



TECHNISCHE UNIVERSITÄT
CHEMNITZ

**Measures of Implicit Self-Esteem.
Psychometric Properties and the Prediction of Anxious,
Self-Confident and Defensive Behavior**

Verfahren zur Erfassung impliziter Selbstwertschätzung.
Psychometrische Eigenschaften und die Möglichkeiten zur Vorhersage von
ängstlichem, selbstsicherem und defensivem Verhalten

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der Philosophischen Fakultät der
Technischen Universität Chemnitz

von Almut Rudolph, geboren am 30.06.1980 in Frankfurt/Oder

Gutachter: Prof. Dr. Astrid Schütz, Prof. Dr. Udo Rudolph

Eingereicht am 10.06.2009, Disputation am 21.09.2009

Online verfügbar unter <http://archiv.tu-chemnitz.de/pub/2009/0151>

Danksagung

Ich möchte einigen Personen danken, die mich auf ihre individuelle Art unterstützt haben und damit auf einzigartige Weise an der Entstehung dieser Arbeit beigetragen haben.

Ich danke Prof. Dr. Astrid Schütz, die mein Interesse für das wissenschaftliche Arbeiten und die implizite Selbstwertschätzung weckte, als ich noch studierte. Ich danke ihr für die Unterstützung und die konstruktive Kritik, aber auch für das Vertrauen, das sie mir entgegenbrachte. Nicht zuletzt gebührt ihr Dank für die unterstützenden Arbeitsbedingungen. So war es mir möglich, in einen intensiven Austausch mit anderen Wissenschaftlern zu treten, der weit über die besuchten Workshops, Tagungen und Auslandsaufenthalte hinweg reichte.

Ich danke Prof. Dr. Udo Rudolph für seine Bereitschaft, diese Arbeit zu begutachten.

Meiner Kollegin Dr. Michela Schröder-Abé danke ich für die intensive Zusammenarbeit. Jederzeit beantwortete sie mir kleinere und größere Fragen, regte mich zum Nachdenken an, gab nützliche Hinweise, und erkundigte sich nach meinem persönlichen Wohlbefinden. Gern erinnere ich mich an die vielen Stunden, die wir uns vermeintlich unlösbaren Problemen mit längeren Syntaxen und komplizierten Quelltexten widmeten.

Dank geht auch an die vielen studentischen Hilfskräfte, die mich in allen Aufgaben engagiert unterstützten, vor allem bei Durchführung und Auswertung der Experimente, aber auch bei kurzfristigen Korrekturen. Ebenso danke ich den Probanden und Probandinnen für ihre Teilnahme an den Studien.

Ich danke Anne Kraft für die unzähligen Stunden am Telefon, in denen wir die Höhen und Tiefen des fernen Alltags mitzerleben versuchten. Die wenigen, doch intensiven gemeinsamen Erlebnisse gaben mir den erforderlichen Abstand zur Arbeit.

Mein ganz besonderer Dank aber gilt meinen Großeltern – meiner Oma Lotte und meinem Opa Erwin. Ich danke ihnen für ihre uneingeschränkte und bedingungslose Unterstützung. Meiner Mutti Christine danke ich für ihre Liebe und das entgegengebrachte Vertrauen. Die aufbauenden Anrufe halfen über manche frustrierende Situation hinweg. Gleichzeitig sorgte sie mit ihrem Engagement und ihren schier unversiegbaren Kräften für das Wohlergehen und die Gesundheit meiner Großeltern, wofür ich ihr sehr dankbar bin.

Besonderer Dank gilt Andrej, der mit seiner ruhigen Art für den nötigen Ausgleich sorgte und mir jederzeit persönlichen Rückhalt gab.

Inhaltsverzeichnis

1	Einführung.....	1
1.1	Selbstwertschätzung.....	3
1.1.1	Explizite Selbstwertschätzung.....	3
1.1.2	Implizite Selbstwertschätzung.....	4
1.1.3	Selbstwert in den Zwei-Prozess-Theorien der Informationsverarbeitung.....	5
1.2	Verfahren zur Erfassung impliziter Selbstwertschätzung.....	7
1.2.1	Initials Preference Task.....	8
1.2.2	LOAD Task.....	9
1.2.3	Implicit Association Test.....	10
1.2.4	Weiterentwicklungen.....	12
1.3	Psychometrische Befunde zu Verfahren zur Erfassung impliziten Selbstwerts.....	14
1.3.1	Reliabilität.....	14
1.3.2	Konvergente und diskriminante Validität.....	15
1.3.3	Prädiktive Validität.....	15
1.3.4	Eigene Befunde zu Reliabilität und Validität.....	16
1.4	Diskrepanzen zwischen expliziter und impliziter Selbstwertschätzung.....	18
1.4.1	Fragile und verletzte Selbstwertschätzung.....	19
1.4.2	Eigene Befunde zu Selbstwertdiskrepanzen.....	19
1.5	Kritische Bemerkungen, offene Fragen und Ausblick.....	20
1.5.1	Grenzen des Implicit Association Test.....	20
1.5.2	Neuere indirekte Verfahren.....	22
1.5.3	Vorhersagemodelle expliziter und impliziter Selbstwertschätzung.....	24
1.5.4	Diskrepante Selbstwertschätzung.....	25
1.6	Fazit.....	26
1.7	Literatur.....	27
2	Zeitschriftenbeiträge.....	35
	Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. & Sedikides, C. (2008). Through a glass, less darkly: Reassessing the convergent and discriminant validity in measures of implicit self-esteem. <i>European Journal of Psychological Assessment</i> , 24, 273-281.....	36
	Rudolph, A., Schröder-Abé, M., Riketta, M. & Schütz, A. (2009). Easier when done than said! Implicit self-esteem predicts observed or spontaneous behaviour, but not self-reported or controlled behaviour. Manuscript submitted for publication.....	56
	Schröder-Abé, M., Rudolph, A., Wiesner, A. & Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. <i>International Journal of Psychology</i> , 42, 174-183.....	78

1 Einleitung

Die Fähigkeit zur Selbstreflexion und Selbstbewertung wird häufig als zentrales Merkmal des Menschseins beschrieben. Individuen können sich selbst zum Gegenstand ihrer Gedanken und Beobachtungen machen. Diese Besonderheit macht sich die Persönlichkeitspsychologie zu nutze, um die Bewertung zur eigenen Person zu erfassen – die Selbstwertschätzung – die eine hohe Bedeutung für menschliches Erleben und Verhalten hat.

Allerdings haben Selbstbeschreibungen, mit denen die so genannte explizite Selbstwertschätzung erhoben wird, einige Schwachstellen. So versuchen Menschen bei ihrem Gegenüber einen bestimmten Eindruck zu hinterlassen, der nicht notwendigerweise mit ihrem Selbstbild übereinstimmen muss. Dabei kann der Eindruck in verschiedenen Situationen günstiger oder bescheidener ausfallen. Diese Täuschung muss nicht immer willentlich geschehen, sondern kann auch mangelnder Introspektion zuzuschreiben sein.

Diesen kritischen Punkten versucht die implizite Selbstwertschätzung zu entgehen. Eine Besonderheit dieser automatischen und überlernten Bewertung der eigenen Person ist, dass sie nicht aus Selbstberichten, sondern mittels indirekter Verfahren aus Reaktionen auf selbstbezogene Stimuli geschlossen wird.

Das zunehmende Interesse an der indirekten Erfassung von Selbstwertschätzung führte zu vielen Neu- und Weiterentwicklungen. Bevor indirekte Verfahren allerdings in persönlichkeitspsychologischer Forschung zum Einsatz kommen, sollte ihre psychometrische Güte nachgewiesen werden. Erst dann kann beurteilt werden, ob sie einen zusätzlichen Nutzen zu bestehenden Selbstberichtsverfahren bieten, und wenn ja welchen.

Die vorliegende Arbeit umfasst drei empirische Beiträge, die diesen Fragen nachgehen. Da die Artikel bei internationalen Zeitschriften eingereicht wurden, sind sie in englischer Sprache verfasst. Die vorangestellten Abschnitte geben eine Einführung in das Themengebiet der Selbstwertschätzung und deren Erfassung mit indirekten Verfahren. Vor allem aber dienen sie dazu, die empirischen Befunde der Beiträge zu integrieren.

Die ersten beiden Artikel widmen sich der Überprüfung der psychometrischen Qualität indirekter Verfahren. Der bereits veröffentlichte Beitrag von Rudolph, Schröder-Abé, Schütz, Gregg und Sedikides (2008) vergleicht die Reliabilität und Validität etablierter und

neuerer Verfahren. Ein besonderes Augenmerk liegt auf der Prüfung, ob verschiedene indirekte Verfahren konvergieren und ob explizite und implizite Selbstwertschätzung korrespondieren, wenn implizite Selbstwertschätzung reliabel erfasst wird. Außerdem werden strukturelle Ähnlichkeiten der indirekten Verfahren in die Betrachtungen mit einbezogen. Der eingereichte Beitrag von Rudolph, Schröder-Abé, Riketta und Schütz (2009) komplettiert die Untersuchungen zur Güte der Verfahren. Der Fokus dieses Beitrags liegt im Nachweis der prädiktiven Validität verschiedener indirekter Verfahren. Dabei werden Kriterien herangezogen, die nicht im Selbstbericht erhoben werden. Im Detail wird eine doppelte Dissoziation zwischen expliziter und impliziter Selbstwertschätzung geprüft.

Der bereits erschienene Beitrag von Schröder-Abé, Rudolph, Wiesner und Schütz (2007) rundet diese Arbeit ab, da er sich auch einem anwendungsbezogenen Aspekt der Persönlichkeitspsychologie widmet. Untersucht wird der Zusammenhang zwischen defensiven Verhaltensweisen und Selbstwertdiskrepanzen, also der Kombination expliziter und impliziter Selbstwertschätzung.

1.1 Selbstwertschätzung

In der einschlägigen Literatur findet der Begriff des „Selbst“ eine vielseitige Verwendung. Das selbstbezogene Wissen stellt hierbei eines der am häufigsten untersuchten Konstrukte in der psychologischen Forschung dar (für einen Überblick, Leary & Tangney, 2003). Es vereint eine deskriptive Komponente, das Bild von der eigenen Person, und eine evaluative Komponente, die Bewertung dieses Bildes (Baumeister, 1998).

1.1.1 Explizite Selbstwertschätzung

Selbstwertschätzung wird definiert als das subjektive Empfinden des eigenen Wertes (Coopersmith, 1967; Schütz, 2003). Das Ergebnis unzähliger Untersuchungen lässt sich nur schwer kurz zusammenfassen, jedoch kann festgehalten werden, dass Selbstwert eine bedeutende Rolle im sozialen und persönlichen Funktionieren innehat. So zeigte sich, dass Personen mit hoher Selbstwertschätzung zufriedener (Diener, 1984), optimistischer (Taylor & Brown, 1988) und emotional stabiler (Judge, Erez, Thoresen & Bono, 2002) sind als Personen mit niedriger Selbstwertschätzung. Außerdem haben diese Personen bessere akademische Leistungen (Newbegin & Owens), schreiben sich eher Erfolge als Misserfolge zu (Blaine & Crocker, 1993) und bevorzugen Interaktionspartner, die sie selbst positiv beurteilen (Swann, Stein-Seroussi & Giesler, 1992). Diese Auswahl von Ergebnissen stellt nur einen kleinen Ausschnitt bisheriger Befunde dar und lässt erahnen, welche beeindruckende Anzahl von Studien bereits durchgeführt wurde. Dennoch wurden in letzter Zeit Stimmen laut, die auf Messproblematiken bei der Erhebung von Selbstwert hinweisen.

Selbstwertschätzung wird klassischerweise auf direktem Weg durch Selbstbeschreibungsfragebogen erfasst. In diesen wird um die Zustimmung zu positiven oder negativen selbstbezogenen Aussagen gebeten. International sehr häufig eingesetzt wird die Rosenberg-Skala (Rosenberg, 1965; deutsche revidierte Version von Collani & Herzberg, 2003), die die globale Selbstwertschätzung erfasst („Alles in allem bin ich mit mir selbst zufrieden.“). Einzelne Facetten des Selbstwerts können mit der Multidimensionalen Selbstwertkala (MSWS; Schütz & Sellin, 2006) unterschieden werden, die u.a. zwischen emotionaler, sozialer und leistungsbezogener Selbstwertschätzung differenziert.

Allerdings sind Selbstberichtsverfahren in der Persönlichkeitsdiagnostik grundsätzlich in ihrer Validität eingeschränkt. So kann die per Selbstbericht erfasste Selbstwertschätzung durch Selbstdarstellung (Schlenker, 1980), Selbsttäuschung (Paulhus, 1984) und Selbstignoranz (Wilson, Lindsey & Schooler, 2000) verzerrt sein.

In der diagnostischen Situation spielt insbesondere das Interesse, sich gegenüber dem Diagnostiker oder dem Auftraggeber positiv darzustellen, eine entscheidende Rolle. Der Schutz und die Erhöhung des eigenen Selbstwerts wird als ein wichtiges Motiv menschlichen Handelns beschrieben (Greenwald, 1980) und erfolgt häufig über Selbstdarstellung, dem Bemühen, anderen Menschen einen bestimmten Eindruck von der eigenen Person zu vermitteln (vgl. Tice & Baumeister, 1990): „Eigentlich bin ich wertlos, ich gebe aber lieber vor, großartig zu sein“. Auch Selbsttäuschung (Paulhus, 1984), die Motivation, sich selbst in einem günstigen Licht zu sehen, fördert entsprechende Antworten im Fragebogen: „Ich bin großartig, wenngleich jeder mich hasst“. Treffen die beiden eben aufgeführten Aspekte zu, erfassen Selbstwertfragebogen nicht nur das subjektive Empfinden des eigenen Wertes, sondern auch individuelle Unterschiede in diesen Selbstdarstellungs- und Selbsttäuschungsstrategien. Außerdem wird in neuerer Literatur der mangelnde Zugang zu eigenen Gedanken und Gefühlen beschrieben, der selbstberichteten Selbstwert verfälschen kann: „Wie schätze ich mich *wirklich* ein?“. Belege dafür liefern empirische Studien, die die Bedeutung von Denkprozessen, deren Inhalte dem Bewusstsein kaum zugänglich sind, belegen (Wilson et al., 2000).

1.1.2 Implizite Selbstwertschätzung

Implizite Selbstwertschätzung ist „the introspectively unidentified (or inaccurately identified) effect of the self-attitude on evaluation of self-associated and self-dissociated objects“ (Greenwald & Banaji, 1995, S. 11). Einige Autoren bringen den Aspekt des Unbewussten in diese Definition mit ein und beschreiben impliziten Selbstwert als unbewusste Assoziationen über die eigene Person (z.B. Koole & Pelham, 2003). Um sich von der Freud'schen Sichtweise des Unbewussten abzugrenzen und den aktuellen Erkenntnisstand darzulegen, werden an dieser Stelle die Begrifflichkeiten kurz voneinander abgegrenzt.

Die Untersuchung weniger bewusster Prozesse in der Persönlichkeitspsychologie ist relativ neu, allerdings schrieb schon Freud (1933) von der Existenz unbewusster mentaler Vorgänge. Dabei unterscheidet sich die Definition impliziter Selbstwertschätzung deutlich

von der Freud'schen Sichtweise der unbewussten Assoziationen. In seinen Theorien finden sich Vorstellungen von einem Unbewusstsein, in welchem Erinnerungen, Wünsche und Gedanken abgelegt werden, die Menschen von ihrem Bewusstsein fernzuhalten versuchen. Im letzten Jahrhundert jedoch wandelte sich dieses Verständnis und neuere Ansätze gehen von zwei parallel arbeitenden Systemen der Informationsverarbeitung aus, welche im nächsten Abschnitt vorgestellt werden.

Demnach ist der Begriff „implizit“ nicht mit „unbewusst“ gleichzusetzen. Vielmehr kann impliziter Selbstwert als die automatische Bewertung der eigenen Person verstanden werden. Automatizität ist dann gegeben, wenn mindestens eine der Bedingungen für kontrollierte Prozesse nicht erfüllt ist (Zielgerichtetheit, Intention, Aufmerksamkeit oder Bewusstheit; Bargh, 1994). Damit kann impliziter Selbstwert zwar unbewusste Prozesse umfassen, jedoch sind diese keine grundlegende Voraussetzung.

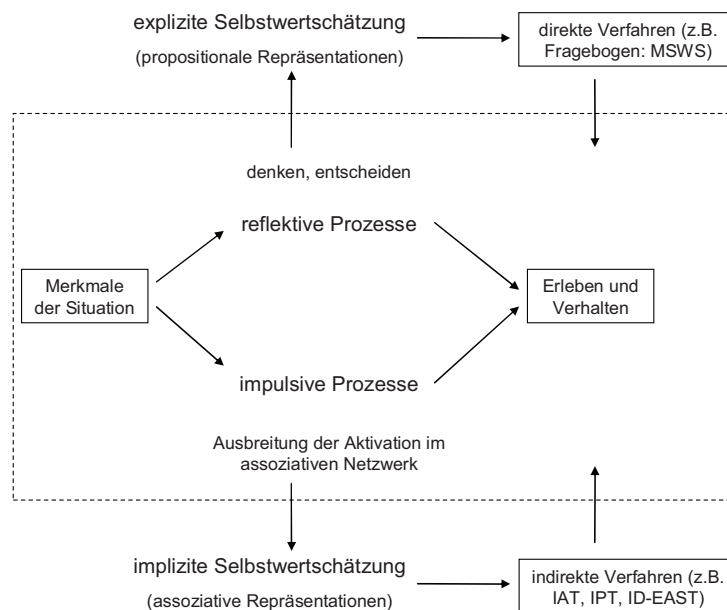
1.1.3 Selbstwertschätzung in den Zwei-Prozess-Modellen der Informationsverarbeitung

Die Zwei-Prozess-Modelle (Epstein, 1994; Fazio, 1990; Wilson et al., 2000) haben einen weiten Anwendungsbereich in Forschung der letzten Jahre gefunden. Mit dem unlängst vorgestellten Behavioral Process Model of Personality (BPMP; Back, Schmukle & Egloff, in press) wird ein umfassender theoretischer Rahmen geliefert, der die Zwei-Prozess-Modelle um den Aspekt der Persönlichkeit erweitert. Somit gibt dieses Modell den derzeitigen Wissenstand zum expliziten und impliziten Selbstwert umfassend wieder. Eine vereinfachte Darstellung ist in Abbildung 1 zu sehen.

In Anlehnung an das Reflective Impulsive Model (RIM; Strack & Deutsch, 2004) werden zwei parallel arbeitende Systeme der Informationsverarbeitung unterschieden – ein reflektives System und ein impulsives System. Im reflektiven System werden selbstrelevante Informationen überlegt und kontrolliert verarbeitet, was sich in propositionalen Repräsentationen des Selbst niederschlägt: „Ich kann sehr zufrieden mit meinen fachlichen Fähigkeiten sein“. Der Inhalt des reflektiven Systems ist somit dahingehend begrenzt, dass er lediglich selbstrelevante Gedanken umfasst, die intentional und bewusst sind. Diese können, wenn genügend kognitive Kapazitäten für die Transformation der Informationen in ein propositionales Format zur Verfügung stehen, durch Befragung erfasst werden. Explizite Selbstwertschätzung bezieht sich demnach auf einen Prozess im reflektiven System, der mittels direkter Methoden erfasst werden kann. Im impulsiven System hingegen werden selbstrelevante Informationen automatisch und

intuitiv verarbeitet. Durch die gleichzeitige Aktivierung von Konzepten bahnt sich mit jeder Wiederholung eine stärker werdende assoziative Repräsentation. Wird beispielsweise eine Person häufig im Alltag für ihre Leistungen gelobt, stärkt sich die assoziative Verknüpfung zwischen den Konzepten „ich“ und „leistungsfähig“. Allerdings werden diese nicht notwendigerweise in ein propositionales Format übersetzt, insbesondere dann, wenn keine kognitiven Kapazitäten zur Verfügung stehen. Demnach hat das impulsive System auch Zugang zu weniger bewussten und automatischen Gedanken, die mit wenigen Ressourcen abgerufen werden können. Implizite Selbstwertschätzung kann daher als Prozess im impulsiven System verstanden werden, der eine Verbindung zwischen dem Selbst und evaluativen Elementen herstellt. Daher wird impliziter Selbstwert nicht im Selbstbericht erhoben, sondern aus anderen Daten abgeleitet.

Abbildung 1. Implizite und explizite Selbstwertschätzung in den Zwei-Prozess-Modellen der Informationsverarbeitung (in Anlehnung an Back et al., in press)



1.2 Verfahren zur Erfassung impliziter Selbstwertschätzung

Wie in den vorangegangenen Abschnitten deutlich wurde, kann impliziter Selbstwert nicht im klassischen Selbstwertfragebogen erhoben werden, sondern wird mit indirekten Verfahren erfasst. Einige Autoren nutzen dabei die Begriffe indirekt und implizit gleichbedeutend, um auf Verfahren zur Erfassung impliziter Selbstwertschätzung Bezug zu nehmen (Wilson et al., 2000). Andere Autoren hingegen grenzen beide Bezeichnungen voneinander ab, da nicht die Verfahren sondern die den Testwerten zugrunde liegenden Eigenschaften implizit sind (z.B. Fazio & Olson, 2003). So spricht De Houwer (2006) sich dafür aus, dass der empirische Nachweis über die „implizite“ Funktionsweise entsprechender Verfahren erst erbracht werden muss. Um von einem impliziten Verfahren sprechen zu können, müsste erst nachgewiesen werden, dass die Probanden (1) nicht erkennen, dass ein bestimmtes Konstrukt gemessen wird, (2) keinen bewussten Zugang zu dem Konstrukt haben, und (3) keine Kontrolle über die Ausprägung des Messergebnisses haben.

Im überwiegenden Teil der für die vorliegende Arbeit gesichteten Literatur allerdings wurde nur gewährleistet, dass die Verfahren den Probanden nicht über den Inhalt der Messung informieren und dass nicht direkt nach dem interessierenden Konstrukt gefragt wird, so dass die Testwertermittlung nicht auf Selbstbeurteilung beruht. Die Verfahren zur Erfassung von Selbstwertschätzung werden daher im Folgenden als direkt oder indirekt bezeichnet, um auf die Eigenschaft der Messverfahren einzugehen, die entweder im Selbstbericht erhoben oder aus anderen Daten abgeleitet werden. Die Ergebnisse dieser Verfahren als auch die zugrunde liegenden Konstrukte werden entsprechend als explizit oder implizit bezeichnet.

Die indirekten Methoden, die zur Erhebung der impliziten Selbstwertschätzung eingesetzt werden, beruhen auf drei Annahmen. Ausgehend von den Zwei-Prozess-Modellen, begründet sich explizite Selbstwertschätzung, beispielsweise eine Selbstbeurteilung im Fragebogen, auf deliberative und kontrollierte Prozesse in der Informationsverarbeitung (Strack & Deutsch, 2004). Mangelt es hingegen an kognitiven Kapazitäten (z.B. bei akutem Stress), sind Individuen nicht in der Lage, ihre deliberativen Selbstbeurteilungen zu aktivieren bzw. ihren spontanen Selbstbewertungen gegenzusteuern. Daher wird angenommen, dass Selbstbeschreibungen, die unter kognitiver Beanspruchung abgegeben

werden, auf assoziative und impulsive Prozesse zurückgehen und somit implizite Selbstwertschätzung widerspiegeln (siehe LOAD Task).

