies der Fall, könnten Reaktivitätseffekte möglicherweise durch den Einsatz paralleler oder sehr großer Stimulusmengen verhindert werden.

Verschiedene Autoren nehmen an, dass eine Ausbalancierung der Kompatibilitätsreihenfolge zu einer Verringerung der Korrelationen zwischen dem IAT und expliziten Maßen führen kann (vgl. Banse et al., 2001; Gawronski, 2002), da eine Ausbalancierung zu zwei unterschiedlichen Verteilungen der IAT-Effekte in den Reihenfolgebedingungen führen sollte. Dadurch sollte der Anteil an Fehlervarianz zunehmen, wenn die Daten aus diesen Bedingungen zusammengeführt werden, und entsprechend eine verringerte Korrelation auftreten. Gleichwohl dies Argument zutrifft, sollte man beachten, dass eine Konstanthaltung der Reihenfolge, in der zwei Tastenzuordnungen bearbeitet werden, nicht das gleiche ist, wie eine Konstanthaltung der Kompatibilitätsreihenfolge. Wenn die IAT-Effekte unterschiedlicher Probanden in ihrer Richtung variieren, führt eine konstante Reihenfolge der Tastenzuordnungsbedingungen nämlich zu einer unerwünschten Konfundierung zwischen Merkmalsausprägung und Reihenfolgebedingung.

Wird etwa in einem IAT zur Einstellung gegenüber Deutschen und Türken immer die Tastenzuordnung zuerst bearbeitet, die für eine Person mit Vorurteilen gegenüber türkischen Mitbürgern leichter zu bearbeiten ist, so vermengt sich der Kompatibilitätsreihenfolgeeffekt mit dem resultierenden IAT-Effekt: Personen mit Vorurteilen gegenüber Türken bearbeiten zuerst die kompatible Tastenzuordnung, dann die inkompatible, während Personen mit einer positiven Einstellung gegenüber Türken die Phasen in umgekehrter Reihenfolge bearbeiten. Dies könnte eine stärkere Differenzierung der Personengruppen zur Folge haben, da die IAT-Effekte einstellungsabhängig mit Kompatibilitätsreihenfolgeeffekten vermengt sind. Unter bestimmten Umständen wäre also tatsächlich ein größerer Zusammenhang mit Außenkriterien zu erwarten, jedoch nicht aufgrund eines geringeren Messfehleranteils, sondern aufgrund der genannten Konfundierung. Zwar könnte man sich auf den Standpunkt stellen, dass eine höhere Validität prinzipiell wünschenswert ist; dabei sollte man jedoch beachten, dass die Interpretierbarkeit individueller Messwerte trotz der höheren Validität möglicherweise geringer ausfällt.

Die tendenzielle Verringerung von IAT-Effekten und internen Konsistenzen bei einer Testwiederholung (Cunningham et al., 2001; Steffens & Buchner, 2003) und die Divergenz zwischen interner Konsistenz und Retest-Reliabilität (Steffens & Buchner, 2003) können ebenso durch längerfristige Trägheitseffekte erklärt werden. In den hier dargestellten Experimenten 9 bis 12 traten solche Effekte nur teilweise auf, jedoch immer in der erwarteten

Richtung. Zwar waren die Effekte immer verringert, wenn der gleiche IAT wiederholt bearbeitet wurde, eine Generalisierung dieses Effekts auf die Wiederholung andersartiger IATs (Experiment 11) zeigte sich jedoch nur für den IAT mit geometrischen Objekten. Der Blumen-Insekten-IAT war von der seriellen Position im Experiment unbeeinflusst. Eine abschließende Klärung dieser Frage bedarf also weiterer Experimente.

Unabhängig von der Generalisierung der Effekte auf neue Materialien ist jedoch zu erwarten, dass ein erschwerter Abruf von Attributinformation IAT-Effekte verringert und eine Reihe anderer so genannter Malleability-Effekte (Blair, Ma & Lenton, 2001; Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001) zur Folge haben kann, etwa eine Veränderung von IAT-Effekten durch die Vorstellung stereotyp-inkongruenter Inhalte. Obwohl diese Tatsache sicher eine Einschränkung der Nützlichkeit des IAT bedeutet, sollte man sehen, dass auch explizite Verfahren ebenso wie offenes Verhalten Nachwirkungen auf anschließende Messungen haben können. Bei Präferenz- oder Temperamentsmessungen kann zum Beispiel ein Konsistenzstreben dazu führen, dass sich frühere Messungen auf aktuelle auswirken. Bei Leistungsmaßen sind Lern- und Übungseffekte zu erwarten, die die Validität und Genauigkeit folgender Messungen betreffen. In diesem Zusammenhang scheint es fast ein glücklicher Umstand, dass die durch den IAT ausgelösten Nachwirkungen von andersartiger Beschaffenheit sind. So wäre beispielsweise zu erwarten, dass die gemeinsame Varianz verschiedenartiger Messverfahren besonders bei Testwiederholungen zu einem relativen Gewinn an Merkmalsvarianz führt, da die spezifischen Fehlerkomponenten der jeweiligen Maße eben nicht gemeinsam sind.

Methodenspezifische Varianz

Auf die praktischen Konsequenzen reliabler methodenspezifischer Varianz wurde an verschiedenen Stellen schon eingegangen. Die zentrale Folgerung aus diesen Befunden ist, dass die gemeinsame Varianz von IATs, die vorgeblich die gleichen Inhalte messen, nur teilweise inhaltspezifisch ist, da sich Korrelationen in vergleichbarer Höhe auch finden, wenn keine inhaltliche Überlappung vorliegt. Dieser Punkt betrifft insbesondere die Interpretation von Schätzungen der Reliabilität. Zumindest implizit findet sich in vielen Arbeiten die Annahme, dass die gemeinsame Varianz von Testhälften und Testwiederholungen des IAT auf deren Inhaltsgleichheit zurückgeht. Dass dies nicht zwingend der Fall ist, zeigen die niedrigen Korrelationen zwischen teilweise hochgradig reliablen impliziten und expliziten Maßen des gleichen Konstrukts (Bosson et al., 2000; Cunningham et al., 2001; Egloff &

Schmukle, 2002; Greenwald et al., 1998; Karpinski & Hilton, 2001; Rudman et al., 1999; Swanson et al., 2001). Dass dieses Befundmuster, wie von einigen Autoren vermutet, auf Unterschiede zwischen den durch implizite und explizite Verfahren erfassten Konstrukten zurückgeht, ist wenig überraschend. Nichtsdestoweniger überraschend ist aber, dass dies nicht auf Unterschiede der Konstrukte zurückgehen muss, die man zu messen glaubt, sondern auch auf scheinbar gänzlich irrelevante Fähigkeitsparameter; nämlich solche, die an exekutiven Kontrollfunktionen beteiligt sind.

Letztlich ist dies jedoch ein trivialer Befund, da schon die Definition der Reliabilität in der klassischen Testtheorie (z. B. Gulliksen, 1950; Lord, 1959; Lord & Novick, 1968) klar macht, dass Reliabilität zwar notwendig, nicht jedoch hinreichend für eine hohe Validität einer Messung ist. Reliabilität beschreibt in der klassischen Testtheorie lediglich die Reproduzierbarkeit einer Messung, nicht ihre Gültigkeit. Die Tatsache, dass *wahre* Varianz gemäß dieser Definition nicht identisch mit *Merkmalsvarianz* ist, scheint aber nicht immer angemessen in die Deutung von Befunden des IAT einzufließen. Gerade die gute Reliabilität des IAT wird häufig als Argument für dessen psychometrische Überlegenheit ins Feld geführt (Bosson et al., 2000; Fazio & Olson, 2003; Greenwald & Nosek, 2001; Nosek et al., 2002). Die hier berichteten Befunde zu methodenspezifischer Varianz könnten eine Neubewertung dieser Einschätzung zur Folge haben.