Außerdem beurteilen Menschen selbstrelevante Objekte (z.B. die Buchstaben des eigenen Namens) im Durchschnitt positiver als nicht-selbstrelevante Objekte (Greenwald & Banaji, 1995). Da es individuelle Unterschiede in diesen Bewertungen gibt (Pelham & Hetts, 1999) und Individuen sich diesen Bewertungen nicht bewusst sind (Koole, Dijksterhuis & van Knippenberg, 2001), können Einstellungen zu einem selbstrelevanten Stimulus als implizite Einstellung zum Selbst beurteilt werden. Zur Erfassung impliziter Selbstwertschätzung lässt sich demnach heranziehen, in welchem Ausmaß Probanden selbstrelevante Stimuli, z.B. ihre Initialen, positiv bewerten (siehe Initials Preference Task).

Schließlich wird angenommen, dass bei der Präsentation eines Einstellungsobjektes die damit verbundene Bewertung automatisch aktiviert wird (Fazio, Sanbonmatsu, Powell & Kardes, 1986). Infolgedessen wird die Verarbeitung ähnlicher Objekte erleichtert bzw. die Verarbeitung unähnlicher Objekte erschwert. Daher können aus den entsprechenden Reaktionen auf Einstellungsobjekte Rückschlüsse auf die implizite Einstellung gezogen werden. Beispielsweise werden aus Reaktionen auf evaluative und selbstbezogene Stimuli automatische Selbstbewertungen abgeleitet, indem die Schnelligkeit erfasst wird, in der positive Stimuli mit dem Selbst assoziiert werden (siehe Implicit Association Test). Es wird davon ausgegangen, dass den meisten Personen die Assoziation von positiven und selbstrelevanten Stimuli (kongruente Bedingung) leichter fällt, als die Assoziation von negativen und selbstrelevanten Stimuli (inkongruente Bedingung). Somit sind die Reaktionszeiten in der kongruenten Bedingung schneller als in der inkongruenten Bedingung.

An diesen Prinzipien wird deutlich, dass individuelle Unterschiede in der Ausprägung von Selbstwertschätzung indirekt erschlossen werden: Aus Selbstbeurteilungen unter kognitiver Beanspruchung, aus Reaktionen auf selbstrelevante Stimuli oder über die Messung von Reaktionszeiten.

1.2.1 Initials Preference Task

Nuttin (1985) deckte erstmals das Phänomen des name-letter-effects auf, wonach Individuen die Buchstaben des eigenen Namens gegenüber den Buchstaben, die nicht im Namen vorkommen, präferieren. Einige Jahre später legten Greenwald und Banaji (1995)

mit ihrer Definition von implizitem Selbstwert – die sich in Reaktionen auf selbstrelevante Stimuli niederschlagende selbstbezogene Haltung – den Grundstein für die Nutzung des name-letter-effects als Maß für die implizite Selbstwertschätzung (Kitayama & Karasawa, 1997). Empirische Befunde belegten, dass (a) die Buchstaben des Namens mit dem Selbst verbunden sind (was beinahe tautologisch erscheint), (b) die Überbewertung der Namensbuchstaben sowohl in westlichen als auch in östlichen Kulturen vorkommt, (c) Personen sich dieser Verzerrung bei der Bewertung der Buchstaben des Alphabets nicht bewusst sind und (d) dieser Bias nicht allein durch andere Faktoren (z.B. mere exposure) erklärt werden kann. Für eine ausführliche Darlegung der Befunde wird auf Kohle und Pelham (2003) verwiesen.

Da der name-letter-effect für die Initialbuchstaben am stärksten ist (Koole et al., 2001), etablierte sich in den letzten Jahren der Einsatz der Initials Preference Task (IPT; Bosson, Swann & Pennebaker, 2000) zur indirekten Erfassung von Selbstwertschätzung. In dieser Aufgabe wird, nachdem die Probanden ihren Gefallen bezüglich eines jeden Buchstaben des Alphabets auf einer siebenstufigen Skala (1 = „mag ich überhaupt nicht“ bis 7 = „mag ich sehr“) zum Ausdruck gebracht haben, berechnet, wie stark sie ihre Initialen gegenüber den anderen Buchstaben des Alphabets bevorzugen. In der Literatur sind mindestens fünf verschiedene Algorithmen zur Berechnung der Initialenpräferenz zu finden, wobei aktuellen Befunden zufolge ein Algorithmus vorzuziehen ist, der sowohl individuelle Antworttendenzen als auch die Präferenz einzelner Buchstaben kontrolliert (für eine Übersicht, siehe LeBel & Gawronski, 2009).

Die Initialenpräferenz hängt darüber hinaus überzufällig häufig mit Verhaltensweisen wie der Wahl des Wohnorts, des Berufs und des Partners zusammen (Jones, Pelham, Mirenberg & Hetts, 2002; Pelham, Mirenberg & Jones, 2002). Nelson und Simmons (2008) zeigten zudem Zusammenhänge mit Abschlussnoten, allerdings nur für Studierende mit C- oder D-Initial (im britischen Notensystem steht D für ungenügend).

1.2.2 LOAD Task

Unlängst stellten Riketta und Gebauer (subm.) eine Selbstbeurteilungsaufgabe unter kognitiver Beanspruchung vor (LOAD Task). In dieser sollen die Probanden angeben, ob positive und negative Eigenschaftswörter sie beschreiben. Die Stimuli erscheinen randomisiert auf dem Bildschirm und die Zustimmung oder Ablehnung soll per Tastendruck abgegeben werden. Dabei wird die Erfassung spontaner Selbstbewertungen

gewährleistet, da die Probanden sich eine achtstellige Ziffer merken und am Ende richtig wiedergeben sollen. Die Differenz zwischen den positiven und den negativen Eigenschaftswörtern, denen zugestimmt wurde, wird als Maß für die implizite Selbstwertschätzung herangezogen. Hohe Werte geben an, dass sich der Proband mehr positive als negative Eigenschaften zuschreibt. Es wird davon ausgegangen, dass implizite Selbstwertschätzung erfasst wird, da der Mangel an kognitiven Ressourcen keine deliberative Selbstbeurteilung erlaubt.

Darüber hinaus bietet diese Selbstbeurteilungsaufgabe auch die Möglichkeit, explizite Selbstwertschätzung zu erfassen, nämlich dann, wenn sie ohne kognitive Beanspruchung durchgeführt wird. Der Vorteil dieser Aufgabe ist, dass sie vergleichbare Indizes liefert: Dieselbe Skala dient, unter Bedingungen, die entweder reflektive oder impulsive Antwortprozesse fördern, zur Erfassung expliziter oder impliziter Selbstwertschätzung.

1.2.3 Implicit Association Test

Greenwald und Kollegen stellten vor 10 Jahren im *Journal of Personality and Social Psychology* den Implicit Association Test (IAT; Greenwald, McGhee & Schwartz, 1998) vor, dessen „rasante Verbreitung in der Geschichte der Psychologie fast schon als einzigartig bezeichnet werden kann“ (Gawronski, 2006, S. 53).

IATs messen implizite Einstellungen durch Reaktionszeiten. Im Falle impliziten Selbstwerts (Greenwald & Farnham, 2000; Rudolph et al., 2008) wird per Reaktionsgeschwindigkeit die Assoziationsstärke zwischen *selbstrelevanten* bzw. *nicht-selbstrelevanten* Wörtern (Targetdimension) und evaluativ *angenehmen* bzw. *unangenehmen* Wörtern (Attributdimension) gemessen. Bei der ca. 15-minütigen Testdurchführung wird der IAT als Zuordnungsaufgabe vorgestellt. In fünf Blöcken sollen die Probanden Zielwörter per Tastendruck klassifizieren (siehe Abbildung 2 für eine veranschaulichende Darstellung und Beispielstimuli). Die Zielwörter erscheinen in der Mitte des Bildschirms und sollen möglichst schnell und fehlerfrei einer von zwei Kategorien zugeordnet werden. Wie in Abbildung 2 zu erkennen ist, handelt es sich bei den Blöcken 1, 2 und 4 um Übungsblöcke, während die Blöcke 3 und 5 zur Auswertung herangezogen werden. Dabei wird, vereinfacht gesprochen, die Differenz der Reaktionszeiten der kombinierten Blöcke gebildet und für die individuelle Reaktionszeit des Probanden kontrolliert (Greenwald, Nosek & Banaji, 2003). Der resultierende IAT Effekt gibt an, wie viel höher die mittlere Reaktionszeit in Block 5 (Klassifikation von

selbstrelevanten und unangenehmen Begriffen mit derselben Taste) im Vergleich zu Block 3 (Klassifikation von selbstrelevanten und angenehmen Begriffen mit derselben Taste) ist. Diese Reaktionsgeschwindigkeit wird als Hinweis darauf interpretiert, wie eng die automatische Assoziation der Begriffe ist, d.h. wie viel einfacher es für den Probanden ist, angenehme Begriffe mit dem Selbst zu verbinden als unangenehme. Es wird angenommen, dass Individuen mit hoher impliziter Selbstwertschätzung eine automatische Assoziation des Selbst mit positivem Affekt aufweisen. Daher antworten sie schneller wenn dieselbe Antworttaste für die selbstrelevante und die angenehme Kategorie steht, als wenn die selbstrelevante und die unangenehme Kategorie gekoppelt sind.

Abbildung 2. Die fünf Blöcke eines Implicit Association Test zur Erfassung von Selbstwertschätzung in Anlehnung an Greenwald und Farnham (2000)

	Kategorie linke Taste	Beispielstimuli	Kategorie rechte Taste
Block 1 Übung	ich ● ○	mein andere	nicht-ich ○ ●
Block 2 Übung	angenehm ● ○	Freude Leid	unangenehm ○ ●
Block 3 Test	ich oder angenehm ● ● ○ ○	mein Freude andere Leid	nicht-ich oder unangenehm ○ ○ ● ●
Block 4 Übung	nicht-ich ○ ●	mein andere	ich ● ○
Block 5 Test	nicht-ich oder angenehm ○ ● ● ○	mein Freude andere Leid	ich oder unangenehm ● ○ ○ ●

Anmerkung. ● = richtige Antwort ○ = falsche Antwort; die Stimuli werden wiederholt dargeboten

Von besonderer Wichtigkeit erscheinen beim Selbstwert IAT die Kontexteinflüsse bei der Interpretation des IAT Effektes (z.B. Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001) zu sein. So liefert der IAT lediglich ein *relatives* Maß für die Präferenz einer Kategorie im Vergleich zu einer anderen, jedoch kein Maß der *absoluten* Bewertung

jeweils einer Kategorie. Genauer gesagt wird die Bewertung der eigenen Person in Abhängigkeit von der Bewertung anderer Personen gemessen. Studien zeigten, dass die Beurteilung selbstrelevanter Stimuli von der Beurteilung nicht-selbstrelevanter Stimuli beeinflusst wird. So fand Karpinski (2004) einen Einfluss der nicht-selbstrelevanten Kategorie, wenn besonders positive oder negative Beispiele repräsentiert werden (z.B. Santa Claus). Ein hoher positiver IAT Effekt kann daher also drei Ursachen haben: (a) die positive Bewertung der einen Kategorie, (b) die negative Bewertung der anderen Kategorie, oder (c) eine Kombination beider Effekte. Derartige Mehrdeutigkeiten können im Bereich von Selbstwertschätzung von großer Bedeutung sein, da die Relation von Selbstbewertungen und Bewertungen anderer noch nicht vollständig geklärt ist (Schütz, 2005). Eine Lösung des Problems wird in der Wahl einer *unspezifischen* nicht-selbstrelevanten Kategorie, mit Stimuli wie fremd oder andere, gesehen (Pinter & Greenwald, 2005). Andere Autoren stellten Alternativen zum IAT vor, die absolute Selbstbewertungen erfassen.

1.2.4 Weiterentwicklungen

Um die Problematik, dass der Selbstwert IAT durch die Verwendung von jeweils zwei Targetkategorien (selbstrelevant vs. nicht-selbstrelevant) relative aber nicht absolute Präferenzen misst, grundsätzlich zu lösen, wurde der Single-Category IAT (SC-IAT; Karpinski & Steinman, 2006) entwickelt. Im SC-IAT werden nur die Reaktionen auf die selbstrelevante Targetkategorie erhoben. Somit werden in getrennten Blöcken die Assoziationen der eigenen Person mit positiven und mit negativen Attributstimuli erfasst. Die Bewertung anderer als Störgröße wird daher ausgeschlossen.

In der Extrinsic Affective Simon Task (EAST; De Houwer, 2003) und deren Weiterentwicklung der Identification EAST (ID-EAST; De Houwer & De Bruycker, 2007) werden die Stimuli der Attributdimension anhand der Valenz kategorisiert. Allerdings sollen die Stimuli der Targetdimension anhand eines irrelevanten Merkmals (z.B. Farbe oder Groß-/ Kleinschreibung) kategorisiert werden. In den Durchgängen, in denen die positiven und negativen Attributstimuli zugeordnet werden, nehmen die Antworttasten deren Valenz an. Diese Verknüpfung erleichtert bzw. erschwert die Zuordnung der Targetstimuli, da letztere anhand eines irrelevanten Merkmals kategorisiert werden sollen. Einerseits benötigen die Antworten gewöhnlich dann länger, wenn Stimulus und Valenz auf unterschiedlichen Antworttasten liegen. Erfolgt beispielsweise die Klassifikation

negativer Attributstimuli und blauer Targetstimuli mit der linken Taste, wird die Antwort für einen blau dargebotenen selbstrelevanten Stimulus inhibiert, da die Antworttaste extrinsisch mit negativer Valenz belegt ist. Andererseits wird angenommen, dass Targetstimuli dann schneller zugeordnet werden, wenn der Stimulus und die Valenz der Antworttaste übereinstimmen. Im Gegensatz zum IAT bieten die Verfahren, denen die Stimulus-Response-Kompatibilität zugrunde liegt (De Houwer, 2008), neben der absoluten Beurteilung der Targetkonzepte auch eine verringerte Anfälligkeit für mögliche Umkodierungen der Aufgabe (siehe Grenzen des Implicit Association Test): Diesen Verfahren liegt keine Blockabfolge zugrunde, so dass die Auswertung für die einzelnen Stimuluskategorien gesondert erfolgen kann. Neueren Befunden zufolge ist die ID-EAST der EAST vorzuziehen, da die Probanden für die Zuordnung der Stimuli zunächst deren Wortbedeutung identifizieren müssen, um die entsprechende Regel für Target- oder Attributstimuli anzuwenden.

Die Go/No-Go-Association Task (GNAT; Nosek & Banaji, 2001) ermöglicht ebenfalls die Messung automatischer Assoziationen in Bezug auf eine einzige Kategorie. Dabei benötigt die GNAT nur eine Antworttaste, da der Proband auf die Target- und Attributstimuli in den einzelnen Blöcken reagieren (Go) oder nicht reagieren (No-Go) soll. Gehört der dargebotene Stimulus einer der beiden Kategorien an, die in diesem Block relevant sind, muss die Testperson reagieren, ist dies nicht der Fall, soll sie die Zeit bis zur Darbietung des nächsten Stimulus verstreichen lassen. Dabei wird angenommen, dass es den Probanden leichter fällt, in den Durchgängen zu reagieren bzw. nicht zu reagieren, die kongruent zu ihrer Evaluation des Zielkonzepts sind. Bedingt durch den Versuchsablauf der GNAT können, neben der gängigen Berechnung automatischer Assoziationen (relativer Index), auch Treffer, Falsche Alarmer, Korrekte Ablehnung und Verpasser ermittelt werden (vgl. Signalentdeckungstheorie, Green & Swets, 1966). So wird beispielsweise die automatische Assoziation zur eigenen Person durch die Differenz der Treffer und der Falschen Alarmer in den Durchgängen berechnet, in denen die Targetstimuli dargeboten wurden (absoluter Index).

Die hier vorgestellten Alternativen bzw. Weiterentwicklungen des IAT erfassen die absolute Beurteilung eines Konzepts. Einige benötigen keine Kontrastkategorie, andere erlauben eine Berechnung absoluter Assoziationsstärken. Für die Selbstwertschätzung bedeutet dies, dass eine Erfassung von Selbstwert möglich ist, die unabhängig von der Bewertung anderer ist.

1.3 Psychometrische Befunde zu Verfahren zur Erfassung impliziter Selbstwertschätzung

Sucht man nach Angaben zur psychometrischen Güte der oben genannten indirekten Verfahren zur Erfassung von Selbstwertschätzung, muss man sich mit wenigen Befunden zufrieden geben. Dies liegt einerseits daran, dass die Weiterentwicklungen des IAT nur in wenigen Untersuchungen für die Erfassung von Selbstwertschätzung adaptiert wurden (De Houwer, 2003; Karpinski & Steinman, 2006). Andererseits wurden einige auch erst kürzlich vorgestellt. Auf Basis des aktuellen Wissensstands, der zunächst kurz zusammengefasst werden soll, konnte demnach kein Favorit unter der Vielzahl der Verfahren ausgemacht werden.

1.3.1 Reliabilität

Die wohl ausführlichste Untersuchung zur Güte indirekter Verfahren führten Bosson et al. (2000) durch. Die Autoren verglichen sieben indirekte Verfahren (u.a. sub- und supraliminales Priming, IAT, IPT). Allerdings bestätigten die Ergebnisse lediglich dem IAT und der IPT befriedigende *Retest-Reliabilitäten* über einen Zeitraum von einem Monat ($r_{tt} = .69$ bzw. $r_{tt} = .63$). Greenwald und Farnham (2000) berichteten für den IAT sogar nur einen Wert von $.52$ über einen Zeitraum von zehn Wochen.

Die *internen Konsistenzen* von Selbstwert IATs liegen meist recht hoch ($.80 < \alpha < .90$; Hofmann, Gawronski, Gschwendner, Le & Schmitt, 2005), während die IPT mit einem Alpha von $.50$ testtheoretisch gesehen nicht ausreichend reliabel erscheint (LeBel & Gawronski, 2009). Allerdings sollte dabei beachtet werden, dass nur zwei Items (Vor- und Nachnamensinitial) in die Berechnung der Testwerte eingehen. Betrachtet man die Inter-Itemkorrelationen der IPT, nimm diese durchschnittlich Werte größer $.30$ an (Bosson et al., 2000). Damit ist sie durchaus mit direkten Verfahren vergleichbar. Die MSWS, die mit $\alpha = .92$ eine sehr hohe interne Konsistenz aufweist, hat beispielsweise eine durchschnittliche Inter-Itemkorrelation von $.26$ (Rudolph et al., 2009; Studie 2). Die interne Konsistenz der IPT ist somit noch in weiteren Untersuchungen zu klären, allerdings macht ihre vergleichbar hohe Inter-Itemkorrelation sie – nicht nur aus messtheoretischer Sicht – zu einem interessanten Verfahren für die Erfassung impliziter Selbstwertschätzung (siehe auch LeBel & Gawronski, 2009).

1.3.2 Konvergente und diskriminante Validität

Zur Berechnung der *konvergenten Validität* werden die Zusammenhänge zwischen verschiedenen indirekten Verfahren untersucht. Diese sind in den meisten Fällen sehr gering und nicht signifikant (Bosson et al., 2000). Allerdings ist anzunehmen, dass einige dieser Befunde auf die mangelnde Reliabilität der Messverfahren zurückzuführen sind.

Ergebnisse bisheriger Studien zeigen zudem, dass auch zwischen impliziter und expliziter Selbstwertschätzung meist nur ein geringer Zusammenhang besteht (Bosson, Brown, Zeigler-Hill & Swann, 2003; Bosson et al., 2000; Greenwald & Farnham, 2000; Jordan, Spencer, Zanna, Hoshino-Browne & Correll, 2003). Allerdings wird dies in Abhängigkeit von der Höhe der Befunde in einigen Fällen als Hinweis auf *konvergente*, in anderen Fällen als Hinweis auf *diskriminante Validität* gewertet (Gawronski & Conrey, 2004), da eine theoretische Vorstellung über die Höhe des Zusammenhangs fehlt. So wird einerseits die geringe Korrelation zwischen implizitem und explizitem Selbstwert als Indikator für die Divergenz der gemessenen Konstrukte angenommen (Greenwald et al., 1998), andererseits wird eine hohe Konvergenz zwischen indirekten und direkten Verfahren als Validitätshinweis interpretiert (Greenwald et al., 2003). Diese widersprüchlichen Erklärungen sind problematisch.

1.3.3 Prädiktive Validität

Befunde zur Kriteriumsvalidität indirekter Verfahren sind verschwindend gering. Nur wenige Studien fanden signifikante Haupteffekte impliziten Selbstwerts, der unabhängig vom expliziten Selbstwert ist. So sagte implizite Selbstwertschätzung negative Emotionen im Alltag (Conner & Barrett, 2005), selbstberichtete depressive Symptome (Franck, De Raedt & De Houwer, 2007), und nonverbale Indikatoren von Angst in einem selbstwertrelevanten Interview vorher (Spalding & Hardin, 1999).

Eine Ursache für die wenigen Befunde könnte die Wahl der Kriterien sein. Ausgehend von den Zwei-Prozess-Modellen (Strack & Deutsch, 2004) kann angenommen werden, dass direkte und indirekte Verfahren eine differenzierte Verhaltensvorhersage ermöglichen. Verhalten ist das Ergebnis von Prozessen des impulsiven und des reflektiven Systems. Im reflektiven System resultiert Verhalten aus Entscheidungsprozessen, die kognitive Ressourcen benötigen. Da direkte Verfahren propositionale Repräsentationen des Selbst

erfassen, sollten sie bessere Prädiktoren für kontrollierte Verhaltensweisen sein. Andererseits werden Verhaltensschemata im impulsiven System auch bei geringen kognitiven Kapazitäten durch die Ausbreitung von Aktivierung im assoziativen Netzwerk ausgelöst. Demnach sollten indirekte Verfahren, die assoziative Repräsentationen des Selbst erfassen, bessere Prädiktoren für spontane Verhaltensweisen sein. Allerdings nutzte nur eine der oben genannten Untersuchungen konkretes Verhalten als Kriterium (Spalding & Hardin, 1999).

Weitere Studien fanden eine Interaktion von implizitem Selbstwert und Lebensereignissen (Steinberg, Karpinski & Alloy, 2007) oder Selbstwertbedrohungen (McGregor & Jordan, 2007). Die meisten Hinweise auf Kriteriumsvalidität indirekter Verfahren allerdings lieferten Studien zu Selbstwertdiskrepanzen (siehe Diskrepanzen zwischen expliziter und impliziter Selbstwertschätzung).

1.3.4 Eigene Befunde zu Reliabilität und Validität

Wie in der Darstellung bisheriger Befunde zur Güte indirekter Verfahren deutlich wurde, gibt es nur eine Studie, die verschiedene Verfahren miteinander verglichen hat (Bosson et al., 2000) und ihnen größtenteils geringe Reliabilitäten bestätigte. Allerdings waren die Weiterentwicklungen des IAT nicht Gegenstand dieses systematischen Vergleichs. Aus diesen Gründen können die oben aufgeworfenen Fragen, ob die Unreliabilität indirekter Verfahren (1) die mangelnde Konvergenz der indirekten Verfahren erklärt, und (2) die Zusammenhänge zwischen impliziter und expliziter Selbstwertschätzung beeinflusst, nicht vollständig beantwortet werden.

Der erste in dieser Arbeit eingebrachte Beitrag versucht, diese Fragen zu beantworten. Rudolph, Schröder-Abé, Schütz, Gregg und Sedikides (2008) verglichen in drei Studien die Reliabilitäten etablierter (IAT, IPT) und neuerer (SC-IAT, ID-EAST, GNAT) indirekter Verfahren. Ziel war es, herauszufinden, ob der Einsatz reliabler Maße höhere Zusammenhänge unter den indirekten Maßen als auch mit direkten Maßen ermöglicht. Außerdem wurde untersucht, ob konzeptuell ähnliche Verfahren – also Verfahren wie der SC-IAT, die ID-EAST und die GNAT, die absolute Selbstbewertungen messen – höher miteinander korrelieren als mit Verfahren, die relative Selbstbewertungen messen.