Die Prüfung von Techniken zur Kontrolle methodenspezifischer Varianz zeigt, dass Aufgabenwechselkosten am Zustandekommen methodenspezifischer Varianz mindestens beteiligt sind. Gleichzeitig scheint der optimierte Scoring-Algorithmus (Greenwald et al., 2003) einer guter Ansatz zur Kontrolle methodenspezifischer Varianz zu sein. Aus theoretischer Sicht könnte eine Erklärung für das gute Abschneiden des optimierten Algorithmus sein, dass bei einer Relativierung der IAT-Scores an der Standardabweichung der Latenzen der Einfluss extrem langsamer Reaktionen auf die Scores verringert wird - möglicherweise also genau solcher Reaktionen, die stark von Aufgabenwechselkosten beeinflusst sind. Eine Prüfung dieser Frage liegt jedoch außerhalb der Zielsetzungen der vorliegenden Arbeit. Festhalten kann man, dass die von Greenwald et al. (2003) verwendete Brute-Force Methodologie durchaus geeignet sein kann, Probleme zu lösen, die weder bekannt noch theoretisch beschrieben sind. Eine zweigleisige Strategie, die diesen rein empirisch-exploratorischen Ansatz mit einer theoretischen Fundierung kombiniert, scheint eine geeignete Strategie zur Lösung der praktischen Probleme des IAT zu sein.

Ausblick

Bei einer Einordnung der vorliegenden Ergebnisse zum IAT muss man sich zunächst einmal klar machen, dass spezifische IATs in verschiedenen Anwendungskontexten kaum mehr gemeinsam haben als etwa verschiedene Likert-Skalen. Die erfolgreiche Validierung eines IAT in einem Anwendungskontext lässt sich ebenso wenig verallgemeinern wie das Scheitern in einem anderen. Ebenso wie für explizite Verfahren gilt, dass ein spezifischer IATs nur so gut ist wie das verwendete Material. Die Effekte des IAT mit geometrischen Objekten oder die Effekte von Figur-Grund Asymmetrien (Rothermund & Wentura, 2001) zeigen in diesem Zusammenhang, dass assoziative Verknüpfungen keineswegs notwendig für das Auftreten von IAT-Effekten sind. Spezifische Effekte schlecht gewählter Exemplare, wie sie etwa in der Auswirkung von „Crosscategory Associations“ (Steffens & Plewe, 2001) sichtbar werden, können die Effekte zusätzlich verzerren. Die übliche Praxis, Exemplare der Antwortkategorien nach Augenschein auszuwählen, sollte also einer sorgfältigeren Selektion der Materialien weichen. Aus dieser Sicht kann man wohl folgern, dass der IAT einen relativ direkten Blick in die Black-Box erlauben *kann*, wenn auch mit der Einschränkung, dass nicht immer sicher ist, was durch diesen Blick ans Licht gebracht wurde. Denn auch im besten Fall gilt, dass IAT-Effekte auf Redundanzen und Überlappungen zwischen vollkommen aufgabenirrelevanten Merkmalen zurückgehen können. Unangenehm ist an diesem Sachverhalt vor allem, dass gegenwärtig keine verbindlichen Kriterien dafür bekannt sind, ob die Resultate eines IAT irrelevante Aspekte des Verfahrens und der verwandten Materialien widerspiegeln oder die zu messenden Inhalte (Fiedler et al., 2003). Zwar gilt dies für die klassischen expliziten Maße in gleichem Maße, die Sensibilität für diesen Sachverhalt ist jedoch entschieden größer, als dies für die scheinbar objektiven indirekten Verfahren der Fall ist.

Ob und inwiefern ein spezifischer IAT tatsächlich die Stärke assoziativer Verknüpfungen abbildet, ist also in jedem Fall im Vorfeld empirisch abzusichern. Insbesondere wenn man auf eine individuelle Interpretation der Resultate abzielt, sollte eine sorgfältige Validierung des spezifischen IAT vorausgehen. Leider fehlen häufig gerade in solchen Bereichen, in denen eine Anwendung indirekter Verfahren besonders attraktiv ist, zuverlässige Aussenkriterien. Wenn etwa soziale Erwünschtheit eine Rolle spielt oder eine geringe introspektive Zugänglichkeit der Inhalte angenommen werden muss, sind hiervon häufig auch Verhalten und explizite Verfahren betroffen. In diesem Zusammenhang stellt sich natürlich die Frage, was die Bedeutung eines Merkmals sein soll, das sich weder im Verhalten noch in der Selbst- und Fremdwahrnehmung manifestiert.

Aus der Position, dass explizite und implizite Verfahren unterschiedliche Konstrukte erfassen (Greenwald et al., 1998; Wilson et al., 2000), erwächst in diesem Zusammenhang ein schwer zu lösendes Dilemma. Einerseits kann man aus dieser Sicht plausibel erklären, warum die Korrelationen zwischen impliziten und expliziten Verfahren in vielen Untersuchungen gering ausfallen oder ganz ausbleiben. Andererseits entsteht der Eindruck, dass diese Erklärung in vielen Fällen post hoc erfolgt, und signifikante Korrelationen ebenso willkommen als erwartungsgemäß gelten würden. Dass dies so ist, zeigt sich schon daran, dass Greenwald et al. (2003) die Höhe der Korrelation zwischen IAT und expliziten Maßen als positives Kriterium bei der Optimierung des Scoring-Algorithmus verwenden. In keinem Fall darf jedoch das Ausbleiben von Zusammenhängen zwischen Maßen, die nominell das Gleiche messen, als Hinweis auf die Überlegenheit eines der beteiligten Verfahren ausgelegt werden. Beim gegenwärtigen Stand der Forschung scheinen implizite Verfahren gerade in sensiblen Bereichen eher ungeeignet.

Damit sei nicht gesagt, dass implizite Verfahren in einem Anwendungskontext keinen Beitrag leisten könnten, denn auch die gegenwärtig üblichen expliziten Verfahren unterliegen spezifischen und schwer kontrollierbaren Verzerrungen (Crowne & Marlowe, 1960; Greenwald, 1980; Gur & Sackeim, 1979; Holden et al., 2001; Orne, 1962; Paulhus, 1998; Rosenberg, 1965; Taylor & Brown, 1988; Tedeschi et al., 1971; Viswesvaran & Ones, 1999). Die schlechte Nachricht, dass auch implizite Verfahren Schwächen aufweisen und Probleme mit sich bringen, kann deshalb kaum überraschen. Die gute Nachricht besteht möglicherweise darin, dass die Probleme des IAT und anderer impliziter Verfahren anders geartet sind als diejenigen expliziter. In diesem Zusammenhang ist wenig plausibel, dass der methodische und methodologische Fortschritt im Bereich impliziter Verfahren zu einer neuen „*Industry of research*“ führen wird, die mehr oder weniger exklusiv auf impliziten Verfahren basiert, wie Greenwald und Banaji (1995) prognostizieren. Vielmehr ist zu vermuten und zu hoffen, dass sich der Blick auf individuelles Verhalten und Erleben in Zukunft vermehrt auf Konvergenz und Divergenz von Ergebnissen *verschiedenartiger* Verfahrensklassen und Methoden stützen wird, deren jeweils spezifische Stärken und Schwächen einander ergänzen und in der Zusammenschau ein vollständigeres Bild ergeben.

9 Literatur

Allport, D. A., Styles, E. A. & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Hrsg.), *Attention and performance XV* (421-452). Cambridge, MA: MIT Press.