Die Ergebnisse wiesen dem IAT, dem SC-IAT und dem GNAT gute, der IPT und der ID-EAST befriedigende Reliabilitäten nach. Die EAST genügte diesen Anforderungen nicht. Interessant erschien außerdem, dass eine Verdoppelung der Testlänge der IPT durch eine

zweifache Darbietung und Beurteilung der Buchstaben des Alphabets (D-IPT), zu einer Steigerung der Reliabilität führte ($\alpha_{IPT} = .51$, $\alpha_{D-IPT} = .69$). Somit kann angenommen werden, dass die mangelnde interne Konsistenz der IPT teilweise auf die geringe Testlänge zurückzuführen ist. Weitere Forschung erscheint vielversprechend.

Wie schon bei Bosson et al. (2000) fanden sich nur geringe Korrelationen zwischen direkten und indirekten Verfahren. Allerdings wurde ein Zusammenhang zwischen konzeptuell ähnlichen Verfahren gezeigt. Dieser zeigte sich zwischen zwei indirekten Verfahren, die absolute Selbstbewertungen erfassen, dem SC-IAT und der ID-EAST. Weitere Hinweise auf die Bedeutung struktureller Ähnlichkeit deuteten die höheren Zusammenhänge zwischen den relativen Indizes der GNAT und des IAT an, während die Zusammenhänge zwischen dem absoluten Index der GNAT und dem IAT geringer ausfielen.

Somit gibt dieser Beitrag einerseits praktische Empfehlungen für die reliable Erfassung impliziter Selbstwertschätzung und liefert andererseits Hinweise auf die Funktionsweise indirekter Verfahren. So scheint nicht die Vielschichtigkeit des Konstrukts des impliziten Selbstwerts für die geringe konvergente Validität verantwortlich zu sein, vielmehr die Struktur der indirekten Verfahren.

Betrachtet man bisherige Forschung zur prädiktiven Validität indirekter Verfahren, fällt auf, dass unter den wenigen Studien lediglich in einem konkreten Verhalten vorhersaget wurde (Spalding & Hardin, 1999). Daher geht der zweite Beitrag der Frage nach, warum, trotz gegebener Reliabilität und Konstruktvalidität (zumindest einiger) indirekter Verfahren, in den letzten Jahren kaum Hinweise für deren Kriteriumsvalidität erbracht wurden.

Rudolph, Schröder-Abé, Riketta und Schütz (2009) untersuchten, ob die mangelnden Belege für die prädiktive Validität indirekter Verfahren auf die Wahl der Kriterien zurückzuführen ist. Hierfür wurden Kriterien jenseits von Selbstbericht herangezogen. Dies waren objektive Beurteilungen von selbstsicherem und spontanem Verhalten in einem selbstrelevanten Interview und in einer Selbstbeschreibungsaufgabe (Gebrauch von ängstlichkeitsrelevanten Wörtern, Beurteilung durch den Interviewer und durch einen Beobachter).

In den beiden Studien kamen der IAT, der implizite Selbstwertschätzung mittels Reaktionszeiten erfasst, und die LOAD Task, die Selbstbeurteilungen unter kognitiver Beanspruchung erhebt, zum Einsatz. Die Ergebnisse von drei Strukturgleichungsmodellen

bestätigten eine doppelte Dissoziation: Indirekte Verfahren sagten beobachtetes ängstliches oder spontanes selbstsicheres Verhalten vorher, das direkte Verfahren nicht vorhersagen konnten; Direkte Verfahren sagten selbstberichtetes ängstliches oder kontrolliertes selbstsicheres Verhalten vorher, das indirekte Verfahren nicht vorhersagen konnten.

Somit liefert diese Studien erstmals Befunde zur prädiktiven Validität verschiedener indirekter Verfahren: Implizite Selbstwertschätzung sagte selbstsicheres Verhalten vorher, dass explizite Selbstwertschätzung nicht vorhersagen konnte. Interessanterweise konnte die doppelte Dissoziation direkter und indirekter Verfahren repliziert werden.

1.4 Diskrepanzen zwischen expliziter und impliziter Selbstwertschätzung

Die Annahme von Diskrepanzen zwischen expliziter und impliziter Selbstwertschätzung erscheint aus drei Gründen gerechtfertigt. Zunächst werden in der Literatur Befunde zu hoher expliziter Selbstwertschätzung zunehmend als uneinheitlich und verwirrend kritisiert, da diese beispielsweise sowohl mit psychischer Gesundheit als auch mit sozial unverträglichem Verhalten einhergeht (z.B. Baumeister, Campbell, Krueger & Vohs, 2003). Außerdem haben, wie oben ausführlicher beschrieben, bisherige Studien kaum direkte Zusammenhänge von implizitem Selbstwert und anderen Konstrukten gefunden. Schließlich war die Korrelation impliziter und expliziter Selbstwertschätzung signifikant niedriger als die entsprechenden Korrelationen bei anderen Konstrukten (Hofmann et al., 2005). Das dies nicht nur auf messtheoretische Ursachen zurückzuführen ist, bestätigten Rudolph et al. (2008). Vielmehr scheinen bei Konstrukten wie dem Selbstwert, die stärker durch personale Faktoren, wie Selbstdarstellung oder Selbsttäuschung beeinflusst sind, inhaltlich bedeutsame Diskrepanzen in den Messwerten direkter und indirekter Verfahren aufzutreten.

Die Kombination expliziter und impliziter Selbstwertschätzung bietet hier die Möglichkeit, eine Differenzierung in kongruente und diskrepante Selbstwertschätzung vorzunehmen. Während bei *kongruenter Selbstwertschätzung* die Höhe expliziter und impliziter Selbstwertschätzung übereinstimmt (hoch/hoch oder niedrig/niedrig), klaffen diese bei *diskrepanter Selbstwertschätzung* auseinander (hoch/niedrig oder niedrig/hoch).

1.4.1 Fragile und verletzte Selbstwertschätzung

Im bisherigen Interesse der Forschung stand hauptsächlich die als *fragile Selbstwertschätzung* bezeichnete Kombination von hohem explizitem und niedrigem implizitem Selbstwert. Man geht davon aus, dass hinter dem äußerlich positiven Selbstbild bei einigen Personen Selbstzweifel verborgen sein können, die nicht notwendigerweise bewusst sein müssen. Ein Bewusstwerden der Selbstzweifel (beispielsweise durch Misserfolge oder Ablehnung) könnte zu defensiven Reaktionen führen, um die positiven Selbstbilder zu schützen bzw. aufrecht zu erhalten. Bereits mehrere Studien lieferten Hinweise, die diese Annahmen bestätigten. So war fragile Selbstwertschätzung im Vergleich zu den anderen Kombinationen von Selbstwert mit besonders hoher Selbstaufwertung (Bosson et al., 2003), Defensivität (Jordan et al., 2003; McGregor, Nail, Marigold & Kang, 2005) und Vorurteilen (Jordan, Spencer & Zanna, 2005) verbunden. Wenig Aufmerksamkeit wurde der zweiten Variante diskrepanter Selbstwertschätzung in bisheriger Forschung gewidmet. Die Kombination von geringer expliziter und hoher impliziter Selbstwertschätzung, von uns als verletzte Selbstwertschätzung bezeichnet, ging allerdings auch mit erhöhter Defensivität und selbsterhöhenden Tendenzen einher, wie die genauere Betrachtung der Ergebnisse von Jordan et al. (2003) und Bosson et al. (2003) zeigte. Die Erklärung, die herangezogen wurde, war eine Puffer-Hypothese: Hohe implizite Selbstwertschätzung wurde als Ressource angesehen, da sie die Wiederherstellung positiver Selbstberichte bei Personen mit niedrigem explizitem Selbstwert motiviert (Jordan et al., 2003) bzw. die negativen Effekte niedrigen expliziten Selbstwerts abschwächt (Bosson et al. 2003).

1.4.2 Eigene Befunde zu Selbstwertdiskrepanzen und Defensivität

Der dritte Beitrag dieser Arbeit überprüft eine Alternativerklärung zu der eben angeführten Puffer-Hypothese, wonach hohe implizite Selbstwertschätzung in Kombination mit geringer expliziter Selbstwertschätzung dysfunktional ist. Somit wird angenommen, dass Selbstwertdiskrepanzen, gleich welcher Variante, Stressoren darstellen, die beispielsweise zu selbstwertschützenden bzw. defensiven Reaktionen führen.

In zwei Studien untersuchten Schröder-Abé, Rudolph, Wiesner und Schütz (2007) den Zusammenhang zwischen Selbstwertdiskrepanzen und defensiven Reaktionen. Die

eingesetzten indirekten Verfahren unterschieden sich hinsichtlich ihrer zugrunde liegenden Annahmen (der IAT als Reaktionszeitmaß, die IPT als Beurteilungsmaß selbstrelevanter Initialen) und zählten in den vorangegangenen Beiträgen zu den Verfahren mit den besten psychometrischen Eigenschaften. Um den vorgestellten eigenen Befunden auch in der Wahl der Kriterien gerecht zu werden, wurden zwei unterschiedliche Maße für die Erfassung von Defensivität gewählt. In einem Selbstbeschreibungsverfahren wurde die Interpretation ambivalenter Aussagen erfasst (Studie 1). Außerdem wurde soziale Ablehnung mittels eines positiven bzw. negativen Feedbacks manipuliert. Die Dauer, in der die Personen das Feedback lasen, wurde als ein nicht-reaktives Verhaltensmaß für Defensivität herangezogen (Studie 2). Es wurde angenommen, dass defensives Verhalten sich in einer kurzen Lesezeit niederschlägt, da negative Rückmeldungen vermieden werden. Die Befunde beider Studien bestätigen, dass Personen mit fragiler als auch verletzter Selbstwertschätzung die defensivsten Reaktionen zeigten.

1.5 Kritische Beurteilung, offene Fragen und Ausblick

Die Ergebnisse der vorliegenden Studien sind bedeutsam, liefern sie doch erstmals eine umfassende und aktualisierte Untersuchung der psychometrischen Eigenschaften indirekter Verfahren. Der folgende Abschnitt diskutiert allgemeine Einschränkungen der vorliegenden Studien und gibt einen Ausblick auf weitere Fragen, an denen zukünftige Forschung ansetzen könnte.

1.5.1 Grenzen des Implicit Association Test

In der vorliegenden Arbeit wurde der IAT, neben der IPT, in jedem der drei Artikel eingesetzt. Damit zählt er nicht nur im Selbstwertbereich zu den indirekten Verfahren, die am häufigsten eingesetzt werden. Auch wenn der IAT grundsätzlich als reliabel und valide angesehen werden kann, werden im folgenden Abschnitt seine Probleme diskutiert und, falls vorhanden, deren Lösungsmöglichkeiten vorgestellt.

So wurde gezeigt, dass IATs verfälscht werden können (Fiedler & Blümke, 2005), wengleich dies nicht im selben Ausmaß wie bei direkten Verfahren möglich ist (Asendorpf, Banse & Mücke, 2002; Banse, Seise & Zerbes, 2001; Egloff & Schmukle, 2002; Steffens, 2004).

Auch die Zuweisung einer Kategorie zur linken oder rechten Antworttaste oder die Händigkeit der Probanden wurden als problematisch diskutiert, allerdings konnte deren Einfluss auf die Messung empirisch nicht nachgewiesen werden (Greenwald & Nosek, 2001). Dafür scheinen weitere externe Einflussgrößen identifiziert, die die Höhe des IAT Effekts beeinflussen, wie zum Beispiel die wohl nicht ganz unabhängig voneinander zu betrachtenden Faktoren der kognitiven Flüssigkeit und des Alters der Testpersonen: Probanden, die die Aufgabe recht langsam bearbeiteten bzw. älter waren, hatten höhere IAT Effekte als Probanden, die Aufgabe schneller bearbeiteten bzw. jünger waren (Greenwald & Nosek, 2001; McFarland & Crouch, 2002). Außerdem zeigte sich ein Zusammenhang zwischen IAT Effekt und der Häufigkeit der Durchführung: Die IAT Effekte wurden kleiner, je häufiger die Testpersonen zuvor einen IAT durchgeführt hatten (Greenwald & Nosek, 2001).

Den bisher genannten externen Einflussgrößen versucht der neue Berechnungsalgorithmus von Greenwald et al. (2003) zu begegnen, der auch in den vorliegenden Studien eingesetzt wurde. Während der herkömmliche Algorithmus die Differenz aus den logarithmierten Reaktionszeiten beider kombinierter Blöcke bildete und lediglich extrem langsame und schnelle Reaktionszeiten transformierte (Greenwald et al., 1998), bezieht der neue Algorithmus auch die durchschnittliche Reaktionszeit eines jeden Probanden in die Berechnung des IAT Effekts mit ein.

Ein weiterer bedeutender Einwand betrifft die interne Validität des IAT. Bis heute sind die dem IAT zugrunde liegenden Prozesse nicht hinreichend geklärt, was die Interpretation des IAT Effekts als Stärke von automatischen Assoziationen in Frage stellt (Greenwald et al., 1998).

Rothermund und Wentura (2001) führten Figur-Grund-Kompatibilitäten als eine zusätzliche Varianzquelle an. Demnach bildet eine der Kategorien die 'Figur' (zumeist die negativ bewertete oder weniger saliente), während die Kontrastkategorie den 'Grund' darstellt. Sollen nun im entsprechenden Block des IAT beide 'Figuren' mit derselben Taste klassifiziert werden, fällt diese Reaktion leichter. Langsamere Reaktionszeiten hingegen resultieren in den Blöcken, in denen die entsprechenden 'Figuren' der Dimensionen mit unterschiedlichen Tasten zu klassifizieren sind.

Außerdem wurden unterschiedliche Reaktionsschwellen für die Unterschiede in den Reaktionszeiten der kombinierten Blöcke angeführt (Brendl, Markman & Messner, 2001). So wird angenommen, dass die Probanden in der inkongruenten Bedingung ihre Reizschwelle höher ansetzen als in der kongruenten Bedingung. Da die Anpassung

bewusst und intentional erfolgen kann, widerspricht dies der Erfassung automatisch aktivierter Assoziationen im IAT.

Schließlich argumentierten andere Autoren, dass individuelle Unterschiede in der Fähigkeit zum Aufgabenwechsel mit den Assoziationsstärken im IAT konfundiert sind (Back, Schmukle & Egloff, 2005; Mierke & Klauer, 2001). Somit kann durch die Blockabfolge im IAT ein beachtlicher Anteil der Messwertvarianz auf Aufgabenwechselkosten beruhen.

Das genaue Verhältnis nicht-assoziativer Prozesse und tatsächlich gemessener automatischer Assoziationen im Selbstwert IAT zu bestimmen, erscheint schwierig. Empirische Befunde, die zeigen, dass sich die Rangfolge der Probanden durch die oben genannten Prozesse unsystematisch ändern, stehen für den Selbstwert IAT noch aus. Allerdings sei angemerkt, dass auch hier der neue Berechnungsalgorithmus einen Teil der Störgrößen minimieren kann (Mierke & Klauer, 2003).

Einen Ansatz, die unterschiedlichen Prozesse im IAT zu identifizieren, stellten Conrey und ihre Kollegen vor (Conrey, Sherman, Gawronski, Hugenberg & Groom, 2005) vor. Das allgemeine Messmodell des IAT isoliert vier Teilleistungen – automatische Aktivierung einer Assoziation, kategoriale Diskrimination des Stimulus, bewusste Kontrolle der Assoziation und allgemeine Reaktionstendenz – und kann zudem deren Anteile im IAT schätzen. Es ist anzunehmen, dass ein Großteil der Befunde zur Validität des IAT auf die automatische Aktivierung einer Assoziation zurückgeht. Die aufgezeigten Probleme der internen Validität des IAT hingegen beruhen auf einen der drei anderen Prozesse des Modells. Aus messtheoretischer Sicht mag es ungewöhnlich erscheinen, dass das Modell die Messkomponenten integriert, die widersprüchliche Ergebnisse zur Validität des IAT liefern. Allerdings könnte gerade dadurch der IAT als besonders erfolgversprechendes Verfahren angesehen werden. Denn auch im Alltag ist Verhalten eine Konfundierung unterschiedlichster Prozesse, die der IAT womöglich besonders gut abzubilden vermag (externe Validität). Eine Anwendung dieses Modells in zukünftigen Untersuchungen könnte eine Antwort auf die Frage geben, welchen Anteil nicht-assoziative Prozesse im Selbstwert IAT tatsächlich haben.

1.5.2 Neuere indirekte Verfahren

Die Ergebnisse des ersten systematischen Vergleichs indirekter Verfahren haben viele Fragen über die Konvergenz und Divergenz von impliziter und expliziter

Selbstwertschätzung aufgeworfen (Bosson et al., 2000). Aber diese Untersuchung inspirierte die Entwicklung neuerer indirekter Verfahren, die die Mängel bestehender Methoden ausgleichen sollten. Für einen aktuelleren Vergleich adaptierten Rudolph, Schröder-Abé, Schütz, Gregg und Sedikides (2008) einige dieser Verfahren für die Erfassung von Selbstwert. Allerdings war es nicht möglich, alle Neuentwicklungen in die vorliegenden Untersuchungen mit einzubeziehen, da einige erst nach Beginn der Datenerhebung publiziert wurden. So wurden außerdem die Affect Misattribution Procedure (Payne, Cheng, Govorun & Stewart, 2005) vorgestellt, aber auch weitere Varianten des IAT, wie der Brief IAT (Sriram & Greenwald, in press), der Single Block IAT (Teige-Mocigemba, Klauer & Rothermund, 2008) und die Implicit Association Procedure (Schnabel, Banse & Asendorpf, 2006). Gebauer, Riketta, Broemer und Maio (2008) präsentierten zudem die Name-Liking Task.

Weiterhin wurden *unipolare* IATs als Variante des IAT untersucht (Jordan et al., 2003). Diese kontrastieren die Selbstbewertungen nicht mit den Bewertungen anderer, sondern vergleichen sie mit einer neutralen Kategorie. Die nicht-selbstrelevante Targetkategorie eines unipolaren IAT enthält, statt der üblichen Stimuli (z.B. andere, ihr, euch), neutrale Stimuli (z.B. es, das, ein). Interessanterweise zeigten Korrelationen mit direkten Verfahren, dass der unipolare IAT höher mit diesen korrelierte als der Standard IAT (Schütz, Schröder-Abé & Rudolph, 2005). Somit scheint das neue Verfahren eine Erfassung impliziter Selbstwertschätzung zu ermöglichen, die weniger durch die Bewertung der anderen Kategorie verzerrt ist. Wenn eine, wie es scheint, minimale Änderung der Stimuli die Zusammenhänge der Verfahren derartig beeinflussen kann, erbringt möglicherweise die systematische Untersuchung *affektiver* und *evaluativer* IAT weitere interessante Hinweise. Beide Versionen unterscheiden sich hinsichtlich der Stimuli der Attributdimension. Im affektiven IAT werden emotionsgetönte Begriffe (z.B. Gesundheit, Gift) dargeboten, beim evaluativen IAT positiv und negativ konnotierte selbstbeschreibende Adjektive (z.B. höflich, hässlich). Im Großteil der Literatur und in der vorliegenden Arbeit findet sich die affektive Version des Selbstwert IAT wieder. Der Einsatz positiver und negativer Eigenschaften in einem evaluativen IAT allerdings würde eine Differenzierung impliziten Selbstwerts erlauben, wie dies auch bei den Facetten des expliziten Selbstwerts der Fall ist. Evaluative IATs unter der Verwendung von Eigenschaftswörtern lieferten im Bereich der Big Five Persönlichkeitsdimensionen bereits vielversprechende Ergebnisse (Back et al., in press; Steffens & Schulze König, 2006).

Insofern steht die Überprüfung der Reliabilität und Validität einer Vielzahl von neuen Verfahren noch aus. Dabei könnte das Beste der indirekten Verfahren aus einem unmittelbaren Vergleich mehrerer Methoden hervorgehen oder aber ein bestimmtes indirektes Verfahren stellt seine Nützlichkeit unter Beweis, indem es ein relevantes Kriterium vorhersagt. Für eine angemessene Beurteilung der Güte indirekter Verfahren ist zunächst einmal aber ein besseres Verständnis der zugrunde liegenden Prozesse sinnvoll. Hierbei scheinen die unterschiedlichen Strukturen der indirekten Verfahren eine bedeutende Rolle zu spielen (Rudolph et al., 2008). Sind diese in einer fundierten theoretischen Annahme formuliert, werden Befunde auch nicht widersprüchlich interpretiert (vgl. konvergente und diskriminante Validität). Zudem könnte die Kenntnis der Prozeduren dabei hilfreich sein, dasjenige unter den Verfahren zu finden, das für eine bestimmte Fragestellung am ehesten geeignet ist.

1.5.3 Vorhersagemodelle expliziter und impliziter Selbstwertschätzung

Ergebnisse zur prädiktiven Validität impliziter Selbstwertschätzung sind selten. Dabei ist die Kriteriumsvalidität eines der wichtigsten Belege für die Güte eines neuen Verfahrens. Die vorliegenden Studien erbringen Hinweise für die prädiktive Validität von impliziter Selbstwertschätzung und von Selbstwertdiskrepanzen. Rudolph et al. (2009) replizierten die doppelte Dissoziation direkter und indirekter Verfahren mit Haupteffekten für explizite und implizite Selbstwertschätzung. Schröder-Abé et al. (2007) sagten Defensivität in zwei Studien durch die Interaktion von expliziter und impliziter Selbstwertschätzung vorher. Die Frage, die sich womöglich stellt, ist: Warum wurden im erstgenannten Artikel keine Interaktionseffekte und im letztgenannten Artikel keine Haupteffekte gefunden?

Zur Erklärung werden die von Perugini (2005) vorgestellten Vorhersagemodelle herangezogen. Diese wurden in den vorliegenden Untersuchungen zur prädiktiven Validität berücksichtigt, da sie klären, welche Verhaltensweisen unter welchen Bedingungen vorhergesagt werden. Demnach liefert ein *interaktives Modell* von explizitem und implizitem Selbstwert eine Vorhersage, wenn sowohl reflektive als auch impulsive Prozesse bei der Aktivierung von Verhaltensschemata beteiligt sind. Andererseits kann eine *doppelte Dissoziation* angenommen werden, in der explizite und implizite Selbstwertschätzung differenzierte Verhaltensvorhersagen ermöglichen. Dafür müssen Verhaltensweisen erfasst werden, die entweder auf reflektive oder auf impulsive Prozesse zurückgehen.

Dieser Überblick verdeutlicht, dass die gewählten Kriterien den Geltungsbereich der verschiedenen Modelle determinieren. Somit sind die in Rudolph et al. (2009) und Schröder-Abé et al. (2007) gefundenen Ergebnisse nicht unvereinbar. Vielmehr zeigen sie die Notwendigkeit, in zukünftiger Forschung die vorherzusagenden Verhaltensweisen sorgfältig auszuwählen und gegebenenfalls unterschiedliche Vorhersagemodelle zu testen.

1.5.4 Diskrepante Selbstwertschätzung

In der Studie von Schröder-Abé et al. (2007) fanden sich die defensivsten Reaktionen für Personen mit fragiler als auch verletzter Selbstwertschätzung. Dementsprechend stellen Selbstwertdiskrepanzen, gleich welcher Variante, Stressoren dar, die insbesondere in selbstwertbedrohlichen Situationen, zu selbstwertschützenden Verhalten und defensiven Reaktionen führen. Da sich dieser Zusammenhang allerdings in weniger bedrohlichen Situationen nur für verletzte Selbstwertschätzung zeigte, sollten weitere Untersuchungen klären, welche Unterschiede zwischen fragiler und verletzter Selbstwertschätzung bestehen. Eine Vermutung wäre, dass andauernde Gefühle der Unzulänglichkeit das Wohlbefinden beeinträchtigen.

Erste Hinweise liefert eine Studie von Schröder-Abé, Rudolph und Schütz (2007), die verletzte Selbstwertschätzung im Zusammenhang mit psychischer und physischer Gesundheit sowie dem Umgang mit Emotionen untersuchte. Die Ergebnisse zeigten, dass Personen mit verletztem Selbstwert mehr Unterdrückung von Ärger, mehr Nervosität und mehr Krankheitstage berichten als Personen mit kongruent niedriger Selbstwertschätzung. Außerdem attribuierten sie negative Ereignisse eher auf internale, stabile und globale Ursachen, zeigten also einen depressiveren Attributionsstil.

Eine weitere offene Frage betrifft das Zustandekommen derartiger Selbstwertdiskrepanzen. Es wird angenommen, dass die Höhe der impliziten Selbstwertschätzung in der frühen Kindheit festgelegt wird (Wilson et al., 2000). Hinweise darauf finden sich im Zusammenhang zwischen impliziter Selbstwertschätzung und dem erinnerten Interaktionstil der Eltern (DeHart, Pelham & Tennen, 2006). Dieser Auffassung widersprechen zwei aktuelle Befunde zur Stabilität indirekter Verfahren. Diese betreffen die geringen Retest-Reliabilitäten (Bosson et al., 2000; Rudolph et al., 2008) und die Möglichkeit der Veränderung impliziter Selbstwertschätzung durch evaluative Konditionierung (Baccus, Baldwin & Packer, 2004; Dijksterhuis, 2004). Möglicherweise unterliegt implizite Selbstwertschätzung ähnlichen, oder sogar stärkeren, Schwankungen

wie explizite Selbstwertschätzung. Betrachtet man die Instabilität von explizitem Selbstwert, so ist diese verbunden mit Sensibilität für evaluatives Feedback, Ärger, Feindseligkeit und depressiven Symptomen (Kernis & Goldman, 2003). Ob Schwankungen impliziter Selbstwertschätzung tatsächlich eine Ursache für die Entstehung von Selbstwertdiskrepanzen darstellen, ist eine Frage für zukünftige Forschung.

1.6 Fazit

Die vorliegende Arbeit zeigt, dass implizite Selbstwertschätzung mit indirekten Verfahren reliabel und valide erfasst werden kann. Außerdem sagt impliziter Selbstwert Verhalten vorher, das expliziter Selbstwert nicht prädiziert. Schließlich werden Zusammenhänge von Selbstwertdiskrepanzen und defensivem Verhalten demonstriert. Impliziter Selbstwertschätzung wird somit eine große Bedeutung über explizite Selbstwertschätzung hinaus bescheinigt und sollte daher auch in zukünftigen Untersuchungen berücksichtigt werden. Wie allerdings im Ausblick der Arbeit deutlich wird, bleiben mindestens ebenso viele Fragen unbeantwortet wie mit dieser Arbeit geklärt werden können. Diese Fragestellungen gilt es, als nächstes anzugehen.

1.7 Literatur

(die mit * gekennzeichneten Artikel sind die in diese Arbeit eingebrachten Beiträge)

- Asendorpf, J. B., Banse, R. & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, 83, 380-393.
- Baccus, J. R., Baldwin, M. W. & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, 15, 498-502.
- Back, M. D., Schmukle, S. C. & Egloff, B. (2005). Measuring task-switching ability in the Implicit Association Test. *Experimental Psychology*, 52, 167.
- Back, M. D., Schmukle, S. C. & Egloff, B. (in press). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology*.
- Baine, B. & Crocker, J. (1993). Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard*. New York: Plenum Press.
- 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, efficiency, intention, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1-40). Hillsdale, NJ: Erlbaum.
- Baumeister, R. F. (1998). The Self. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The Handbook of social psychology* (Vol. 1). New York: McGraw-Hill.
- Baumeister, R. F., Campbell, J. D., Krueger, J. I. & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4, 1-44.
- Bosson, J. K., Brown, R. P., Zeigler-Hill, V. & Swann, W. B., Jr. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity*, 2, 169-187.

- 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.
- 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.
- Conner, T. & Barrett, L. F. (2005). Implicit self-attitudes predict spontaneous affect in daily life. *Emotion*, *5*, 476-488.
- Conroy, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K. & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The Quad Model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469-487.
- Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco: Freeman.
- Dasgupta, N. & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800-814.
- De Houwer, J. (2003). The Extrinsic Affective Simon Task. *Experimental Psychology*, *50*, 77-85.
- De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 11-28). Thousand Oaks, CA: Sage Publishers.
- De Houwer, J. (2008). Comparing measures of attitudes at the procedural and functional level. In R. Petty, R. H. Fazio & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 361-390). Mahwah, NJ: Lawrence Erlbaum.
- De Houwer, J. & De Bruycker, E. (2007). The identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 133-143.
- DeHart, T., Pelham, B. W. & Tennen, H. (2006). What lies beneath: Parenting style and implicit self-esteem. *Journal of Experimental Social Psychology*, *42*, 1-17.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*, 542-575.
- Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology*, *86*, 345-355.

- 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.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, 49, 709-724.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE Model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75-109). Orlando, FL: Academic Press.
- 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.
- Fiedler, K. & Blümke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, 27, 307-316.
- Franck, E., De Raedt, R. & De Houwer, J. (2007). Implicit but not explicit self-esteem predicts future depressive symptomatology. *Behaviour Research and Therapy*, 45, 2448-2455.
- Freud, S. (1933). *New introductory lectures on psychoanalysis*. New York: Norton.
- Gawronski, B. (2006). Die Technik des Impliziten Assoziationstests als Grundlage für Objektive Persönlichkeitstests. In T. M. Ortner, R. T. Proyer & K. Kubinger (Eds.), *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 53-69). Bern: Hans Huber.
- Gawronski, B. & Conrey, F. R. (2004). Der Implizite Assoziationstest als Maß automatisch aktivierter Assoziationen: Reichweite und Grenzen. *Psychologische Rundschau*, 55, 118-126.
- Gebauer, J. E., Riketta, M., Broemer, P. & Maio, G. R. (2008). "How much do you like your name?" An implicit measure of global self-esteem. *Journal of Experimental Social Psychology*, 44, 1346-1354.
- Green, D. M. & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
- 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. & 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.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369-1385.
- Jones, J. T., Pelham, B. W., Mirenberg, M. C. & Hetts, J. J. (2002). Name letter preferences are not merely mere exposure: Implicit egotism as self-regulation. *Journal of Experimental Social Psychology*, 38, 170-177.
- Jordan, C. H., Spencer, S. J. & Zanna, M. P. (2005). Types of high self-esteem and prejudice: How implicit self-esteem relates to ethnic discrimination among high explicit self-esteem individuals. *Personality and Social Psychology Bulletin*, 31, 693-702.
- Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E. & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology*, 85, 969-978.
- Judge, T. A., Erez, A., Thoresen, C. J. & Bono, J. E. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, 83, 693-710.
- 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.
- Karpinski, A. & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, 91, 16-32.

- Kernis, M. H. & Goldman, B. M. (2003). Stability and variability in self-concept and self-esteem. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 106-127). Guilford Press: New York.
- Kitayama, S. & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin*, 23, 736-742.
- Koole, S. L., Dijksterhuis, A. & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, 80, 669-685.
- Koole, S. L. & Pelham, B. W. (2003). On the nature of implicit self-esteem: The case of the name letter effect. In S. J. Spencer & S. Fein (Eds.), *Motivated social perception: The Ontario symposium* (pp. 93-116). Hillsdale, NJ: Lawrence Erlbaum.
- Leary, M. R. & Tangney, J. P. (2003). The self as an organizing construct in the behavioral and social science. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 1-14). Guilford Press: New York.
- LeBel, E. P. & Gawronski, B. (2009). How to find what's in a name: Scrutinizing the optimality of five scoring algorithms for the Name-Letter Task. *European Journal of Personality*, 23, 85-106.
- McFarland, S. G. & Crouch, Z. E. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, 20, 483-510.
- McGregor, I. & Jordan, C. H. (2007). The mask of zeal: Low implicit self-esteem, threat, and defensive extremism. *Self and Identity*, 6, 223-237.
- McGregor, I., Nail, P. R., Marigold, D. C. & Kang, S.-J. (2005). Defensive pride and consensus: strength in imaginary numbers. *Journal of Personality and Social Psychology*, 89, 978-996.
- Mierke, J. & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for the 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.
- Nelson, L. D. & Simmons, J. P. (2008). Moniker Maladies: When names sabotage success. *Psychological Science*, 18, 1106-1112.
- Newbegin, I. & Owens, A. Self-esteem and anxiety in secondary school achievement. *Journal of Social Behavior and Personality*, 11, 521-530.

- Nosek, B. A. & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, 19, 625-666.
- Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, 15, 353-361.
- Paulhus, D. L. (1984). Two component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Payne, B. K., Cheng, S. M., Govorun, O. & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89, 277-293.
- Pelham, B. W. & Hetts, J. J. (1999). Implicit and explicit personal and social identity: Toward a more complete understanding of the social self. In T. R. Tyler, R. M. Kramer & O. P. John (Eds.), *The psychology of the social self* (pp. 115-143). New York: Lawrence Erlbaum.
- Pelham, B. W., Mirenberg, M. C. & Jones, J. T. (2002). Why Susie Sells Seashells by the Seashore: Implicit Egotism and Major Life Decisions. *Journal of Personality and Social Psychology*, 82, 469-487.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, 44, 29-45.
- Pinter, B. & Greenwald, A. G. (2005). Clarifying the Role of the 'Other' Category in the Self-Esteem IAT. *Experimental Psychology*, 52, 74-79.
- Riketta, M. & Gebauer, J. (subm.). Loving oneself deep inside: Implicit self-esteem is higher than explicit self-esteem. Manuscript submitted for publication.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rothermund, K. & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift für Experimentelle Psychologie*, 48, 94-106
- *Rudolph, A., Schröder-Abé, M., Riketta, M. & Schütz, A. (2009). Easier when done than said. Implicit self-esteem predicts observed or spontaneous behavior, but not self-reported or controlled behavior. Manuscript submitted for publication.
- *Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. P. & Sedikides, C. (2008). Through a glass, less darkly? Reassessing convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment*, 24, 273-281.

- Schlenker, B. R. (1980). *Impression management: The self-concept, social identity and interpersonal relations*. Monterey, CA: Brooks/Cole.
- Schnabel, K., Banse, R. & Asendorpf, J. (2006). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: The Implicit Association Procedure (IAP). *Experimental Psychology*, 53, 69-76.
- Schröder-Abé, M., Rudolph, A. & Schütz, A. (2007). High implicit self-esteem is not necessarily advantageous: Discrepancies between explicit and implicit self-esteem and their relationship with anger expression and psychological health. *European Journal of Personality*, 21, 319-339.
- *Schröder-Abé, M., Rudolph, A., Wiesner, A. & Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *International Journal of Psychology*, 42, 174-183.
- Schütz, A. (2003). *Selbstwertgefühl. Zwischen Selbstakzeptanz und Arroganz*. Stuttgart: Kohlhammer.
- Schütz, A. (2005). *Je selbstsicherer desto besser. Licht und Schatten positiver Selbstbewertung*. Weinheim: Beltz.
- Schütz, A., Schröder-Abé, M. & Rudolph, A. (2005). *Auf die 'andere' Kategorie kommt es an! Vergleich zweier Versionen des IAT*. Posterpräsentation auf der 8. Arbeitstagung der Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, Marburg.
- Schütz, A. & Sellin, I. (2006). *Die Multidimensionale Selbstwertkala (MSWS)*. Göttingen: Hogrefe.
- Spalding, L. R. & Hardin, C. D. (1999). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science*, 10, 535-539.
- Sriram, N. & Greenwald, A. G. (in press). The Brief Implicit Association Test. *Experimental Psychology*.
- Steffens, M. C. (2004). Is the Implicit Association Test Immune to Faking? *Experimental Psychology*, 51, 165-179.
- Steffens, M. C. & Schulze König, S. (2006). Predicting spontaneous Big Five behavior with Implicit Association Tests. *European Journal of Psychological Assessment*, 22, 13-20.

- Steinberg, J. A., Karpinski, A. & Alloy, L. B. (2007). The exploration of implicit aspects of self-esteem in vulnerability - stress models of depression. *Self and Identity*, 6, 101-107.
- Strack, F. & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220-247.
- Swann, W. B., Jr, Stein-Seroussi, A. & Giesler, R. B. (1992). Why people self-verify. *Journal of Personality and Social Psychology*, 62, 392-401.
- Taylor, S. & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210.
- Teige-Mocigemba, S., Klauer, K. C. & Rothermund, K. (2008). Minimizing method-specific variance in the IAT: A Single Block IAT. *European Journal of Psychological Assessment*, 24, 237-245.
- Tice, D. M. & Baumeister, R. F. (1990). Self-esteem, self-handicapping, and self-presentation: The strategy of inadequate practice. *Journal of Personality*, 58, 449-464.
- von Collani, G. & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 3-7.
- Wilson, T. D., Lindsey, S. & Schooler, T. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.

2 Zeitschriftenbeiträge

- Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. & Sedikides, C. (2008). Through a glass, less darkly: Reassessing the convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment, 24*, 273-281. 36
- Rudolph, A., Schröder-Abé, M., Riketta, M. & Schütz, A. (2009). Easier when done than said! Implicit self-esteem predicts observed or spontaneous behaviour, but not self-reported or controlled behaviour. Manuscript submitted for publication. 56
- Schröder-Abé, M., Rudolph, A., Wiesner, A. & Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *International Journal of Psychology, 42*, 174-183. 78

**Through a Glass, Less Darkly?
Reassessing Convergent and Discriminant Validity in Measures of Implicit Self-
Esteem**

Almut Rudolph¹

Michela Schröder-Abé¹

Astrid Schütz¹

Aiden P. Gregg²

Constantine Sedikides²

¹Chemnitz University of Technology

²University of Southampton

Abstract

Self-esteem has been traditionally assessed via self-report (explicit self-esteem: ESE). However, the limitations of self-report have prompted efforts to assess self-esteem indirectly (implicit self-esteem: ISE). It has been theorized that ISE and ESE reflect the operation of largely distinct mental systems. However, although low correlations between measures of ISE and ESE empirically support their discriminant validity, similarly low correlations between different measures of ISE do not support their convergent validity. We explored whether such patterns would re-emerge if more newly developed, specific, and reliable ISE measures were used. They did, although some convergent validity among ISE measures emerged once confounds due to conceptual mismatch, individual differences, and random variability were minimized. Nonetheless, low correlations among ISE measures are not primarily due to the usual psychometric suspects, and may be the result of other factors including subtle differences between structural features of such measures.

Key words: implicit self-esteem, explicit self-esteem, Implicit Association Test, Extrinsic Affective Simon Task, Single Category Implicit Association Test, Go/No-Go Association Task

To investigate people's attitude towards themselves—their *self-esteem*—psychologists have traditionally relied on self-report (*explicit self-esteem* or ESE; Rosenberg, 1965). Fortunately, when reporting their self-esteem, people are reasonably knowledgeable about themselves, honest with themselves, and honest with others. Nonetheless, people sometimes lack self-insight (“How do I feel about myself *really*?”; Wilson, 2002), deceive others (“I really think I’m useless, but I better pretend to be great!”; Schlenker & Leary, 1982), or even deceive themselves (“I’m great—even if everyone hates me!”; Paulhus, Fridhandler, & Hayes, 1997). Hence, self-reports of self-esteem, though tolerably valid, still contain some systematic error.

One possible way to curtail such error is to employ *indirect measures* of self-esteem (Greenwald & Farnham, 2000). Such measures are designed to reveal people's attitude towards themselves from their reactions to self-related stimuli (e.g., first and last names, first-personal pronouns), typically under conditions where people are either unaware of, or lack control over, the measurement process (e.g., Rudolph, Schröder, & Schütz, 2006). Consider unawareness: the *initials preference task* (IPT; Koole, Dijksterhuis, & van Knippenberg, 2001) requires respondents to rate all letters of the alphabet for likeability, whereupon people typically exhibit an unknowing preference for their initials.¹ Or consider uncontrollability: the *Implicit Association Test* (IAT; Greenwald, McGhee, & Schwartz, 1998) requires respondents to co-classify self-related and self-unrelated stimuli alongside positive and negative stimuli. Respondents are also required to go as quickly as they can without making errors. However, they typically find responding more difficult—and hence go more slowly—when the four target categories are configured one way (e.g., *Self* with *Bad*, *Non-Self* with *Good*) rather than another (e.g., *Self* with *Good*, *Non-Self* with *Bad*). What indirect measures assess is often termed *implicit self-esteem* (ISE). The properties of ISE have been assumed to reflect properties of the indirect measures used to assess it (e.g., ISE is unconscious and automatic; Greenwald & Farnham, 2000). However, it remains controversial whether and to what extent indirect measures operate via wholly implicit processes (for a discussion, see De Houwer & Moors, 2007).

Standard dual-process models (e.g., Strack & Deutsch, 2004) suggest that ISE and ESE, being subserved by modular cognitive systems, should yield measures that are largely independent. In addition, different measures of each construct, in virtue of tapping into the same modular system, should exhibit reasonable intercorrelations (though see Marsh & Craven, 2006; Sakellaropoulou & Baldwin, 2007). Given such expectations, prior empirical research has yielded two patterns, one reassuring, the other troubling. The first suggests

discriminant validity. In particular, measures of ISE and ESE typically show weak correlations (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005), although these can be augmented under theoretically specified conditions (e.g., Jordan, Whitfield, & Zeigler-Hill, 2007; Koole et al., 2001; Olson, Fazio, & Hermann, 2007). However, the second pattern suggests a *lack of convergent validity*: different measures of ISE typically fail to exhibit the predicted intercorrelations (Bosson, Swann, & Pennebaker, 2000). If valid, this pattern implies one of two things: either (a) ISE exists and is heterogeneous; or (b) ISE does not even exist. The matter remains unresolved. Nonetheless, most measures of ISE *do* converge insofar as they register a pronounced average self-positivity bias (e.g., Greenwald & Farnham, 2000; Gregg & Sedikides, 2008). In addition, such measures exhibit meaningful antecedents and consequences (see Koole & DeHart, 2007, for an overview). Hence, there are some reasons to believe (a) over (b).

However, an even gloomier possibility exists: *both* patterns could be artifacts of measurement error. Measures of ISE have a reputation for unreliability (Bosson et al., 2000). Such unreliability could obscure latent correlations, and falsely suggest that ESE and ISE diverge when they do not, or that different indices of ISE fail to converge when they do.

In this article, we reconsider the convergent and discriminant validity of ISE and ESE. In particular, we investigate whether and to what extent two factors—(a) the reliability and sensitivity to self-positivity bias of different measures of ISE, and (b) the conceptual correspondence between what different measures of ISE assess—moderate the relation between measures of ISE and ESE, and between different measures of ISE. We then attempt to formulate concrete and constructive recommendations for future research, and make some empirically informed theoretical interpretations.

Some years ago, a study concluded that the IAT and the IPT were the most reliable and valid measures of ISE available (Bosson et al., 2000). Since then, however, indirect measures have proliferated. In particular, three new measures have emerged that—unlike the IAT—permit associations toward an object to be assessed in isolation: the *Single-Category IAT (SC-IAT)* (Karpinski & Steinman, 2006), the *Extrinsic Affective Simon Task (EAST)* (De Houwer, 2003), and the *Go/No-Go Association Task (GNAT)* (Nosek & Banaji, 2001). These new measures are of interest, because they assess self-related evaluations independently of other-related evaluations—an advantage, given that variations in the theoretically irrelevant non-self category confound performance on the self-esteem IAT (Karpinski, 2004). In addition, due to acknowledged problems concerning the effect size and reliability of the EAST, an improved variant of the EAST, namely the *Identification EAST (ID-EAST)*, has

been devised (De Houwer & De Bruycker, 2007). Furthermore, the present authors have devised a potentially more reliable version of the IPT, namely the *Duplicate IPT (D-IPT)*. To update the literature, we conducted three studies to compare and contrast the older IAT and IPT with the newer SC-IAT, EAST, ID-EAST, GNAT, and D-IPT as putative indices of ISE. In addition to using new measures of ISE, we also applied more recently developed algorithms (e.g., *D-index*; Greenwald, Nosek, & Banaji, 2003) to maximize validity, and employed standard indices of internal consistency (based on equivalent split-halves, and incorporating warranted Spearman-Brown adjustments).

Study 1

In our first study, we evaluated three different measures of ISE: an IPT, an EAST, and an IAT. We quantified their internal consistency, their test-retest stability one-week apart, their intercorrelations, and their correlations with ESE.

Method

Participants and Procedure

A total of 102 students (80 female; $M_{AGE} = 22.7$) participated.² They began by providing basic demographic data and by generating an ID code that contained their first and last initials. Next, they completed three measures of ISE in a fixed order: an IPT, an IAT, and an EAST. Finally, participants completed a measure of ESE. They were then dismissed, but returned exactly one week later to redo the three measures of ISE.

Measures of ISE and ESE

IPT. As in Bosson et al. (2000), participants rated each letter of the alphabet on a scale from 1 (*I dislike this letter very much*) to 7 (*I like this letter very much*). To derive an initials preference index that controlled for general letter popularity and personal rating tendencies, we followed the guidelines provided by Koole et al. (2001). We derived Spearman-Brown corrected split-half estimates of internal consistency from correlating ratings of first and last initials.

IAT. The IAT conformed to the canonical five-block structure and procedure (see Appendix; Greenwald & Farnham, 2000). We kept critical block order constant to reduce method variance. We computed the IAT index using the scoring algorithm (the *D-index*) recommended by Greenwald et al. (2003). Higher scores reflect an automatic preference for *Self* over *NonSelf*. The IAT's internal consistency was based on a Spearman-Brown corrected split-half correlation, the split-halves being derived from alternating pairs of trials in both critical blocks. This served to ensure (a) that both halves were maximally comparable and (b) that attribute and target trials were equivalently represented in each half.³

EAST. The EAST featured the same general structure and response options as described by De Houwer (2003). As for the IAT, we computed an EAST index using the *D*-algorithm. Higher scores reflect stronger automatic liking for *Self*. We computed internal consistency on the basis of a split-half, as above.

ESE. The total score from the 32-item Multidimensional Self-Esteem Scale (*MSES*; Schütz & Sellin, 2006) served as index of ESE. Each item featured a seven-point scale with one of two types of endpoints (1 = *Not at All* to 7 = *Very Much*; 1 = *Never* to 7 = *Always*.)

Results and Discussion

Positivity Bias. All indices of ISE, like the index of ESE, yielded significant effects that were both positive in sign and large in magnitude, although the EAST index lagged behind the others (Table 1). Thus, all indices of self-esteem, implicit and explicit, converged at a directional level, revealing a general bias towards positive self-evaluation.

Internal Consistency and Temporal Stability. As Table 1 shows, the IAT displayed satisfactory levels of internal consistency. However, levels were less satisfactory for the IPT, and unsatisfactory for the EAST. In addition, only the IAT ($r_{tt} = .54$) and the IPT ($r_{tt} = .56$) but not the EAST ($r_{tt} = .18$) showed satisfactory levels of temporal stability over one week.

Intercorrelations. Correlations between the three indices of ISE hovered around zero on both measurement occasions (Table 2). In addition, neither the IPT nor the IAT index correlated significantly with the ESE index on either occasion; and although the EAST index did on one occasion, its failure to do so on another, combined with its trifling internal consistency, suggest sampling error as the most likely explanation.