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

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

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

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Hrsg.), *Handbook of social cognition, Vol. 1: Basic processes; Vol. 2: Applications (2nd ed.)* (1-40). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Hrsg.), *The automaticity of everyday life: Advances in social cognition, Vol. 10. Advances in social cognition* (1-61). Mahwah, NJ: Lawrence Erlbaum Associates.

Bargh, J. A., Chaiken, S., Govender, R. & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893-912.

Bargh, J. A., Chaiken, S., Raymond, P. & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 104-128.

Bellezza, F. S., Greenwald, A. G. & Banaji, M. R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments and Computers*, 18, 299-303.

Blair, I. V., Ma, J. E. & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, 81, 828-841.

Blümke, M. & Friese, M. (2004). *Do irrelevant features of stimuli influence IAT effects?* Unveröffentlichtes Manuskript, Universität Heidelberg.

Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) - Handanweisung*. Göttingen: Hogrefe.

Bosson, J. K., Swann, W. B., Jr. & 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.

Brauer, M., Wasel, W. & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology*, 4, 79-101.

Bredenkamp, J., Klein, K. M., von Hayn, S. & Vaterrodt, B. (1988). Gedächtnispsychologische Untersuchungen eines Rechenkünstlers. *Sprache and Kognition*, 7, 69-83.

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

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

Cepeda, N. J., Kramer, A. F. & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, 37, 715-730.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 240-248.

Costa, P. T. & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory. Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349-354.

Cunningham, W. A., Preacher, K. J. & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *121*, 163-170.

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.

Dasgupta, N., McGhee, D. E., Greenwald, A. G. & Banaji, M. R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology*, *36*, 316-328.

De Houwer, J. (1998). The semantic Simon effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *3*, 683-688.

De Houwer, J. (2001a). Contingency awareness and evaluative conditioning: When will it be enough? *Consciousness and Cognition: An International Journal*, *10*, 550-558.

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

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

De Houwer, J. & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion, 12*, 45-61.

De Houwer, J., Hermans, D. & Eelen, P. (1998). Affective Simon effects using facial expressions as affective stimuli. *Zeitschrift für Experimentelle Psychologie, 45*, 88-98.

De Jong, P. J., van den Hout, M. A., Rietbroek, H. & Huijding, J. (2003). Dissociations between implicit and explicit attitudes toward phobic stimuli. *Cognition and Emotion, 17*, 521-545.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B. & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510-540.

Draine, S. C. & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286-303.

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). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Hrsg.), *Research methods in personality and social psychology* (74-97). Thousand Oaks, CA: Sage Publications.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115-141.

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 uses. *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 (Hrsg.), *Dual process theories in social psychology* (97-116). New York, NY: Guilford Press.

Fiedler, K., Messner, C. & Blümke, M. (2003). *Unresolved problems with the "I", the "A" and the "T": Logical and psychometric critique of the Implicit Association Test (IAT)*. Unveröffentlichtes Manuskript, Universität Heidelberg.

Fitts, P. M. & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 199-210.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A. & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion*, 14, 61-92.

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., Ehrenberg, K., Banse, R., Zukova, J. & Klauer, K. C. (2003). 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.

Gopher, D. (1996). Attention control: Explorations of the work of an executive controller. *Cognitive Brain Research*, 5, 23-38.

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task-set switching. In S. Monsell & J. Driver (Hrsg.), *Control of cognitive processes: Attention and Performance XVIII* (331-356). Cambridge, MA: MIT Press.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35, 603-618.

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

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A. & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3-25.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A. & Rosier, M. (2000). Prologue to a unified theory of attitudes, stereotypes, and self-concept. In J. P. Forgas (Hrsg.), *Feeling and thinking: The role of affect in social cognition. Studies in emotion and social interaction, second series* (308-330). New York, NY: Cambridge University Press.

Greenwald, A. G., Draine, S. C. & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, *273*, 1699-1702.

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.