The overall pattern implies that, although indices of ISE and ESE show directional convergence, they do not show convergence at the level of individual scores. It seems that either the underlying “elephant” of ISE (Bosson et al., 2000) is an illusory beast, or that different indices of ISE map on to very different parts of that underlying “elephant.” However, given that the IAT and IPT at least showed a degree of internal consistency, the lack of emergent relations does not appear to have been solely an artifact of measurement unreliability.

Study 2

We proceeded to test a further indirect measure of ISE designed to assess automatic attitudes specifically toward the *Self*—the SC-IAT—in conjunction with structurally improved versions of previous measures (i.e., the ID-EAST and D-IPT), plus an IAT. We examined their relative psychometric properties, intercorrelations, and correlations with ESE.

Method

Participants and Procedure

A total of 60 students (11 female; $M_{AGE} = 22.4$) participated. In fixed order, participants completed four measures of ISE (an IAT, a D-IPT, an ID-EAST, and a SC-IAT) followed by one measure of ESE. Due to computer problems, data from eleven participants on the ID-EAST were lost. One participant was excluded due to extreme scores.

Measures of ISE and ESE

D-IPT. We administered the IPT as in Study 1, except that all letters were now presented, not just once, but *twice*, in the same fixed random order. The repetition was designed to increase its reliability. We derived an overall initial preference index by averaging the two initial preference scores (calculated as before) computed separately from each of the rated alphabets. Internal consistency was based on a Spearman-Brown corrected split-half correlation using the (now averaged) preference scores for first and last initials.

IAT. The layout of the IAT, and the computation of its results, was the same as in Study 1. The only difference was the addition of a few extra target stimuli (e.g., *I, mine, their, them*) to vary the representation of self (also added to the SC-IAT and ID-EAST).

SC-IAT. Unlike the IAT, the two critical blocks of the SC-IAT required participants to classify stimuli into one of three categories using two keys. We indexed automatic liking for self by an analogue of the *D*-index. Internal consistency was estimated as for the IAT.

ID-EAST. The ID-EAST contains a structural modification designed to ensure that the stimuli it contains are processed based on semantics rather than on features.⁴ We adopted a single category ID-EAST, using only the *Me* target category (Appendix).⁵ Trial data were aggregated, and internal consistency was estimated, just as in the original EAST.

MSES. As in Study 1, ESE was measured with the MSES (Schütz & Sellin, 2006).

Results and Discussion

Positivity Bias. Again, all indices yielded significant effects that were both positive in sign and large in magnitude, with the ID-EAST index lagging behind the others (Table 1). Thus, a convergent directional bias towards positive self-evaluation re-emerged.

*Internal Consistency.*⁶ Both the IAT and SC-IAT displayed high levels of internal consistency, with the value for the latter numerically exceeding that for the former (Table 1). In addition, both the D-IPT and the ID-EAST showed reasonable levels of internal consistency, higher than with the IPT and original EAST in Study 1.

Intercorrelations. Overall, the pattern replicated Study 1 (Table 2). First, no correlation between any index of ISE and the ESE index approached significance. Second, none of the intercorrelations between indices of ISE attained significance, although one

approached it. However, given that, thanks to methodological innovations to the EAST and IPT, the internal consistency was even higher than in Study 1, measurement unreliability is unlikely to be the whole explanation for the absence of significant correlations.

Nonetheless, from an exploratory perspective, one might ask why the sole marginal correlation found was between the ID-EAST and the SC-IAT. We believe the answer is that, despite some structural differences, both these indirect measures *corresponded conceptually*, in that they reflected specific self-evaluations uncomplicated by any salient other-comparisons. By reflecting the same target, they achieved greater convergent validity.

Study 3

Our investigations of measures of ISE had yet to include a promising methodology: the GNAT. This measure was originally designed to provide an alternative to the IAT assessing automatic associations towards individual objects. In Study 3, therefore, we duly examined the reliability and sensitivity to self-positivity bias of the GNAT as a potential measure of ISE. We additionally examined the relation of the GNAT to another indirect measure of self-esteem, the IAT, as well as to a traditional direct measure (Rosenberg, 1965).

One notable feature of the GNAT is that, when used to assess automatic evaluations towards X and Y individually, its results can be combined to create a relative index that conceptually corresponds to the standard IAT index. In particular, if two GNAT blocks respectively assess positive and negative evaluations of *self*, and two more GNAT blocks respectively assess positive and negative evaluations of *non-self*, then all four blocks can assess positive and negative evaluations of self *relative* to non-self—precisely what the IAT assesses. Building on the suggestive results of Study 2 with regard to the lone marginal correlational observed, we tested in Study 3 whether the IAT index would correlate better with a relative GNAT index than with individual GNAT indices, given that the conceptual correspondence would be exact in the former case, but inexact in the latter cases.

We additionally explored the impact on levels of convergent validity of attempting to reduce, first, systematic error (i.e., variance due to individual differences in classification ability), and second, random error (i.e., variance due to measurement unreliability). Specifically, we (a) compared correlations obtained using the original IAT index (Greenwald & Farnham, 2000) to those based on the newer algorithm (Greenwald et al., 2003), thereby controlling for some systematic error, and (b), compared the second set of correlations to corresponding coefficients estimated in a structural model, thereby controlling for random error. We predicted that the combined use of the improved algorithm and structural modeling would increase convergent validity.

Method

Participants and Procedure

The sample comprised 195 students⁷, predominantly young ($M_{AGE} = 20.5$) and female (85%). Due to participant dropout, technical failures, task non-compliance, or extreme scores, listwise N s ranged from 182 to 195 across various analyses. The dataset comprised scores from two self-esteem IATs and GNATs, both run twice with a time interval of one week, along with measures of ESE. The IAT and GNAT featured identical categories and stimuli.

Measures of ISE and ESE

IAT. The IAT in Study 3 resembled those in preceding studies except that it comprised only two critical blocks presented in random order.

GNAT. The GNAT comprised four blocks (*Self-Positive*, *Self-Negative*, *NonSelf-Positive*, and *NonSelf-Negative*) presented in random order, each of which featured two target categories out of a possible four (i.e., two of *Self*, *NonSelf*, *Positive*, *Negative*). Participants attempted to press a key within 750 ms when a word presented matched those categories, and not to press it when a word did not (see Appendix). Accordingly, a response on each trial could be classed as *hit*, *false alarm*, *correct rejection*, or *miss*. Overall accuracy within each block at distinguishing target from non-target items was duly quantified by d' , the normalized hit rate minus the normalized false alarm rate (Nosek & Banaji, 2001).

We derived *five* GNAT indices. First, we computed all four *individual* d' indices: *Self P>N* (d' in the *Self-Positive* block minus d' in the *Self-Negative* block); *NonSelf N>P* (d' in the *NonSelf-Negative* block minus d' in the *NonSelf-Positive* block); *Positive S>NS* (d' in the *Self-Positive* block minus d' in the *NonSelf-Positive* block); and *Negative NS>S* (d' in the *NonSelf-Negative* block minus d' in the *Self-Negative* block). Second, we derived a *relative* index (*Overall*), by computing the average of *Self P>N* and *NonSelf N>P*.

ESE. We used a 10-item questionnaire (*RSES*; Rosenberg, 1965) to assess participants' overall liking for themselves (1 = *strongly agree*; 4 = *strongly disagree*).

Results and Discussion

Positivity Bias and Reliability. Averaging across sessions, the relative GNAT index showed high sensitivity to self-positivity bias ($d = 1.56$), as well as reasonable internal consistency ($r = .65$) and modest test-retest reliability ($r = .51$). Individual GNAT indices also showed high sensitivity to self-positivity bias (d s = .80 to 1.60), but more modest internal consistency (r s = .52 to .59) and low test-retest reliability (r s = .23 to .38). Corresponding IAT indices were comparable (original index $d = 1.66$, $r_{ic} = .85$, $r_{tt} = .39$; new index $d = 1.54$, $r_{ic} = .67$, $r_{tt} = .29$).⁸

Convergent Validity. Table 3 displays the correlation between each of the five GNAT indices and (a) the original IAT index, (b) the new IAT index, and (c) the new IAT index, estimated as part of a different structural model. In each model, we estimated (assuming unequal loadings, unequal error variances, and uncorrelated errors) correlations among three common factors: each GNAT index, the IAT, and the RSES. To derive manifest variables on the basis of which to estimate latent correlations, we created four parcels for each measure, consisting of equivalent split-halves derived from each of the two measurement occasions.⁹

As predicted, the correlations involving GNAT indices were always numerically (albeit not significantly) larger when the new IAT index was used as opposed to the original IAT index. Moreover, these correlations were numerically larger again when estimated from a structural model. Finally, the numerically highest correlation was always obtained between the IAT index and the *Overall* GNAT index. Although the increment in the magnitude and reliability of correlations at each step was admittedly small and nonsignificant, the *combined* increment after taking all three steps was nonetheless theoretically meaningful. For example, if a researcher using our dataset had attempted *neither* to minimize systematic and random error *nor* to examine a pair of conceptually convergent indices, then he or she might well have falsely concluded that the IAT and GNAT (e.g., using the *Positive S>NS* index) did *not* converge ($r = .11, p = .14$), whereas had they taken all these steps, they would have correctly concluded that the IAT and GNAT (using the *Overall* index) *did* converge ($r = .27, p < .05$). Indeed, when underlying relations are weak, it is critical to maximize all available statistical power and conceptual correspondence.

Discriminant Validity. The RSES failed to covary even marginally with either index of ISE, both at the level of raw correlations (GNAT: $-.07 < r < -.01$; IAT: $.01 < r < .10$) and estimated structural coefficients (GNAT: $-.11 < r < -.01$; IAT: $r \approx .01$). Thus, both the GNAT and IAT indices were still independent of the RSES index.

General Discussion

We close by making some practical recommendations for researchers wishing to explore ISE empirically and drawing some theoretical conclusions from our findings.

Practical Recommendations

Across three studies, the IAT—and two of its methodological offshoots designed to capture ISE more specifically, namely the SC-IAT and the relative GNAT index—exhibited satisfactory to good levels of reliability. The IPT and EAST exhibited comparatively lower levels, although their improved methodological variants, the D-IPT and ID-EAST, fared somewhat better. Hence, IAT, SC-IAT and GNAT are, for psychometric reasons, to be

recommended over IPT and EAST (at least in their original form) for use in research on individual differences in ISE. In addition, the smaller aggregate self-positivity biases obtained for the EAST and ID-EAST (although not for the IPT and D-IPT) suggest less sensitivity to self-positivity bias, and may counterindicate their use.

Theoretical Conclusions

If ISE is a single construct distinct from ESE, then one would expect, all else equal, different indices of ISE to correspond more strongly with one another than with an index of ESE. Having employed several newly developed, reliable, and specific measures of ISE (i.e., the SC-IAT, EAST, and GNAT), what did we find?

First, we found that, despite a pronounced positivity bias for ESE and ISE indices at a directional level, individual ISE scores remained independent of individual ESE scores. Moreover, this independence could not be attributed to measurement unreliability: most indices exhibited satisfactory internal consistency, and estimated structural coefficients were little higher than observed raw correlations. Nor could this independence be attributed to a lack of correspondence between direct and indirect measures of self-esteem: the more specific ISE indices did not correlate any better with ESE than the relative IAT indices. Thus, in the absence of other explanations, our results are in keeping with dual-process models of cognition applied to self-esteem. Indeed, our results are *starkly* in keeping with such models: even though a moderate degree of explicit-implicit correlation is typically observed across a range of topics (Nosek & Smyth, 2007), we found almost *none*. Perhaps self-reported self-esteem is particularly prone to reflect the impact of carefully pondered propositions and of self-presentational concerns (Upshaw & Yates, 1968), and is only more rarely a reflection of introspective insights into spontaneous self-feelings (Jordan et al., 2007; Koole et al., 2001; Olson et al., 2007), given that (a) the self is habitually the focus of so much cognition (e.g., Rogers, Kuiper, & Kirker, 1977), and (b) maintaining a positive self-view is such an urgent motivational priority (Sedikides & Gregg, 2003, 2008; Sedikides & Strube, 1997).

Second, we found, in Studies 1 and 2, that nearly all correlations between indices of ISE—even involving several newer, more specific, and more reliable measures—fell well short of significance. These results echoed previous research casting doubt on the convergent validity of measures of ISE (Bosson et al., 2000). The sole exception here was the marginal correlation obtained between two measures of ISE that were conceptually correspondent (SC-IAT and ID-EAST). Taking our cue from this suggestive result, we took pains in Study 3 *simultaneously* to minimize confounding sources of variance (i.e., conceptual mismatch between indices, individual differences in reaction time, random error of measurement).

When we did so, some evidence of convergent validity finally *did* emerge. Yet the level of convergence remained curiously low. The question is why.

We suspect these answer may lie, not merely in the heterogeneity of ISE itself, but also in the contrasting *structural features* that characterize even similar-seeming indirect measures (De Houwer & Moors, 2007; Deutsch & Gawronski, 2008; Payne, Burkley, & Stokes, 2008). For example, due to the “bipolar” layout of the IAT (i.e., A&X *versus* B&Y), asymmetry in the salience of its categories alone is sufficient to engender effects (Rothermund & Wentura, 2004). However, this dynamic is less likely to confound its cousin, the more “unipolar” GNAT, to a similar degree. Moreover, whereas in the IAT accuracy is held constant and reaction time let vary, in the GNAT reaction time is held constant and accuracy let vary. Such disparities in structure may cause corresponding disparities in performance, not only because they recruit different types of classification skill, but also because they elicit different types of classification strategy. We suspect that understanding the structural features of ISE measures will shed light on the reasons for their low convergent validity, and that the devil may be in the procedural details not the underlying construct.

References

- Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. (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.
- Byrne, B. M. (2001). *Structural Equation Modeling with AMOS, Basic Concepts, Applications, and Programming*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- De Houwer, J. (2003). The Extrinsic Affective Simon Task. *Experimental Psychology, 50*, 77-85.
- De Houwer, J., & Moors, A. (2007). How to define and examine the implicitness of implicit measures. In B. Wittenbrink, & N. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford Press.
- De Houwer, J., & De Bruycker, E. (2007). The identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 95-104.
- Deutsch, R., & Gawronski, B. (2008). When the method makes a difference: Antagonistic effects on “automatic evaluations“ as a function of task characteristics of the measure. Manuscript submitted for publication.
- 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. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.
- Gregg, A. P., & Sedikides, C. (2008). *Implicit self-esteem and narcissism: Rethinking the link*. Manuscript submitted for publication.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369-1385.

- Jordan, C. H., Whitfield, M., & Zeigler-Hill, V. (2007). Intuition and the correspondence between implicit and explicit self-esteem. *Journal of Personality and Social Psychology, 93*, 1067-1079.
- 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., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*, 16-32.
- Koole, S. L., & DeHart, T. (2007). Self-affection without self-reflection: Origins, models, and consequences of implicit self-esteem. In C. Sedikides, & S. J. Spencer (Eds.), *The Self*. New York: Psychology Press.
- Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80*, 669-685.
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*, 133-163.
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition, 19*, 625-666.
- Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54*, 14-29.
- Olson, M. A., Fazio, R. H., & Hermann, A. D. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological Science, 18*, 287-291.
- Paulhus, D. L., Fridhandler, B., & Hayes, S. (1997). Psychological defense: Contemporary theory and research. In J. Johnson, R. Hogan, & S. R. Briggs (Eds.), *Handbook of Personality Psychology*. New York: Academic Press.
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*, 16-31.
- Rogers T. B., Kuiper, N. A., & Kirker W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality Social Psychology, 35*, 677-688.

- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*, 139-165.
- Rudolph, A., Schröder, M., & Schütz, A. (2006). Ein Impliziter Assoziationstest zur Erfassung von Selbstwertschätzung [An Implicit Association Test measuring self-esteem]. In T. M. Ortner, R. T. Proyer, & K. D. Kubinger (Hrsg.), *Theorie und Praxis Objektiver Persönlichkeitstests*. Bern: Verlag Hans Huber.
- Sakellaropoulo, M., & Baldwin, M. W. (2007). The hidden sides of self-esteem: Two dimensions of implicit self-esteem and their relation to narcissistic reactions. *Journal of Experimental Social Psychology*, *43*, 995-1001.
- Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization and model. *Psychological Bulletin*, *92*, 641-669.
- Schmukle, S. C., & Egloff, B. (2006). Assessing anxiety with Extrinsic Simon Tasks. *Experimental Psychology*, *53*, 149-160.
- Schütz, A., & Sellin, I. (2006). *Die Multidimensionale Selbstwertkala (MSWS)*. [The multidimensional self-esteem scale]. Göttingen: Hogrefe.
- Sedikides, C., & Gregg, A. P. (2003). Portraits of the self. In M. A. Hogg & J. Cooper (Eds.), *Sage handbook of social psychology* (pp. 110-138). London: Sage Publications.
- Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science*, *3*, 102-116.
- Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*, *29*, 209-269. New York, NY: Academic Press.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychological Review*, *8*, 220-247.
- Upshaw, H. S., & Yates, L. A. (1968). Self-persuasion, social approval, and task success as determinants of self-esteem following impression management. *Journal of Experimental Social Psychology*, *4*, 143-152.
- Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.

Footnotes

¹ Although the bias extends to all letters in one's name, it is most pronounced for one's initials (Koole et al., 2001), a pattern that measures of ISE tend to exploit.

² We reanalyzed a relevant subset of variables of the data used in Schröder-Abé et al. (2007).

³ As Schmukle and Egloff (2006) note, this alternating method of deriving split-halves is liable to estimate internal consistency better than one in which the data is simply split into earlier and later trials, because temporal order effects are avoided. A further advantage is that they can be computed across every index of ISE and ESE, thereby increasing comparability.

⁴ We adopted the convention to classify target stimuli in the ID-EAST on the basis of their being shown in uppercase or lowercase (De Houwer & De Bruycker, 2007).

⁵ In addition, one *Self* stimulus was idiosyncratic: each participant's first name.

⁶ For a subset of the sample ($N = 39$), we also computed the six-month temporal stability of IAT, D-IPT, and SC-IAT scores. The IAT and D-IPT showed satisfactory levels ($r_{stt} = .60$ and $.68$ respectively), the SC-IAT more modest levels ($r_{tt} = .44$).

⁷ We reanalyzed a relevant subset of variables reported by Gregg and Sedikides (2008).

⁸ One oddity that emerged was that the newer IAT index exhibited lower internal consistency than the original IAT index, despite the fact that the former was specifically developed to yield higher internal consistency than the latter (Greenwald et al., 2003). In our case, this could have been a procedural artifact, given that our IAT was somewhat abbreviated.

⁹ Due to generally low intercorrelations, the SEM fit indices were, according to Byrne's (2001) criteria, located at the boundary values of a mediocre fit (e.g., value for the relative GNAT: $RMSEA = .088$, $CFI = .910$, $TLI = .863$).

Table 1

Study 1 and 2: Descriptive Statistics, Internal Consistency, Directional Significance, and Effect Size of Explicit and Implicit Self-Esteem Indices

Self-Esteem Index	Mean (SD)	Split-Half Reliability	One-Sample <i>t</i>	Cohen's <i>d</i>
Study 1				
Explicit				
MSES (T1)	4.69 (.90)	.94	13.38	1.33
Implicit				
IPT ₁	.87 (.66)	.51	13.37	1.33
IPT ₂	.88 (.72)	.50	12.29	1.22
IAT ₁	.62 (.33)	.85	19.18	1.91
IAT ₂	.58 (.32)	.83	18.21	1.81
EAST ₁	.34 (.55)	.16	6.22	.62
EAST ₂	.45 (.61)	.24	7.39	.74
Study 2				
Explicit				
MSES	4.75 (.87)	.93	11.15	1.45
Implicit				
D-IPT	.72 (.64)	.69	8.41	1.09
IAT	.64 (.30)	.80	16.67	2.17
SC-IAT	.46 (.29)	.88	12.42	1.62
ID-EAST	.67 (.76)	.64	6.19	.81

Note. $N_{\text{Study 1}} = 102$, $N_{\text{Study 2}} = 60$. MSES = Multidimensional Self-Esteem Scale; (D-) IPT = (Duplicate) Initials Preference Task; (SC-) IAT = (Single Category) Implicit Association Test; (ID-) EAST = (Identification) Extrinsic Affective Simon Task. SD = Standard deviation; Subscripts (1, 2) indicate measurement occasion (one week apart). One-sample *t*-tests to compare the mean of each index with the theoretical midpoint of its scale. All *t*-values reported are significant at $p < .0001$. Split-Half Reliability based on split-half correlations incorporating Spearman-Brown adjustments. Cohen's *d* refers to the standardized difference between theoretical midpoint of scale and the observed mean for each index.

Table 2

Study 1 and 2: Correlations between Indices of Explicit and Implicit Self-Esteem

		Study 1				
Self-Esteem Index		1	2	3	4	
Explicit						
1	MSES	-	.03	.14	.28**	
Implicit						
2	IPT	-.07	-	-.07	.07	
3	IAT	-.06	.06	-	-.08	
4	EAST	.09	.07	-.09	-	
		Study 2				
Self-Esteem Index		1	2	3	4	5
Explicit						
1	MSES	-				
Implicit						
2	D-IPT	.11	-			
3	IAT	-.04	.07	-		
4	SC-IAT	-.05	-.07	.09	-	
5	ID-EAST	-.04	-.06	-.03	.25 [†]	-

Note. $N_{\text{Study 1}} = 102$, $N_{\text{Study 2}} = 60$. MSES = Multidimensional Self-Esteem Scale. (D-) IPT = (Duplicate) Initials Preference Task; (SC-) IAT = (Single Category) Implicit Association Test; (ID-) EAST = (Identification) Extrinsic Affective Simon Task. In Study 1, Time 1 correlations appear below the diagonal, Time 2 correlations above the diagonal (with the MSES administered only once).

[†] $p < .10$, ** $p < .01$.

Table 3

Study 3: Intercorrelations between Indices of Implicit Self-Esteem (zero-order coefficients and coefficients estimated in a structural model)

GNAT Index	IAT original algorithm	IAT new algorithm	IAT new algorithm (SEM)
<i>Differential Indices</i>			
Self P>N	.14*	.14*	.21 [†]
NonSelf N>P	.11	.15*	.21 [†]
Positive S>NS	.11	.13 [†]	.25*
Negative NS>S	.15*	.17*	.24 [†]
<i>Relative Index</i>			
Overall	.15*	.18*	.27*

Note. $N = 195$. GNAT = Go/No-Go Association Task; IAT = Implicit Association Test; Self P>N = Self-Positive block minus Self-Negative block; NonSelf N>P = NonSelf-Negative block minus NonSelf-Positive block; Positive S>NS = Self-Positive block minus NonSelf-Positive block; Negative NS>S = NonSelf-Negative block minus Self-Negative block.

[†] $p < .10$, * $p < .05$.