Gulliksen, H. (1950). *Theory of mental tests*. Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Gur, R. C. & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology*, *37*, 147-169.

Hoeck-Gradenwith, E. (1959). Rohrschachtest mit kriminellen Psychopathen. *Psychologische Rundschau*, 105-118.

Hofmann, W., Gschwendner, T., Gawronski, B., Le, H. & Schmitt, M. (2003). *A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures*. Unveröffentlichtes Manuskript, Universität Trier.

Holden, R. R., Wood, L. L. & Tomaszewski, L. (2001). Do response time limitations counteract the effect of faking on personality inventory validity? *Journal of Personality and Social Psychology*, *81*, 160-169.

Hsieh, S. (2002). Two-component processes in switching attention: A study of event-related potentials. *Perceptual and Motor Skills*, *94*, 1168-1176.

Huth, A. (1949). Möglichkeiten und Grenzen des Rohrschach-Tests. *Grenzgebiete Medizin*, 430-433.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *14*, 81.

Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin*, *30*, 22-34.

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

Kawakami, K. & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin*, *27*, 212-225.

Klauer, K. C. (1998). Affective priming. In W. Stroebe & M. Hewstone (Hrsg.), *European Review of Social Psychology*, Vol. 8 (67-103). Chichester, England: John Wiley & Sons.

Klauer, K. C., Ehrenberg, K. & Wegener, I. (2003). Crossed categorization and stereotyping: Structural analyses, effects patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology*, *39*, 332-354.

Klauer, K. C. & Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *54a*, 727-751.

Klauer, K. C. & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Hrsg.), *The psychology of evaluation: Affective processes in cognition and emotion* (7-49). Mahwah, NJ: Lawrence Erlbaum Associates.

Klauer, K. C., Rossnagel, C. & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 246-255.

Klauer, K. C. & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology*, 75, 1155-1178.

Klauer, K. C., Wegener, I. & Ehrenberg, K. (2002). Perceiving minority members as individuals: The effects of relative group size in social categorization. *European Journal of Social Psychology*, 32, 223-245.

Klotz, W. & Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 976-992.

Koch, I. (2003). The role of external cues for endogenous advance reconfiguration in task switching. *Psychonomic Bulletin & Review*, 10, 488-492.

Kornblum, S. & Lee, J. W. (1995). Stimulus-response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 855-875.

Kramer, A. F., Hahn, S. & Gopher, D. (1999). Task coordination and aging: Explorations of executive control processes in the task switching paradigm. *Acta Psychologica*, 101, 339-378.

Kray, J., Li, K. Z. H. & Lindenberger, U. (2002). Age-related changes in task-switching components: The role of task uncertainty. *Brain and Cognition*, 49, 363-381.

Kray, J. & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, 15, 126-147.

Kühnen, U., Schiebl, U., Bauer, N., Paulig, N., Pöhlmann, N. & Schmidhals, N. (2001). How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. *Zeitschrift für Experimentelle Psychologie*, *48*, 135-144.

Logan, G. D. & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 575-599.

Logan, G. D. & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*, 393-434.

Lord, F. M. (1959). Inferences about true scores from parallel test forms. *Educational and Psychological Measurement*, *19*, 331-336.

Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Oxford, England: Addison-Wesley.

Los, S. A. (1996). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica*, *94*, 145-188.

Los, S. A. (1999). Identifying stimuli of different perceptual categories in pure and mixed blocks of trials: Evidence for stimulus-driven switch costs. *Acta Psychologica*, *103*, 173-205.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163-203.

Maison, D., Greenwald, A. G. & Bruin, R. (2001). The Implicit Association Test as a measure of implicit consumer attitudes. *Polish Psychological Bulletin*, *2*, 61-79.

Marsh, K. L., Johnson, B. T. & 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.

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

McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio (Hrsg.), *(Prejudice, discrimination, and racism (91-125))*. San Diego, CA: Academic Press.