Appendix

Studies 1, 2, and 3: Structural and Categorical Features of All Indirect Measures of Self-Esteem

<i>Block</i>	<i>Trial N</i>	<i>Task</i>	<i>Press Left Key</i>	<i>Press Right Key</i>
Extrinsic Affective Simon Task (Study 1)				
1	20	P: Attribute's semantic discrimination	Unpleasant	Pleasant
2	20	P: Target's color discrimination	Green ^a	Blue ^a
3-6	30	T: Combined task	Unpleasant + Green ^a	Pleasant + Blue ^a
Single Target Implicit Association Test (Study 2)				
1	40	P: Attribute discrimination	Pleasant	Unpleasant
2-3	40+80	T: Initial combined task	Pleasant + Me	Unpleasant
4-5	40+80	T: Reversed combined task	Pleasant	Unpleasant + Me
Identification Extrinsic Affective Simon Task (Study 2)				
1	30	P: Attribute discrimination	Unpleasant	Pleasant
2	30	P: Target's letter case discrimination	Lower case ^a	Upper case ^a
3-5	50 each	T: Combined task	Unpleasant + lower case ^a	Pleasant + upper case ^a
Implicit Association Test (Study 1 & 2)				
1	24	P: Attribute discrimination	Pleasant	Unpleasant
2	24	P: Target discrimination	Me	Not-Me
3	96	T: Initial combined task	Pleasant + Me	Unpleasant + Not-Me
4	24	P: Reversed target discrimination	Not me	Me
5	96	T: Reversed combined task	Pleasant + Not-Me	Unpleasant + Me
Implicit Association Test (Study 3)				
1	48	T: Initial combined task	Nice + Me	Nasty + Not-Me
2	48	T: Reversed combined task	Nice + Not-Me	Nasty + Me
Go/No-Go Association Task (Study 3)				
			<i>Press Space Bar</i>	<i>Don't Press Space Bar</i>
1	16+48	P+T: Target / Non-target discrimination	Nice + Me	Nasty + Not-Me
2	16+48	P+T: Target / Non-target discrimination	Nasty + Not-Me	Nice + Me
3	16+48	P+T: Target / Non-target discrimination	Nice + Not-Me	Nasty + Me
4	16+48	P+T: Target / Non-target discrimination	Nasty + Me	Nice + Not-Me
Sample stimuli (Study 1 & 2):				
Pleasant (<i>smile, joy</i>); Unpleasant (<i>pain, war</i>); Me (<i>self, my</i>); Not-Me (<i>other, yours</i>)				
Sample stimuli (Study 3):				
Nice (<i>excellent, love</i>); Nasty (<i>bomb, hatred</i>); Me (<i>myself, my</i>); Not-Me (<i>they, them</i>)				

Note. P = practice blocks; T = test blocks. ^a Stimuli of the target category (*Self* and *Non-Self*) are presented in the defined color or letter case. A complete list of the stimuli can be obtained from the authors.

Easier when Done than Said!
**Implicit Self-Esteem predicts Observed or Spontaneous Behavior, but not Self-
Reported or Controlled Behavior**

Almut Rudolph¹

Michela Schröder-Abé¹

Michael Riketta^{†2}

Astrid Schütz¹

¹Chemnitz University of Technology

²University of Tübingen

[†]Deceased

Abstract

Evidence for the criterion validity of indirect self-esteem measures is still limited, with only some studies finding effects of implicit (ISE) independent of explicit (ESE) self-esteem. This may be due to the fact that studies predicting actual behavior are particularly rare. The present study contributes evidence to the predictive validity of the Implicit Association Test and a recently developed self-judgment task under cognitive load. We used criteria beyond self-report: experimenter-ratings of anxiety, linguistic aspects of anxiety, and spontaneous self-confident behavior. Using paired criteria, we tested a double dissociation of ISE and ESE. Results supported our hypothesis: ISE predicted self-confident behavior or aspects of anxiety that ESE was not able to predict, and vice versa, which demonstrated differential predictive validity of both measures of self-esteem. With behavioral criteria that tapped into impulsive processes, ISE was a better predictor than ESE.

Key words: implicit self-esteem, explicit self-esteem, Implicit Association Test, behavior prediction, LIWC

In the past twelve years of research on indirect measures, only few studies have found evidence for the criterion validity of implicit self-esteem (ISE): ISE rarely showed main effects independent of explicit self-esteem (ESE). How can this be explained, given the fact that evidence for reliability and construct validity has been established (e.g., Bosson, Swann, & Pennebaker, 2000; Rudolph, Schröder-Abé, Schütz, Gregg, & Sedikides, 2008)? In other domains, such as anxiety or shyness, incremental validity of indirect measures over and above direct measures has already been demonstrated (Asendorpf, Banse, & Mücke, 2002; Egloff & Schmukle, 2002). A closer look at those studies reveals that behavioral outcomes have been used as criterion measures. In the present article, we therefore examined the usefulness of indirect measures of self-esteem in predicting behavior. In line with a dual-process approach (Strack & Deutsch, 2004), we argue that it is essential to use non-self-report criteria – and above all spontaneous and observed behaviors – when the goal is to evaluate the criterion validity of measures of ISE.

Most studies on implicit attitudes (for reviews see De Houwer, 2006; Fazio & Olson, 2003) used the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998). However, variants and alternatives have been developed such as the Single-Category IAT (Karpinski & Steinman, 2006) – some of them specifically for the assessment of ISE, such as the Initials Preference Task (Koole, Dijksterhuis, & van Knippenberg, 2001) and a self-judgment task under cognitive load (Riketta & Gebauer, *subm.*). For most of those measures, satisfactory reliability and construct validity have been shown in recent research (Rudolph, et al., 2008).

Criterion validity of ISE measures has, however, rarely been demonstrated, yet. Most of the evidence has been provided in research on self-esteem discrepancies (high ESE/low ISE or low ESE/high ISE). Several studies found the interaction of ESE and ISE to be related to self-enhancement (Bosson, et al., 2000), defensiveness (Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; Schröder-Abé, Rudolph, Wiesner, & Schütz, 2007), and problems with physical and psychological health (Schröder-Abé, Rudolph, & Schütz, 2007). Only very few studies found significant main effects of ISE: ISE predicted negative emotions in everyday life (Conner & Barrett, 2005), self-reported depressive symptoms (Franck, De Raedt, & De Houwer, 2007), and nonverbal indicators of anxiety (Spalding & Hardin, 1999).

The limited evidence for criterion validity of indirect measures of self-esteem may be due to the fact that studies predicting criteria *beyond* self-report (e.g., behavior or observer judgments) are particularly rare. However, self-reported criteria may not be the

via regia in testing the predictive validity of ISE as these criteria may be distorted. Research has shown that people's reports about how they feel might be inaccurate because of self-presentational concerns (Upshaw & Yates, 1968) and limits of introspection (Wilson & Dunn, 2004).

Using observable behavior as criterion has yielded promising results concerning criterion validity and even incremental validity, as has been shown in the domain of implicit self-concept aspects such as anxiety, and neuroticism, extraversion, agreeableness and conscientiousness (Egloff & Schmukle, 2002; Steffens & Schulze König, 2006). One such study has demonstrated a double dissociation for the construct of shyness: Direct measures predicted controlled behavior better than indirect measures whereas indirect measures predicted spontaneous behavior better than direct measures (Asendorpf, et al., 2002).

In the domain of ISE, two of the three studies mentioned earlier that found a main effect of ISE used criteria beyond self-report questionnaires: Conner and Barrett (2005) used an experience sampling procedure of emotional experiences, and Spalding and Hardin (1999) used interviewers' ratings of participants' anxiety. Actual behavior was only used in the latter study. That study may however, not be perfectly conclusive as ISE was measured with a subliminal automatic attitude activation task (e.g., Fazio, Jackson, Dunton, & Williams, 1995). Although priming tasks were applied to the assessment of ISE, their reliability has been shown to be low (Bosson, et al., 2000) and their construct validity has rarely been examined. Also the fact that the finding has not been replicated or extended for a whole decade raises the question whether it can be relied upon.

Contemporary dual-process approaches provide an elaborate framework of how implicit and explicit attitudes influence specific aspects of behavior. These models assume two structurally distinct, albeit interrelated, systems of information processing that activate behavioral schemata, and thus form a common pathway to behavior. Particularly, we build on the Reflective-Impulsive Model (Strack & Deutsch, 2004). In the *reflective system*, behavior is a consequence of consciously accessible information, whereas in the *impulsive system* behavior is elicited based on spreading activation without conscious effort (Strack & Deutsch, 2004). ESE and ISE are assumed to reflect structural components of the two systems. ESE is the "positivity of a person's evaluation of the self" (Baumeister, 1998) which is mentally represented in a propositional format in the reflective system and is therefore assessed with direct measures. ISE is an individual's overlearned, automatic, and not necessarily conscious self-evaluation (Greenwald & Banaji, 1995). It is represented in

the associative store of the impulsive system and can be assessed with the help of indirect measures.

Based on theoretical assumptions and empirical findings, we focused on observable and spontaneous behavioral criteria to test the criterion validity of measures of ISE. As little is known about the behavioral correlates of self-esteem, choosing the criteria to validate ISE measures is not a trifle. Empirically, ESE has been shown to predict behaviors such as assertiveness and confidence: People with high self-esteem are more sociable, outgoing, assertive, confident, and less socially anxious and shy than others (Leary & MacDonald, 2003). In general, the effects of self-esteem are especially pronounced in self-relevant situations. Based on theoretical dual-process models (e.g., Strack & Deutsch, 2004), direct and indirect self-esteem measures should predict different behaviors. Direct measures of self-esteem that tap into propositional representations of the self should therefore be more strongly related to consciously controlled self-confident behaviors, because those behaviors are activated deliberately and need introspection. That is, in a specific situation, a behavioral option is weighted and integrated by the value and the probability of its potential consequences. Thus, in the reflective system, a behavioral schema is activated when a decision has been reached. However, ISE should have a stronger link to automatic and spontaneous self-confident behaviors because indirect measures tap into associatively structured aspects of the self. In the impulsive system, behavior is activated without the need of deliberation and introspection. When a situational cue emerges, spreading activation in the associative store triggers behavioral schemata. In line with the preceding explanation, the lack of correspondence between measures and criteria may have reduced the criterion validity of ISE measures in previous studies.

Based on the considerations elaborated, we designed situations that provoked self-confident behaviors and used behavioral criteria as outcomes. We chose observer-ratings of self-confidence because observers were found to have better information about a person's appearance from the "outside" than does the person him or herself (Vazire & Mehl, 2008). Further, we employed word use as an objective indicator of self-confidence. Word use and directly observed behavior in laboratory and real life setting were found to be strongly related (Fast & Funder, 2008). Finally, we used a fine grained analysis of video-rated material on self-confident nonverbal behavior.

Taken together, given that previous studies were hardly able to establish that measures of ISE predicted relevant criteria (Bosson, et al., 2000), the present research aims at investigating the criterion validity of ISE measures. We extended previous research in

contributing evidence on the criterion validity of two different measures of ISE: the well established Implicit Association Test (IAT; Greenwald & Farnham, 2000) and a recently developed self-judgment task under cognitive load (LOAD; Riketta & Gebauer, *subm.*). We used criterion variables beyond self-report, i.e., experimenter and observer-ratings of self-confident behavior and word use. Specifically, we hypothesized that there would be a double dissociation pattern of ISE and ESE in which ISE alone should predict aspects of self-confident behavior that ESE is not able to predict – and vice versa.

Study 1

Study 1 aimed at a conceptual replication and extension of results that found ISE to predict experimenter-ratings of anxiety (Spalding & Hardin, 1999). The study of Spalding and Hardin (1999) was extended in two ways. First, we used a different measure of ISE, the newly developed and validated LOAD task (Riketta & Gebauer, *subm.*). Second, we included a measure of linguistic indicators of anxiety. In detail, a situation was designed in which participants completed a writing task and an interview. Non self-reported anxiety was inferred from experimenter-ratings as well as usage of anxiousness related words. Assuming a double dissociation, we hypothesized that ESE, but not ISE, would predict self-reported nervousness and anxiety, whereas ISE, but not ESE, would predict linguistic aspects of anxiety and experimenter-rated anxiety.

Method

Participants and Procedure

Seventy-nine students (62 female) participated in exchange for partial course credit. Their average age was 23.2 years ($SD = 5.2$). Because of task non-compliance and extreme scores, listwise N s ranged from 77 to 79 across various analyses. Individually, participants first completed measures of ISE and ESE as well as a measure of nervousness. Furthermore, they completed a writing task in which they briefly introduced themselves by writing ten sentences about themselves. Word usage in that personality and character sketch served as an indirect measure of anxiety. Following some other questionnaires not relevant to the present study, the experimenter asked the participants to remember and describe examples of a social situation where participants had been rejected and then to answer several questions about that situation (e.g., How did you feel? What did you think?). After that interview participants rated their anxiety during the interview and the experimenter also rated the participant's anxiety. Finally, participants were thanked and fully debriefed.

Measures

ISE. The self-judgment task under cognitive load (LOAD; Riketta & Gebauer, *subm.*) was used to assess self-esteem indirectly. After three practice trials, a block of 30 trials followed. That block consisted of 15 positive and 15 negative trait words. Participants were asked to press one of two response keys to indicate whether a trait described them or not. Traits randomly appeared on the screen until the participant responded. To ensure spontaneous responding, cognitive load was established by asking participants to learn an eight-digit number, rehearse it during the task and re-enter it at the end of the task. They were told that their data would not be usable when entering the wrong number at the end of the task. 85% of participants entered the right number, the remaining participants had only minor errors (such as transposed digits), and so all participants were retained in further analyses. ISE scores were computed by subtracting the number of negative traits endorsed from the number of positive traits endorsed. Higher LOAD scores reflect higher ISE.

ESE. The 32-item Multidimensional Self-Esteem Scale (MSES; Schütz & Sellin, 2006) was used to measure ESE. Responses were made on seven-point scales with end points labeled *not at all* (1) and *very much* (7) or *never* (1) and *always* (7), respectively.

Nervousness. Nervousness was measured with the respective subscale of the Trier Personality Inventory (Becker, 1989). Items were answered on a four-point scale (1 = *always*, 4 = *never*). To obtain comparability to the measures used in this study, answers were inverted so that high scores represent the tendency to be anxious.

Linguistic Aspects of Anxiety. The task was based on a self-reported reflection of participants' inner thoughts and feelings, however, word use served as an indirect measure of anxiety. We used the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis, & Booth, 2001; German version by Wolf, et al., 2008) to determine the frequency with which participants used anxiety related words (e.g., nervous, afraid) in their free self-descriptions. The score expressed the proportion of words that fell into the Anxiety category of the LIWC.

Self- and Experimenter-Rated Anxiety after Interview. Self-ratings of how anxious participants felt during the interview were assessed with a ten-item scale (e.g., "anxious", "nervous", "tense"). Experimenter-ratings were obtained by the experimenter who was unaware of the participants' scores on ISE and ESE. It was calculated as the mean of a two-item scale ("anxious", "nervous") on which the experimenter rated overall impression

of participants' apparent anxiety. All ratings were administered on the same seven-point scale (1 = *not at all*, 7 = *very much*).

Results and Discussion

Table 1 shows the descriptive statistics and correlations of the main variables. All measures of self-esteem had satisfactory to high reliabilities. Further, ESE and ISE were not significantly correlated, albeit the correlation was somewhat higher than in previous studies. ESE was marginally significantly correlated with self-rated anxiety and significantly with nervousness but not with experimenter-rated and LIWC anxiety. ISE did not correlate with self-rated anxiety and nervousness but with experimenter-rated and anxiety in the LIWC.

In two structural equation models, we tested the double dissociation pattern, controlling for random error and systematic variance of the measurement methods. To replicate the results reported by Spalding and Hardin (1999), self-rated and experimenter-rated anxiety were included as paired criteria in the first model. In the second model, we analyzed self-rated nervousness and anxiety measured with the LIWC. ESE and ISE were measured with the subscales of the MSES subscales and with equivalent split-halves of the LOAD task, respectively (see Figure 1). In addition to the predictive relations between self-esteem and criterion measures, we analyzed the overall fit of both models.

As can be seen in Table 2, ESE and ISE were not significantly correlated. Also, anxiety ratings of the participant and the experimenter were not correlated. However, as hypothesized, ESE predicted self-rated anxiety and ISE predicted experimenter-rated anxiety with both coefficients being significant. Moreover, data confirmed the specified model and goodness-of-fit indices can still be regarded as an acceptable fit: $\chi^2(33) = 59.22$, $p = .003$, $\chi^2/df = 1.79$, RMSEA = .10, SRMR = .08, CFI = .90, and TLI = .86.

For the second model (see Table 2), neither correlations between ESE and ISE nor between self-rated nervousness and LIWC anxiety were significant. However, as hypothesized, ESE predicted participants' ratings of nervousness, and ISE predicted indirectly measured anxiety with the LIWC. The specified model was confirmed as indicated by fit-indices, located at the boundary values of an acceptable fit: $\chi^2(33) = 65.88$, $p = .001$, $\chi^2/df = 2.00$, RMSEA = .11, SRMR = .08, CFI = .90, and TLI = .86.

In sum, the findings supported our hypotheses. A double dissociation pattern of ESE and ISE was clearly confirmed in two structural equation models. ISE predicted the use of anxiety related words in a writing task and experimenter-rated anxiety after an

interview but not self-rated anxiety and nervousness. Furthermore, ESE predicted self-rated anxiety and nervousness but not experimenter-rated or LIWC anxiety.

The results of this initial study were the first to replicate Spalding and Hardin (1999), thereby extending the findings by using a recently developed measure of ISE, the LOAD task, and by using an additional unobtrusive measure of anxiety that is based on word use. To sum up, ISE predicted non-self-report criteria that can be considered indirect indicators of self-confidence (i.e., observer-judgments, word use).

Study 2

Our investigation of the criterion validity of measures of ISE had yet to include the most promising and most widely used methodology, the IAT. This method was found to be a reliable measure of ISE (Rudolph, et al., 2008) and, furthermore, it was successfully validated when measuring aspects of the self-concept such as anxiety and shyness (Asendorpf, et al., 2002; Egloff & Schmukle, 2002). Further, we extended Study 1 in various ways. To begin with, we were able to test predictive validity because ISE and ESE were measured one week before the assessment of the criteria. Furthermore, behavioral criteria were not provided by the experimenter during the experimental situation but by independent observers who rated videotaped sequences. Those ratings were also tested for interrater-reliabilities. And, most importantly, we employed more fine-grained behavioral observations of controlled and spontaneous nonverbal self-confident behaviors.

For this purpose, we analyzed assertive nonverbal behavior during a presentation. We distinguished between *adaptors* and *illustrators* – both nonverbal body movements that vary with respect to the required level of intentionality and awareness (Ekman & Friesen, 1969). Whereas *adaptors* are unintentional gestures that are used spontaneously and less consciously, *illustrators*, on the other hand, are intentionally used hand movements that underline spoken words. *Adaptors* might therefore be triggered more by behavioral schemata that are directly activated through spreading activation in the impulsive system, whereas *illustrators* might be viewed as “reasoned action” that should be influenced more by processes in the reflective system, such as reasoning and intending (Strack & Deutsch, 2004).

As in Study 1, we assumed a double dissociation pattern. We hypothesized that ISE (here conceptualized as the associative representation of the self and measured with the IAT), would predict spontaneous behavior better than ESE, whereas ESE, the propositional representation of the self, was expected to predict controlled behavior better than ISE.

Method

Participants and Procedure

Forty nine students (37 female) participated in exchange for partial course credit. Their average age was 22.8 years ($SD = 4.3$). Students signed up for two sessions in the laboratory. In the first session, groups of up to three participants completed an IAT as a measure of ISE and a self-report measure of ESE. Exactly one week later, participants individually attended a second session during which they were videotaped. They received instructions for a public speaking task that was adopted from Egloff and Schmukle (2002). Participants were asked to give a three minute oral presentation on a scientific text. The instructor emphasized that the ability to comprehend and present scientific issues under time pressure is an important ability in successfully completing academic studies. Participants were told that they had ten minutes to prepare a speech that would later be scored by a panel of judges who would rate their ability.

After completing the presentation, participants were debriefed in detail. They were informed that the study was not designed to measure their ability and that the videotapes would only be viewed by trained raters for scientific purposes.

Measures

ISE. Participants completed the generic affective variant of the self-esteem Implicit Association Test (IAT; Greenwald & Farnham, 2000; Rudolph, et al., 2008). In this test, participants are asked to co-classify self-related and self-unrelated stimuli alongside positive and negative stimuli in a five-block structure and procedure (for a detailed description see Rudolph, et al., 2008). Critical block order was kept constant so as not to confound person variance with method variance (Banse, Seise, & Zerbes, 2001). IAT scores were computed using the scoring algorithm (D -index) recommended by Greenwald, Nosek, and Banaji (2003). Higher IAT scores reflect an automatic preference for Self over NonSelf.

ESE. As in Study 1, ESE was measured with the MSES (Schütz & Sellin, 2006).

Controlled and Spontaneous Self-Confident Behavior. Measures of controlled and spontaneous self-confident behavior were rated by a trained observer who was unaware of participants' ESE and ISE. A category system distinguished between spontaneous and controlled nonverbal behavior, adaptors and illustrators (Ekman & Friesen, 1969), respectively. The observer provided the rating separately on four-point scales (1 = *not at all*, 4 = *very often*) in each category: facial adaptors (nervous mouth movements, e.g., lip biting or licking and pressing the lips together), body adaptors (nervous hand movements,

e.g., rubbing or squeezing fore arm or upper arm and playing with one's fingers), illustrators (hand movements, e.g., emphasizing a particular phrase and sketching the direction of ones thoughts), control of posture and control of speech (perceived control of posture and fluency of speech on four-point scales, 1 = *very tense* or *dysfluent*, 4 = *very relaxed* or *fluent*). Ratings of illustrators, control of posture and speech are assumed to reflect determinants of the reflective system and were aggregated into a mean rating of controlled behavior. High scores indicate a high level of *controlled self-confident behavior*. In contrast, adaptors are assumed to reflect determinants of the impulsive system. Those ratings were reversed coded and formed a mean rating of *spontaneous self-confident behavior*. Higher scores reflect a higher level of self-confident spontaneous behavior (i.e. less frequent use of adaptors during the speech).

Results and Discussion

The bottom section of Table 1 shows the descriptive statistics and correlations of the variables. The MSES was highly reliable, and the reliabilities of the IAT and the behavior ratings were satisfactory. Again, ESE and ISE were not correlated. Also controlled and spontaneous self-confident behaviors were not correlated.

In order to test our hypothesized double dissociation patten, we employed structural equation modeling. As in Study 1, ESE and ISE were measured with the MSES subscales and four equivalent quarters of the IAT, respectively. With nonverbal behavior, we did not construct latent variables because, as in previous research (Asendorpf, et al., 2002) strong individual response hierarchies, indicated by low correlations ($-.20 < r < .30$) between aggregates of the observed variables, were found. Figure 1 illustrates the hypothesized model and corresponding results are displayed in Table 2. Estimated correlations between self-esteem indices and behavioral variables were not significant. However, ESE predicted controlled self-confident behavior and ISE predicted spontaneous self-confident behavior. The model constituted a good fit: $\chi^2(52) = 60.01$, $p = .21$; $\chi^2/df = 1.16$, $RMSEA = .06$, $SRMR = .11$, $CFI = .94$, and $TLI = .93$.

Overall, the pattern replicated the results of Study 1 and established predictive validity for direct and indirect measures of self-esteem. The direct measure of self-esteem was significantly correlated with controlled self-confident behavior while the indirect measure of self-esteem was not. Spontaneous indicators of self-confident behavior correlated significantly with ISE but not with ESE. Most interestingly, this study replicated the double dissociation pattern of ESE and ISE using another indirect measure of self-esteem, the IAT.

General Discussion

The present study is one of the first that provides compelling evidence for the criterion validity of indirect measures of self-esteem. Across two studies and three sets of dependent variables, measures of ISE (the LOAD task and the IAT) predicted self-confident behavior (either rated by an experimenter or by an observer or as reflected in word usage) that the measure of ESE was not able to predict. In accordance with dual-model approaches, the data support the hypothesized double dissociation pattern between ESE and ISE. Indirect measures of self-esteem tap into spontaneous and automatic self-evaluation, and therefore predict automatic and spontaneous behavior that direct measures of self-esteem are not able to predict. On the other hand, direct measures of self-esteem rely on self-reports on participants' inner feelings toward the self, and therefore predict controlled and deliberative behaviors that indirect measures of self-esteem are not able to predict.

The present study contributes to the existing literature in several ways. First, we provided converging evidence for the criterion validity of indirect measures of self-esteem. In spite of promising reliability and construct validity (see Rudolph, et al., 2008), results showing significant main effects of ISE had been rare following the initial study by Spalding and Hardin (1999). In the present research, we have shown independent effects of two different measures of ISE in predicting observed behavior. Using indicators of self-confidence that tap into impulsive processes, we found replicable evidence of the predictive validity of ISE. Thus, a lack of correspondence between measures and criteria may in fact have led to low validity of ISE measures in previous studies.

Second, our findings support the differential predictive validity of ISE and ESE in a double dissociation pattern. More precisely, ESE was uniquely related to self-reported and controlled nonverbal indicators of self-confident behavior, whereas ISE was uniquely related to non-self-reported, indirectly assessed, and spontaneous nonverbal indicators of self-confidence. There were neither significant relations between ESE and observed or spontaneous self-confident behavior nor between ISE and self-reported or controlled self-confident behavior.

Third, our results once again demonstrate that questionnaire ratings should be complemented by observer ratings of personality or observations of actual behavior. As expected, ISE predicted observer-rated anxiety and spontaneous behavior. Correlations were slightly above $|.30|$ which is as high as the typical coefficient that has been found for the relation between behavior and self-report measures (Mischel, 1968). Overall, there are

only few studies available that show the prediction of behavior on the basis of personality traits. Thus, our research responds to recent calls for a broader range of data in personality psychology (Funder, 2001), and for using behavioral data (Baumeister, Vohs, & Funder, 2007). In doing so, we extend previous research by focusing on indirect measures and ISE.

Finally, the present research provides evidence on the usefulness of the *construct* of ISE in the study of human behavior: By predicting behavior that ESE was not able to predict, we extend findings on the well established construct of self-esteem. For a thorough understanding of self-esteem processes, ISE apparently is a very important complement to ESE.

Despite the interesting and consistent results, the present study has some limitations. First, in both experimental designs, data was gathered during laboratory observations and it is not clear how they generalize to everyday settings. In weighing pros and cons, we preferred this approach to benefit from standardized setting (Lakes & Hoyt, 2008). Being aware of this constraint, we used different sources of ratings in the two studies to generalize the results beyond one specific design. As some nonverbal behavior we assessed (e.g., illustrators) might be relatively inconsistent between one type of situation (e.g., giving a talk) and another (e.g., first date), it would still be interesting to see whether our results are replicable in everyday settings and when using acquaintance ratings of behavior. First evidence for the usefulness of acquaintance ratings has been provided by (Schröder-Abé, Rudolph, & Schütz, 2009) who found participants' ISE to be related to friends' ratings of the participants' ISE.

Due to time-consuming experimental sessions and behavior coding, the second limitation of the present study is its sample size which may be considered as small for structural equation modeling. While some researchers suggest a *the more the merrier* strategy, 20% of studies still use small samples (Kline, 2005). Although the power of statistical tests in small samples is limited, we replicated results between the two studies.

Our findings also provide some hints for future research. As has been shown by Friese, Hofmann and Schmitt (2008), the predictive validity of indirect measures is typically high when there is little opportunity or motivation to control behavior. We differentiated between aspects of behavior that vary with respect to controllability (e.g., illustrators and adaptors), and found differential criterion validity, as predicted. We did not manipulate behavioral control in the present study, however. Future research should, therefore, include settings that vary with respect to such aspects. For example, replicating

the results in an experimental setting in which social desirability is included as an independent variable, may provide even more differentiated results.

Clearly, after a decade of research on measures of ISE, a lot of questions have not yet been answered. But the results of this study helped to provide evidence on the predictive validity of those measures. Importantly, we used theoretically and empirically derived behavioral indicators and found them to be predicted by ISE and ESE. Providing a double dissociation, direct and indirect measures of self-esteem differentially predicted behavior. In predicting behavior over and above ESE, it seems therefore essential to consider ISE in future studies.

References

- Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83*, 380-393.
- 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.
- Baumeister, R. F. (1998). The Self. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The Handbook of social psychology* (Vol. 1). New York: McGraw-Hill.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger-movements. Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*, 396-403.
- Becker, P. (1989). Trierer Persönlichkeitsfragebogen (TPF) [Trier Personality Inventory]. Göttingen: Hogrefe.
- Bosson, J. K., Swann, W. B. J., & 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.
- Conner, T., & Barrett, L. F. (2005). Implicit self-attitudes predict spontaneous affect in daily life. *Emotion, 5*, 476-488.
- De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction*. (pp. 11-28). Thousand Oaks, CA: Sage Publishers.
- 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.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage and coding. *Semiotica, 1*, 49-98.
- Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology, 94*, 334-346.
- 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.
- Franck, E., De Raedt, R., & De Houwer, J. (2007). Implicit but not explicit self-esteem predicts future depressive symptomatology. *Behaviour Research and Therapy, 45*, 2448-2455.
- Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behavior? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology, 19*, 285-338.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology, 52*, 197-221.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022-1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & 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.
- Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology, 85*, 969-978
- Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*, 16-32.
- Kline, R. B. (2005). Principles and practice of structural equation modeling. New York: Guilford.
- Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80*, 669-685.

- Lakes, K. D., & Hoyt, W. T. (2008). What sources contribute to variance in observer ratings? Using generalizability theory to assess construct validity of psychological measures. *Infant & Child Development, 17*, 269-284.
- Leary, M. R., & MacDonald, G. (2003). Individual differences in self-esteem: A review and theoretical integration. In M. R. Leary & T. J. Tangney (Eds.), *Handbook of Self and Identity* (pp. 3-14). New York: Guilford.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition, 19*, 625-666.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic Inquiry and Word Count (LIWC): LIWC 2001. Mahwah, NJ: Erlbaum.
- Riketta, M., & Gebauer, J. (subm.). Loving oneself deep inside: Implicit self-esteem is higher than explicit self-esteem.
- Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. P., & Sedikides, C. (2008). Through a glass, less darkly? Reassessing convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment, 24*, 273-281.
- Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007). High implicit self-esteem is not necessarily advantageous: Discrepancies between explicit and implicit self-esteem and their relationship with anger expression and psychological health. *European Journal of Personality, 21*, 319-339.
- Schröder-Abé, M., Rudolph, A., & Schütz, A. (2009). A Single Category IAT to measure implicit self-esteem: Results on psychometric properties and discrepancies with explicit self-esteem. Paper presented at the International Small Group Meeting on the Transformation of Reaction-Time Based Measurement Approaches.
- Schröder-Abé, M., Rudolph, A., Wiesner, A., & Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *International Journal of Psychology, 42*, 174-183.
- Schütz, A., & Sellin, I. (2006). Die Multidimensionale Selbstwertkala (MSWS) [The multidimensional self-esteem scale] Göttingen: Hogrefe.
- Spalding, L. R., & Hardin, C. D. (1999). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10*, 535-539.

- Steffens, M. C., & Schulze König, S. (2006). Predicting spontaneous Big Five behavior with Implicit Association Tests. *European Journal of Psychological Assessment, 22*, 13-20.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220-247.
- Upshaw, H. S., & Yates, L. A. (1968). Self-persuasion, social approval, and task success as determinants of self-esteem following impression management. *Journal of Experimental Social Psychology, 4*, 143-152.
- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology, 95*, 1202-1216.
- Wilson, T. D., & Dunn, E. W. (2004). SELF-KNOWLEDGE: Its limits, value, and potential for improvement. *Annual Review of Psychology, 55*, 493-518.
- Wolf, M., Horn, A. B., Mehl, M. R. H., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computergestützte quantitative Textanalyse. Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count [Computer-aided quantitative textanalysis: Equivalence and reliability of the German adaptation of the Linguistic Inquiry and Word Count]. *Diagnostica, 54*, 85-98.

Table 1
Correlations between indices of self-esteem and criteria

Study 1										
Measures	Min	Max	M	SD	α^a	1	2	3	4	5
1. MSES	1.97	6.25	4.22	1.00	.93	-				
2. LOAD	-8	15	8.42	4.13	.73	.19	-			
3. Self-Rated Anxiety	1.00	6.10	2.51	1.05	.89	-.20 [†]	-.08	-	-	
4. Self-Rated Nervousness	1.09	3.82	2.52	.85	.95	-.74**	-.13	.24*		
5. Experimenter-Rated Anxiety	1.00	5.50	1.85	1.21	.79	-.11	-.25*	.11	.12	-
6. LIWC Anxiety	.00	4.17	.43	.92	n.a.	-.04	-.30**	-.01	.10	-.13
Study 2										
Measures	Min	Max	M	SD	α^a	1	2	3		
1. MSES	2.94	6.28	4.88	.72	.92	-				
2. IAT	-.08	1.22	.83	.30	.78	-.07	-			
3. Controlled Behavior ^b	1.21	2.86	2.01	.40	.72	.35*	.15			
4. Spontaneous Behavior ^b	1.00	2.80	1.85	.40	.82	-.08	.29*		-.05	

Note. $N_{\text{Study 1}} = 77 - 79$, $N_{\text{Study 2}} = 49$. MSES = Multidimensional Self-Esteem Scale, LOAD = self-judgment task under cognitive load, IAT = Implicit Association Test, LIWC = Linguistic Inquiry and Word Count. ^a Internal consistency for self-esteem measures and anxiety ratings; Interrater agreement between two observer judgments of behavioral ratings using independent ratings of 15% randomly selected participants. ^b Average rating of behavioral ratings: Higher scores indicate more self-confident controlled and spontaneous behavior (i.e., more illustrators, less adaptors). [†] $p < .10$, * $p < .05$, ** $p < .01$.

Table 2

Estimated path coefficients and correlation coefficients of the structural equation models testing a double dissociation (cf. Figure 1)

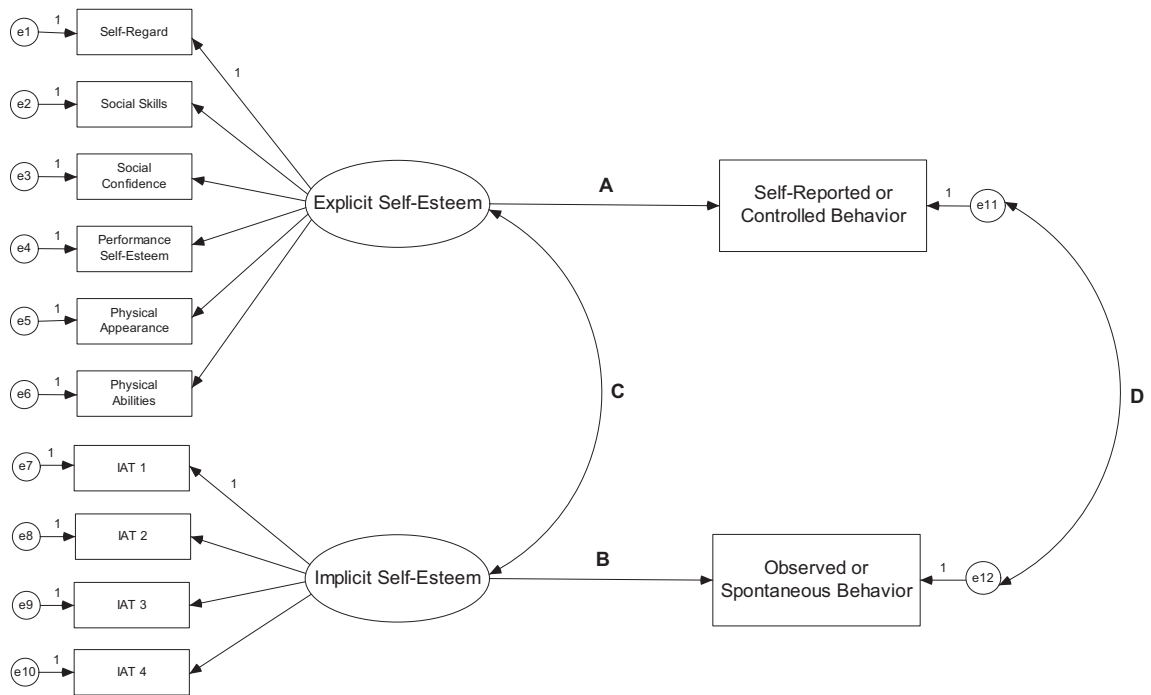
	A	B	C	D
Model 1 (Study 1)	-.26***	-.33*	.20	.09
Model 2 (Study 1)	-.84***	-.35*	.29	.00
Model 3 (Study 2)	.40**	.31*	-.01	-.10

Note. Within all models, A represents the path coefficients of ESE predicting self-reported or controlled behavior, B represents the path coefficients of ISE predicting observed or spontaneous behavior, C is the correlation coefficient between the self-esteem measures, D is the correlation coefficient between the behavioral indicators. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure caption

Figure 1. The structural equation model illustrates the double dissociation between explicit (ESE) and implicit (ISE) self-esteem. With ESE, the latent variable was measured with the subscales of the Multidimensional Self-Esteem Scale. With ISE, two equivalent halves of the LOAD task were created (Study 1), or four equivalent quarters of the Implicit Association Test (Study 2). A represents the estimated path coefficients of ESE predicting self-reported or controlled behavior, B represents the estimated path coefficients of ISE predicting observed or spontaneous behavior, C is the correlation between ESE and ISE, D is the correlation between the behavioral indicators. Respective values of A, B, C and D are displayed in Table 2.

Figure 1



Self-Esteem Discrepancies and Defensive Reactions to Social Feedback

Michela Schröder-Abé¹

Almut Rudolph¹

Anja Wiesner²

Astrid Schütz¹

¹Chemnitz University of Technology, Chemnitz, Germany

²Johannes Kepler University, Linz, Austria

Abstract

Recent research has provided increasing evidence that discrepancies between implicit and explicit self-esteem (SE) are related to defensiveness. However, only one pattern, called fragile SE (a combination of high explicit and low implicit SE), has been the focus of research so far. Relatively little attention has been paid to the second possible SE discrepancy (low explicit/high implicit). We propose that both types of discrepancies are maladaptive, because they indicate a lack of integration in self-representation. We conducted two studies on the correlates of discrepant SE in the sphere of defensiveness. We hypothesized that persons with discrepant SE exhibit more defensive behavior than individuals with congruent SE. In two student samples explicit SE was measured by means of the Multidimensional Self-Esteem Scale. Implicit SE was measured with the Name Letter Technique (Study 1) and an Implicit Association Test (Study 2). In Study 1 (N = 102) we examined effects of implicit and explicit SE on defensive reactions to ambiguous statements. For this purpose, we adapted the Ambiguous Statements Task (AST, Tafarodi, 1998), a procedure that measures participants' tendency to interpret ambiguous expressions in a positive vs. negative manner. We found that the combination of low explicit SE and high implicit SE was associated with a more positive interpretation of the ambiguous statements. Study 2 (N = 87) investigated the effects of discrepant SE on reactions to SE threat in an experimental setting. The time participants spent reading positive or negative social feedback was used as an unobtrusive measure of defensiveness. We found that people with discrepant SE went through the negative feedback faster than people with congruent SE. In sum, the studies provide converging evidence that both kinds of discrepancies between implicit and explicit SE are related to defensive behavior.

Key words: implicit self-esteem, explicit self-esteem, discrepancies, defensiveness, fragile self-esteem

Nothing is easier than self-deceit. For what each man wishes, that he also believes to be true. (Demosthenes, trans. 1930, Olynthiac III, para. 19)

The Greek orator Demosthenes had already addressed the phenomenon of self-deception in the fourth century B.C. Psychological research has examined defensive behavior starting with Freud (1894). Since then, there has been a growing body of research dealing with this issue (see Baumeister, Dale, & Sommer, 1998, for a review). Modern personality and social psychology has mainly focused on defensive behavior aimed at protecting self-esteem (SE). The present research studies defensive reactions as a function of discrepancies between implicit and explicit SE.

Discrepancies between Implicit and Explicit SE

Self-esteem, the “positivity of a person’s evaluation of self” (Baumeister, 1998), is related to various positive outcomes, such as life satisfaction (Diener, 1984) and emotional stability (Robins, Hendin, & Trzesniewski, 2001). Highly positive self-views, however, are also related to aggressive and defensive behavior (Baumeister, Smart, & Boden, 1996; Blaine & Crocker, 1993; Bushman & Baumeister, 1998). In order to resolve these apparently contradictory findings, several attempts have been made to distinguish secure and fragile forms of high SE (Kernis, 2003). One way to do so is the examination of discrepancies between explicit and implicit SE (Bosson, Brown, Zeigler-Hill, & Swann, 2003; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003). Implicit SE is defined as an individual’s overlearned, automatic and nonconscious self-evaluation (Greenwald & Banaji, 1995). Several methods of measuring implicit SE have been proposed, but only the Name Letter Preference Task (Nuttin, 1985) and the Implicit Association Test (IAT; Greenwald & Farnham, 2000) have demonstrated satisfactory reliability and validity (Bosson, Swann, & Pennebaker, 2000). Implicit SE refers to the experiential mode, while explicit SE refers to the cognitive mode as distinguished by dual-process models of information processing (e.g., Epstein, 1994; Wilson, Lindsey, & Schooler, 2000). Accordingly, explicit SE is attained through conscious and rational processing of self-relevant information, whereas implicit SE is the product of automatic, intuitive processing of affective experiences (Epstein & Morling, 1995) and is at least partly influenced by social interactions in early life (DeHart, Pelham, & Tennen, 2006). If people receive diverging information through these different pathways, discrepancies between implicit and explicit SE might develop.

There are two possible variants of discrepant SE: (a) the combination of high explicit SE and low implicit SE, and (b) the combination of low explicit SE and high implicit SE. So far, most research has concentrated on the former, referred to as “defensive high SE” (Jordan et al., 2003), or “fragile high self SE” (Bosson et al., 2003). Recent studies have found self-enhancement (e.g., unrealistic optimism, and smaller differences between actual and ideal self; Bosson et al., 2003), defensiveness (Jordan et al., 2003; McGregor & Marigold, 2003; McGregor, Nail, Marigold, & Kang, 2005), and prejudice (Jordan, Spencer, & Zanna, 2005) to be related to this discrepancy. This dovetails with clinical accounts on narcissism (e.g., Horney, 1937; Kernberg, 1975), which describe individuals with overtly positive self-views who unconsciously evaluate themselves negatively. When their fragile SE is threatened they engage in defensive behaviors to protect it.

Relatively little attention has been paid to the second possible SE discrepancy: the combination of low explicit and high implicit SE. Although not focusing on this group, Bosson et al. (2003) found higher self-enhancement and Jordan et al. (2003) found higher defensiveness among people with low explicit and high implicit SE compared with people low in both types of SE. It has been argued that implicit SE buffers against the negative effects of low explicit SE (Bosson et al., 2003) or motivates tendencies to restore historically positive self-views (Jordan et al., 2003). So far, these interpretations are inconclusive, and it remains unclear whether this discrepancy is adaptive or not.

In contrast to the buffer hypothesis presented above, one could reason that discrepancies of any kind are maladaptive because they indicate deficient integration of self-representation. This interpretation parallels findings on ambivalent attitudes, i.e. attitudes that simultaneously comprise positive and negative dimensions (Thompson & Zanna, 1995). Such conflicting attitudes often result in internal tensions and negative emotions (Hass, Katz, Rizzo, Bailey, & Moore, 1992). Moreover, clinical theorists have identified subtypes of narcissism characterized by overtly negative self-views that conceal less conscious feelings of entitlement and superiority. Similarly, Dickinson and Pincus (2003) described individuals with “vulnerable narcissism” as overtly modest, but having underlying grandiose expectations for themselves. Some theorists (Millon & Davis, 1996; Volkan & Ast, 1994) state that this form of discrepancy might emerge from originally high SE which has been lessened later in life, because the person was unable to live up to overly high expectations. Accordingly, we call this kind of discrepancy “damaged SE.”

Forms of Defensiveness

Defensive behavior can take many forms (Baumeister et al., 1998). For example, recent studies investigating SE discrepancies have operationalized defensiveness as conviction and perceived social consensus about social issues, in-group favoritism, and dissonance-reduction. The present research conceptualizes defensiveness as biased processing of social feedback information.

Baumeister and Newmann (1994) identified two pathways through which cognitive processes can be altered: (a) controlling the interpretation of the information, and (b) regulating the collection and encoding of information. Accordingly, people can (a) try to interpret negative feedback in a more favorable way; for example, by discrediting sources of criticism (Pyszczynski, Greenberg, & Holt, 1985). Individuals can also (b) try to prevent the information from entering mind altogether; for example, by processing the information in a rushed and superficial way (Baumeister & Cairns, 1992; Bonanno, Davis, Singer, & Schwartz, 1991).

The present research investigated how discrepancies between implicit and explicit SE are connected with defensive reactions to social feedback. We hypothesized that both kinds of SE discrepancies would predict defensiveness.

Study 1

We first examined whether SE discrepancies are related to defensiveness as a reaction to relatively neutral social feedback. Defensiveness was operationalized as the positively biased interpretation of ambiguous scenarios. We predicted that individuals with discrepant SE would interpret this ambiguous social feedback in a more favorable way than individuals with congruent SE.

Method

Participants and Procedure

A total of 102 students (80 female) participated in exchange for partial course credit. Mean age of participants was 22.69 years ($SD = 3.74$). The study was conducted in the laboratory in groups of up to four individuals. Participants first completed measures of implicit and explicit SE. Subsequently, they completed several personality scales irrelevant to the present study as well as the Ambiguous Statements Task (AST). Finally, participants were debriefed and thanked.

Measures

Initials Preference Task (IPT). The IPT as an index of implicit SE is based on the assumption that peoples' initials are closely connected to the self and that peoples'

preferences for their initials relative to other letters express their implicit attitudes towards themselves (name letter effect, cf. Nuttin, 1985). Participants evaluated each letter of the alphabet on scales ranging from 1 (I dislike this letter very much) to 7 (I like this letter very much). The letters were presented one by one in a random order on a computer screen. In order to calculate initials preference scores, each participant's average liking for all letters that were not his/her initials were subtracted from the liking for his or her initials. This idiographic score (Bosson et al., 2000; 2003) controls for individual differences in rating tendencies and liking for objects (e.g., letters) in general. As the liking for non-initial letters was weakly correlated to the extent to which ambiguous statements were interpreted positively ($r = .18$), this score seems most appropriate.

Multidimensional Self-Esteem Scale (MSES). Explicit SE was measured using the total score of the MSES (Schütz & Sellin, 2006), the German adaptation of the scale by Fleming & Courtney (1984). Responses were made on seven-point scales with end points labeled *not at all* (1) and *very much* (7) or *never* (1) and *always* (7), respectively.

Ambiguous Statements Task (AST). Participants' tendency to interpret ambiguous statements in a positive way was measured by means of the AST by Tafarodi (1998). The instruction asked participants to vividly imagine a scenario in which an acquaintance directed an ambiguous phrase at them (e.g., "You must be kidding."). After imagining each of the 13 scenarios, participants indicated whether the expression reflected a positive or negative feeling toward them and then rated the intensity of the feeling on scales from *very slightly intense* (1) to *extremely intense* (7). Intensity ratings of each statement were assigned a negative sign if the statement was interpreted in a negative manner. Average intensity ratings across all 13 statements were computed. Higher scores reflected a tendency to interpret ambiguous statements in a positive manner.

Results and Discussion

Descriptive statistics and internal consistencies (Cronbach's Alpha) of all measures are displayed in Table 1. The mean of $-.28$ for the Ambiguous Statements Task indicates that, on average, participants perceived the ambiguous scenarios as slightly negative. The correlation of IPT scores and scores on the MSES was $-.12$ (*ns*), indicating that implicit and explicit SE were only weakly related to each other.