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.

McCrae, R. R. & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality, 57*, 17-40.

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

Meiran, N., Chorev, Z. & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*, 211-253.

Meyer, D. E. & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Mierke, J. & Klauer, K. C. (1999). *Switch costs as an explanation for association effects in the IAT*. Vortrag auf der Meeting of European Social Cognition Network, Nijmegen, The Netherlands.

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.

Monsell, S., Yeung, N. & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research/Psychologische Forschung, 63*, 250-264.

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

Musch, J. (2000). *Affektives Priming: Kongruenzeffekte bei der evaluativen Bewertung*: Inauguraldissertation zur Erlangung des Grades eines Doktors der Philosophie an der Fakultät für Sozial- und Verhaltenswissenschaften der Ruprecht-Karls-Universität Heidelberg.

Nebel, A., Strack, F. & Schwarz, N. (1989). Tests als Treatment: Wie die psychologische Messung ihren Gegenstand veraendert. *Diagnostica*, 35, 191-200.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Hrsg.), *Basic processes in reading: Visual word recognition* (264-336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Neumann, O. & Klotz, W. (1994). Motor responses to nonreportable, masked stimuli: Where is the limit of direct parameter specification? In C. Umiltà & M. Moscovitch (Hrsg.), *Attention and performance 15: Conscious and nonconscious information processing. Attention and performance series* (123-150). Cambridge, MA, USA: The Mit Press.

Neumann, R., Ebert, M., Gabel, B., Guelsdorff, J., Krannich, H., Lauterbach, C., et al. (1998). Vorurteile zwischen Bayern und Norddeutschen: Die Anwendung einer neuen Methode zur Erfassung evaluativer Assoziationen in Vorurteilen. *Zeitschrift für Experimentelle Psychologie*, 45, 99-108.

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. & 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 web site. *Group Dynamics*, 6, 101-115.

Nosek, B. A., Greenwald, A. G. & Banaji, M. R. (2003). *Understanding and using the Implicit Association Test: II. Method variables and construct validity*. Unveröffentlichtes Manuskript, University of Virginia at Charlottesville, VA.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776-783.

Ostendorf, F. (1994). Zur Taxonomie deutscher Dispositionsbegriffe. In W. Hager & M. Hasselhorn (Hrsg.), *Handbuch deutschsprachiger Wortnormen* (382-441). Göttingen: Hogrefe.

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

Pashler, H. & Baylis, G. C. (1991a). Procedural learning: I. Locus of practice effects in speeded choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 20-32.

Pashler, H. & Baylis, G. C. (1991b). Procedural learning: II. Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 33-48.

Paulhus, D. L. (1998). *The balanced inventory of desirable responding*. Toronto, Buffalo: Multi-Health Systems.

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.

Roefs, A. & Jansen, A. (2002). Implicit and explicit attitudes toward high-fat foods in obesity. *Journal of Abnormal Psychology*, *111*, 517-521.

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

Rosenberg, M. J. (1965). When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology*, *1*, 28-42.

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

Rubinstein, J. S., Meyer, D. E. & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 763-797.

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.

Salthouse, T. A., Fristoe, N., McGuthry, K. E. & Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging*, 13, 445-461.

Schwibbe, M., Räder, K., Schwibbe, G., Borchardt, M. & Geiken-Pophanken, G. (1994). Zum emotionalen Gehalt von Substantiven, Adjektiven und Verben. In W. Hager & M. Hasselhorn (Hrsg.), *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe.

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

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Hrsg.), *Stimulus response compatibility: An integrated perspective. Advances in psychology, Vol. 65* (31-86). Amsterdam, Netherlands: North-Holland.

Skinner, B. F. (1974). *About behaviorism*. Oxford, England: Alfred A. Knopf.

Spector, A. & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology*, 89, 669-679.

Stangor, C., Lynch, L., Duan, C. & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62, 207-218.

Steffens, M. C. & Buchner, A. (2003). Situational and transsituational components of implicit attitudes towards gay men. *Experimental Psychology*, 50, 33-48.

Steffens, M. C., Jelenec, P., Anheuser, J., Goergens, N. K., Hülsebusch, T., Lichau, J., et al. (2004). *A two-factor model of reaction time differences in the Implicit Association Test*. Unveröffentlichtes Manuskript, Universität Trier.

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.

Strack, F. (1994). Response processes in social judgment. In R. S. Wyer, Jr. (Hrsg.), *Handbook of social cognition, Vol. 1: Basic processes; Vol. 2: Applications* (2 ed., 287-322). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*.

Sudevan, P. & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 89-103.

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.

Taylor, S. E. & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210.

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.

Tedeschi, J. T., Schlenker, B. R. & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, Vol. 26, 685-695.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529-554.

Verhaeghen, P. & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience and Biobehavioral Reviews*, 26, 849-857.

Viswesvaran, C. & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*, 197-210.

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*, 919-934.

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

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

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1-27.

Anhang A: Material des Blumen-Insekten Einstellungs-IAT in Experiment 1

Positive Wörter: Leckerbissen, Echtheit, Erlösung, Freundlichkeit, Lebhaftigkeit, Nektar, Gemüt, Diamant, Weise, Säugling, Photographie, Humor, Duft, Leichtigkeit, Musikinstrument, Eingebung, Bücherei, Fröhlichkeit, Busen, Aufnahmefähigkeit, Flexibilität, Charme, Redlichkeit, Herzlichkeit

Negative Wörter: Missbilligung, Torheit, Nässe, Verheimlichung, Mutmaßungen, Einbuße, Stoß, Ermüdung, Eigensinn, Notfall, Beteuerungen, Phrase, Makel, Hirngespinnst, Untauglichkeit, Bettler, Zwielflicht, Rückschritt, Meineid, Eigenmächtigkeit, Trugschluß, Orkan, Banalität, Dumpfheit

Blumen: Schwertlilie, Flieder, Lilie, Maiglöckchen, Butterblume, Primel, Nelke, Edelweiß, Feuerlilie, Orchidee, Chrysantheme, Rose, Iris, Gänseblümchen, Vergissmeinnicht, Veilchen, Krokus, Sonnenblume, Tulpe, Stiefmütterchen, Pusteblume, Mohn, Klatschmohn, Seerose

Insekten: Springspinne, Moskito, Wespe, Wespenspinne, Borkenkäfer, Mücke, Fliege, Termiten, Mistkäfer, Spinne, Fruchtfliege, Laus, Made, Stubenfliege, Kartoffelkäfer, Schabe, Hornisse, Küchenschabe, Wanze, Blatthornkäfer, Blattlaus, Floh, Blattwanze, Kakerlake

Anhang B: Zusätzliches Material des Blumen-Insekten Einstellungs-IAT in den folgenden Experimenten

Positive Wörter: Grundrecht, Gelassenheit, Barmherzigkeit

Negative Wörter: Abrechnung, Einbildung, Fragwürdigkeit

Blumen: Narzisse, Pfingstrose, Alpenveilchen

Insekten: Ameise, Rüsselkäfer, Nachtfalter

Anhang C: Material des Selbst-Andere Extraversions-IAT

Selbst: Ich, selbst, mir, mich, eigen

Andere: Sie, ihnen, euer, ihr, andere

Extraversion: Wagemutig, personenorientiert, vergnügungsfreudig, ungehemmt, draufgängerisch, redselig, dominierend, leidenschaftlich, energiegeladen, expressiv, mitteilsam, dominant, geschwätzig, gesprächig

Introversion: Zurückhaltend, schüchtern, zurückgezogen, still, distanziert, einzelgängerisch, reserviert, unauffällig, scheu, schweigsam, unspontan, verschwiegen, vorsichtig, ruhig