To determine whether implicit and explicit SE were related to defensiveness, we conducted a multiple regression analysis with explicit SE, implicit SE, and the interaction between them as predictors. Scores on the MSES and the IPT were centered, and the interaction was represented by the cross-product vector (Aiken & West, 1991). We

regressed defensiveness (i.e. people's tendency to interpret ambiguous statements positively) onto these variables.

We found a significant main effect of explicit SE ($\beta = .20$; $t(97) = 2.04$; $p = .04$) and the predicted interaction between implicit and explicit SE ($\beta = -.23$; $t(97) = -2.39$; $p = .02$). The main effect of implicit SE was not significant ($\beta = .08$; $t(97) = .79$; $p = .43$). Following Cohen and Cohen (1983), we tested simple slopes at values one standard deviation above and below the mean of explicit SE to further explore these interactions. As Figure 1 shows, among people with high explicit SE, implicit SE did not moderate the tendency to interpret ambiguous statements positively ($\beta = -.16$; $t(97) = -1.25$; $p = .22$). It did, however, among people with low SE, as individuals low in explicit SE and high in implicit SE interpreted ambiguous statements more positively than individuals low in explicit and implicit SE ($\beta = .32$; $t(97) = 2.12$; $p = .04$).

The results thus partially supported our predictions. We found that individuals with damaged SE (low explicit, high implicit) interpreted ambiguous scenarios more positively than individuals with congruent low SE. However, contrary to expectations, individuals with fragile SE (high explicit, low implicit) did not interpret the ambiguous statements significantly more positively than individuals with congruent high SE. One reason for this finding might be that individuals with fragile SE only react defensively when their overtly positive self-views are threatened (e.g. McGregor & Marigold, 2003; McGregor et al., 2005).

Study 2

In Study 2, we sought to extend the findings of Study 1 in several ways. Firstly, we introduced a SE threat by manipulating social feedback (rejection versus acceptance). By doing so, we aimed to test whether defensive behavior occurs in individuals with fragile SE (high explicit, low implicit) under threat only. Secondly, we used a different measure of implicit SE. We administered a SE IAT in order to increase generalizability of the findings to other measures of implicit SE. Thirdly, we operationalized defensiveness in a different way. We measured how fast participants read social feedback and interpreted higher speed of reading as defensiveness. We theorized that this would enable participants to avoid deeper processing of potentially negative information (cf. Baumeister & Cairns, 1992; Bonanno et al., 1991). We predicted that, especially in the rejection condition, individuals with both forms of discrepant SE (fragile and damaged SE) would react more defensively to social feedback than individuals with congruent SE.

Method

Participants and Procedure

Eighty-seven students (58 female) participated in exchange for partial course credit. Mean age of participants was 22.67 years ($SD = 4.86$).

Participants were run individually in the laboratory. After completing implicit and explicit SE measures, participants were subjected to a procedure for manipulating social rejection versus acceptance (modified from Buckley, Winkel, & Leary, 2004). Participants were told that the study was on new media and getting to know people via the Internet. Participants completed an online questionnaire about themselves. Questions were adapted from the Relation Closeness Induction Task (Sedikides, Campbell, Reeder, & Elliott, 1999). Participants were told that the information about themselves would be sent electronically to another participant in the laboratory room next door, who would then send the participant feedback and evaluate him or her on the basis of the information provided in the questionnaire. The other participant would also indicate how much s/he wanted to work with the participant during the remainder of the study.

Participants were given feedback after three minutes of waiting. In fact, there was no other participant, but the questionnaire and the feedback were programmed in Inquisit 1.33 software. Participants were randomly assigned to one of two experimental groups: One group (39 participants) received positive feedback indicating that the other participant had evaluated them positively (e.g., likable) and wanted to cooperate with them (acceptance condition). The other group (48 participants) received negative feedback indicating that the other participant had formed a negative impression of them (e.g., boring) and did not want to work with them (rejection condition). The feedback consisted of eight ratings on several traits that were rated on 7-point semantic differentials (e.g., from *dull* to *intelligent*) and answers to three questions (e.g., “How much do you want to work with the participant?”) on 10-point scales ranging from 1 (*not at all*) to 10 (*very much*). In the acceptance condition, participants received ratings in the range of 6 or 7 (semantic differentials) and 8 to 10 (rating scales). In the rejection condition, participants received ratings of 1 or 2 (semantic differentials) and 1 to 3 (rating scales). As an unobtrusive indicator of defensiveness, we summed up the time participants spent on reading the feedback items on their computer screen.

After completing a series of measures irrelevant for the present study, participants completed two manipulation check items worded “How much did the other participant

want to work with you?” and “How much did the other participant want to get to know you?” Finally, participants were debriefed and thanked.

Measures

Implicit Association Test (IAT). As a measure of implicit SE participants completed an IAT (Greenwald, McGhee, & Schwartz, 1998). We used the affective generic variant of the SE IAT (Greenwald & Farnham, 2000; Rudolph, Schröder, & Schütz, 2006) with target-concept discrimination between 3 self-relevant and 3 non-self-relevant words, and attribute discrimination between 6 pleasant and 6 unpleasant words. The IAT consisted of five blocks of trials (see Table 2 for a detailed illustration of the SE IAT). Data from the combined blocks (3 and 5) were used to compute IAT scores (*D* measure) following Greenwald, Nosek, and Banaji (2003).

In order to compute internal consistency, the IAT was split into four subtests of equal length. Each subtest contained equal numbers of target and attribute stimuli. To avoid biases caused by order effects, each subtest contained trials of all combined blocks (block 3 and 5). Internal consistency was evaluated by computing Cronbach’s Alpha with the four IAT subtests (see Table 1).

Multidimensional Self-Esteem Scale (MSES). As in Study 1, explicit SE was measured using the MSES (Schütz & Sellin, 2006). For economical reasons, we administered a short version of the scale.

Results and Discussion

Manipulation checks

Participants accurately perceived the degree to which the alleged other participant wanted to work with them, $t(85) = -42.73, p < .001, d = -7.95$. Participants also correctly identified the degree to which the other participant wanted to get to know them, $t(85) = -40.70, p < .001, d = -8.64$. Internal consistency of these two measures was .98.

Descriptive statistics

Descriptive statistics and internal consistencies (Cronbach’s Alpha) of all measures are displayed in Table 1. We omitted the first item of the reading speed measure (i.e., the time participants spent on reading the first trait rating on the dimension “insecure vs. secure”), because of its lengthened latency, heightened standard deviation and low item-total correlation. This produced an increase in internal consistency from .75 to .87.¹ The correlation of IAT scores (*D* measure) and scores on the MSES was .12 (*ns*), indicating that implicit and explicit SE were only weakly related to each other.

Regression analyses

We conducted multiple regression analyses with rejection-acceptance condition, explicit SE, implicit SE, and all interactions entered as predictor variables. Scores on the MSES and the SE IAT (*D* measure) were centered. We then regressed defensiveness (time spent reading the feedback) onto these variables.

There were no significant main effects of condition, MSES, or IAT scores ($ps > .25$) predicting reading speed. However, we found a significant two-way interaction between implicit and explicit SE ($\beta = .33$; $t(77) = 2.27$; $p = .03$) and a significant three-way interaction between condition, implicit and explicit SE ($\beta = -.30$; $t(77) = -2.05$; $p = .04$). Subsequently, we tested simple slopes at values one standard deviation above and below the mean of explicit SE. We did not find any significant effects in the acceptance condition ($ps > .19$), but we did find marginally significant effects in the rejection condition: Among people high in explicit SE, implicit SE was positively related to the time spent reading the feedback ($\beta = .40$; $t(77) = 1.71$; $p = .09$). That is to say, individuals with fragile SE (high explicit, low implicit) showed a tendency to read the negative feedback faster than individuals with congruent high SE. Among people low in explicit SE, implicit SE was negatively related to the time spent reading the feedback ($\beta = -.37$; $t(77) = -1.79$; $p = .08$). That is, there was a trend in individuals with damaged SE (low explicit, high implicit) to read the negative feedback faster than individuals with congruent low SE.

For a more thorough analysis, we computed separate sum scores of the trait ratings and questions concerning cooperation in the further course of the study (see Table 1 for descriptives) and conducted regression analyses for those scores separately. There were no significant main effects or interactions with the trait ratings as dependent variable ($ps > .13$). However, when using the questions about further cooperation as dependent variable, we found highly significant effects. There was a significant two-way interaction between implicit and explicit SE ($\beta = .42$; $t(77) = 3.00$; $p < .01$) and a significant three-way interaction (Figures 2 and 3) between condition, implicit and explicit SE ($\beta = -.40$; $t(77) = -2.84$; $p < .01$). Except for a marginally significant main effect of condition ($\beta = .20$; $t(77) = 1.84$; $p = .07$) there were no other significant effects of the remaining predictors ($ps > .28$). When conducting simple slope tests, we did not find significant effects ($ps > .15$) in the acceptance condition (Figure 2). In the rejection condition (Figure 3), however, the effects were significant. That is, implicit and explicit SE predicted the speed of reading the feedback mainly in the rejection condition. Among people high in explicit SE, implicit SE was positively related to the time spent reading the feedback ($\beta = .58$; $t(77) = 2.58$; $p =$

.01). Among people low in explicit SE, implicit SE was negatively related to the time spent reading the feedback ($\beta = -.40$; $t(77) = -2.00$; $p = .04$).

In conclusion, the results supported our predictions. Individuals with both forms of discrepant SE (high explicit/low implicit, and low explicit/high implicit) spent less time on reading negative social feedback than individuals with congruent SE. Interestingly, these effects were especially strong for feedback that implied rejection by another person, i.e. feedback that indicated the other person was not interested in cooperation. By contrast, no such effects were found with feedback that only implied negative evaluation (without rejection).

General Discussion

Taken together, these studies provide converging evidence of defensive behavior in both possible patterns of discrepant SE (fragile SE, i.e. high explicit/low implicit, and damaged SE, i.e. low explicit/high implicit).

Individuals with fragile SE read negative feedback faster than individuals with congruent high SE, which suggests that they avoided negative social information implying rejection by another person. This finding corresponds to recent studies showing heightened defensiveness in individuals with fragile SE (Bosson et al., 2003; Jordan et al., 2003; Jordan et al., 2005; McGregor & Marigold, 2003; McGregor et al., 2005). In contrast to Bosson et al. (2003) and Jordan et al. (2003) we did not find elevated defensiveness in individuals with fragile SE in neutral situations (Study 1). Our results correspond to the findings of Jordan et al. (2005), McGregor and Marigold (2003), and McGregor et al. (2005), who also found defensiveness in individuals with fragile SE only after threat. The results show that individuals with fragile SE get along pretty well as long as they do not experience setbacks, such as failure or rejection. Only when their SE is threatened, defensive patterns emerge.

Individuals with damaged SE displayed defensive behavior across both measures we used. Compared with individuals holding congruent low SE, they interpreted ambivalent social feedback more positively and read negative social feedback more quickly. These findings support the notion that both SE discrepancies are maladaptive and dovetail with clinical work on vulnerable narcissism (Dickinson & Pincus, 2003; Volkan & Ast, 1994).

Clinical theorists (e.g., Kernberg, 1975; Millon & Davis, 1996; Volkan & Ast, 1994) have proposed that dysfunctional early interactions lead to discrepancies, but these notions have thus far received only preliminary empirical support (DeHart et al., 2006).

Clearly more research on the development of SE discrepancies is warranted. In future research SE discrepancies and defensiveness should also be analyzed in community samples and clinical populations, such as patients with very low (Major Depression) or very high (Narcissistic Personality Disorder) explicit SE.

Defensiveness may be the result of internal tensions and negative emotions that occur when ambivalent attitudes are held (Hass et al., 1992). Whereas individuals with fragile SE showed defensiveness only after SE threat, individuals with damaged SE behaved defensively even in relatively neutral situations. They seem to almost constantly see their SE at stake and may therefore try to defend it. To further explore this hypothesis, outcomes and behaviors other than defensiveness should be observed in individuals with damaged SE. Experiencing feelings of inadequacy and threat over a long period of time may, for example, lead to impaired well-being. Preliminary evidence suggests that individuals with damaged SE show even lower indicators of psychological and physical health than individuals with congruent low SE (Schröder-Abé, Rudolph, & Schütz, in press). Additionally, it may be interesting to investigate long-term consequences of defensive behavior. Individuals who repeatedly avoid negative feedback are prone to miss the chance to learn from mistakes. In the long run, this may cause interpersonal problems. Future research could investigate possible consequences and examine how SE discrepancies are related to contentment with social relationships, and social skills.

References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. London: Sage.
- Baumeister, R. F. (1998). The self. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The Handbook of social psychology*, Vol. 1, 4th ed. New York: McGraw-Hill.
- Baumeister, R. F., & Cairns, K. J. (1992). Repression and self-presentation: When audiences interfere with self-deceptive strategies. *Journal of Personality and Social Psychology*, *62*, 851-862.
- Baumeister, R. F., Dale, K., & Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality*, *66*, 1081-1124.
- Baumeister, R. F., & Newmann, L. F. (1994). Self-regulation of cognitive inference and decision processes. *Personality and Social Psychology Bulletin*, *20*, 3-19.
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, *103*, 5-33.
- Blaine, B., & Crocker, J. (1993). Self-esteem and self-serving biases in reaction to positive and negative events: An integrative review. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 55–85). Hillsdale, NJ: Erlbaum.
- Bonanno, P. A., Davis, G. J., Singer, J. L., & Schwartz, G. E. (1991). The repressor personality and avoidant information processing: A dichotic listening study. *Journal of Personality and Social Psychology*, *62*, 386-401.
- Bosson, J. K., Brown, R. P., Zeigler-Hill, V., & Swann, W. B., Jr. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity*, *2*, 169-187.
- Bosson, J. K., Swann, W. B., & Pennebaker, J. W., Jr. (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.
- Buckley, K. E., Winkel, R. E., & Leary, M. R. (2004). Reactions to acceptance and rejection: Effects of level and sequence of relational evaluation. *Journal of Experimental Social Psychology*, *40*, 14-28.

- Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75*, 219-229.
- Cohen, J., & Cohen, P. (1983). Applied multiple regression/ correlation analyses for the behavioral sciences. Hillsdale, NJ: Erlbaum.
- DeHart, T., Pelham, B. W., & Tennen, H. (2006). What lies beneath: Parenting style and implicit self-esteem. *Journal of Experimental Social Psychology, 42*, 1-17.
- Demosthenes & Vince, J. H. (Trans.) (1930). *Orationes, olynthiacs, philippics, minor public orationes*. Cambridge: Harvard University Press.
- Dickinson, K. A., & Pincus, A. L. (2003). Interpersonal analysis of grandiose and vulnerable narcissism. *Journal of Personality Disorders, 17*, 188-207.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*, 542-575.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709-724.
- Epstein, S., & Morling, B. (1995). Is the self motivated to do more than enhance and/or verify itself? In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 9-29). New York: Plenum.
- Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology, 46*, 404-421.
- Freud, S. (1894). Die Abwehr-Neuropsychosen [The neuro psychoses of defense]. In S. Freud, *Gesammelte Werke, Volume I*, pp. 57-74. Frankfurt: Fischer.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022-1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: 1. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.

- Hass, R. G., Katz, L., Rizzo, N., Bailey, J., & Moore, L. (1992). When racial ambivalence evokes negative affect, using a disguised measure of mood. *Personality and Social Psychology Bulletin*, *18*, 786-797.
- Horney, K. (1937). *The neurotic personality of our time*. New York: Norton.
- Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and Defensive High Self-Esteem. *Journal of Personality and Social Psychology*, *85*, 969-978.
- Jordan, C. H., Spencer, S. J., & Zanna, M. P. (2005). Types of high self-esteem and prejudice: How implicit self-esteem relates to ethnic discrimination among high explicit self-esteem individuals. *Personality and Social Psychology Bulletin*, *31*, 693-702.
- Kernberg, O. (1975). *Borderline Conditions and Pathological Narcissism*. New York: Jason Aronson.
- Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, *14*, 1-26.
- McGregor, I., & Marigold, D. (2003). Defensive zeal and the uncertain self: What makes you so sure? *Journal of Personality and Social Psychology*, *85*, 838-852.
- McGregor, I., Nail, P. R., Marigold, D. C., & Kang, S.-J. (2005). Defensive pride and consensus: Strength in imaginary numbers. *Journal of Personality and Social Psychology*, *89*, 978-996.
- Millon, T., & Davis, R. (1996). *Disorders of personality: DSM-IV and beyond*. New York: Wiley-Interscience.
- Nuttin, J. M. (1985). Narcissism beyond Gestalt awareness: The name letter effect. *European Journal of Social Psychology*, *15*, 353-361.
- Pyszczynski, T., Greenberg, J., & Holt, K. (1985). Maintaining consistency between self-serving beliefs and available data: A bias in information processing. *Personality and Social Psychology Bulletin*, *11*, 179-190.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *27*, 151-161.
- Rudolph, A., Schröder, M., & Schütz, A. (2006). Ein impliziter Assoziationstest zur Erfassung von Selbstwertschätzung [An Implicit Association Test measuring self-esteem]. In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Eds.), *Theorie und Praxis Objektiver Persönlichkeitstests*, (pp. 153-163). Bern: Verlag Hans Huber.

- Schröder-Abé, M., Rudolph, A., & Schütz, A. (in press). High implicit self-esteem is not necessarily advantageous: Discrepancies between explicit and implicit self-esteem and their relationship with anger expression and psychological health. *European Journal of Personality*.
- Schütz, A., & Sellin, I. (2006). *Die multidimensionale Selbstwertkala (MSWS)* [The multidimensional self-esteem scale]. Göttingen: Hogrefe.
- Sedikides, C., Campbell, W. K., Reeder, G. D., & Elliot, A. J. (1999). The relationship closeness induction task. *Representative Research in Social Psychology*, 23, 1–4.
- Thompson, M. M., & Zanna, M. P. (1995). The conflicted individual: Personality-based and domain-specific antecedents of ambivalent social attitudes. *Journal of Personality*, 63, 259-288.
- Tafarodi, R. W. (1998). Paradoxical self-esteem and selectivity in the processing of social information. *Journal of Personality and Social Psychology*, 74, 1181-1196.
- Volkan, V. D., & Ast, G. (1994). *Spektrum des Narzissmus* [Spectrum of narcissism]. Göttingen: Vandenhoeck & Ruprecht.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.

Footnotes

¹ We thank an anonymous reviewer for suggesting to analyze the internal consistency of the measure in more detail.

Table 1
Descriptive statistics of all variables

	#	α	M	SD	Min	Max
Study 1						
Explicit SE	32	.94	4.69	.90	2.41	6.41
Implicit SE (initials preferences)	2	.63	3.47	2.44	-4.74	7.33
Ambiguous Statements Task	13	.71	-.28	1.97	-5.85	5.38
Study 2						
Explicit SE	22	.93	4.70	.91	2.82	6.55
Implicit SE (IAT, D measure)	n/a ^a	.80	.73	.33	-.13	1.33
Implicit SE (IAT, ms)	n/a ^a	.80	325.66	207.13	-52.22	837.39
Time Spent Reading Feedback (ms)						
total	10	.87	35103.91	10517.83	12402	63984
Trait ratings	7	.85	20664.87	7081.72	8188	42221
Cooperation questions	3	.64	14439.03	4462.46	3804	29434

Note. ^a: See the description of the IAT in the text for details on how internal consistencies were computed, and see Table 2 for a detailed description of the IAT trials. #: Number of items, α : Internal Consistency (Cronbach's Alpha), IAT: Implicit Association Test.

Table 2

Illustration of the Self-esteem Implicit Association Test (SE IAT)

Block	No. of trials	Task	Response key assignment		Sample items
			Left key	Right key	
1	24	Attribute discrimination	Pleasant	Unpleasant	Happiness, friend, harm, agony
2	24	Target discrimination	Me	Not Me	Me, my, it, that
3	96	Initial combined task	Pleasant or Me	Unpleasant or Not Me	Happiness, friend, harm, agony; me, my, it, that
4	24	Reversed target discrimination	Not Me	Me	It, that, me, my
5	96	Reversed combined task	Pleasant or Not Me	Unpleasant or Me	Happiness, friend, harm, agony; it, that, me, my

Note. The original German stimuli can be obtained from the authors.

Figure captions

Figure 1. Predicted values for positive interpretation of ambiguous stimuli as a function of explicit SE and implicit SE. SE = self-esteem, SD = standard deviation.

Figure 2. Acceptance Condition: Predicted values for time spent reading feedback about cooperation as a function of implicit SE and explicit SE. SE = self-esteem, SD = standard deviation.

Figure 3. Rejection Condition: Predicted values for time spent reading feedback about cooperation as a function of implicit SE and explicit SE. SE = self-esteem, SD = standard deviation.

Figure 1

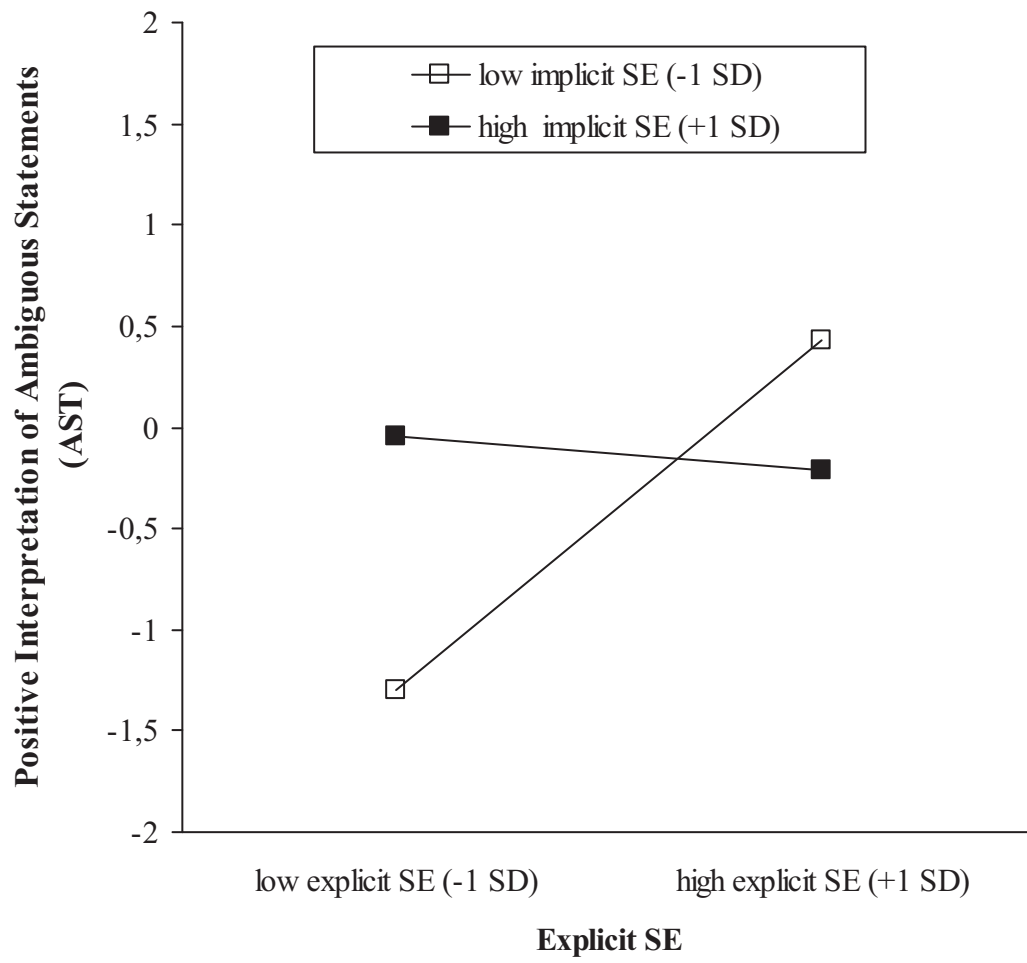


Figure 2

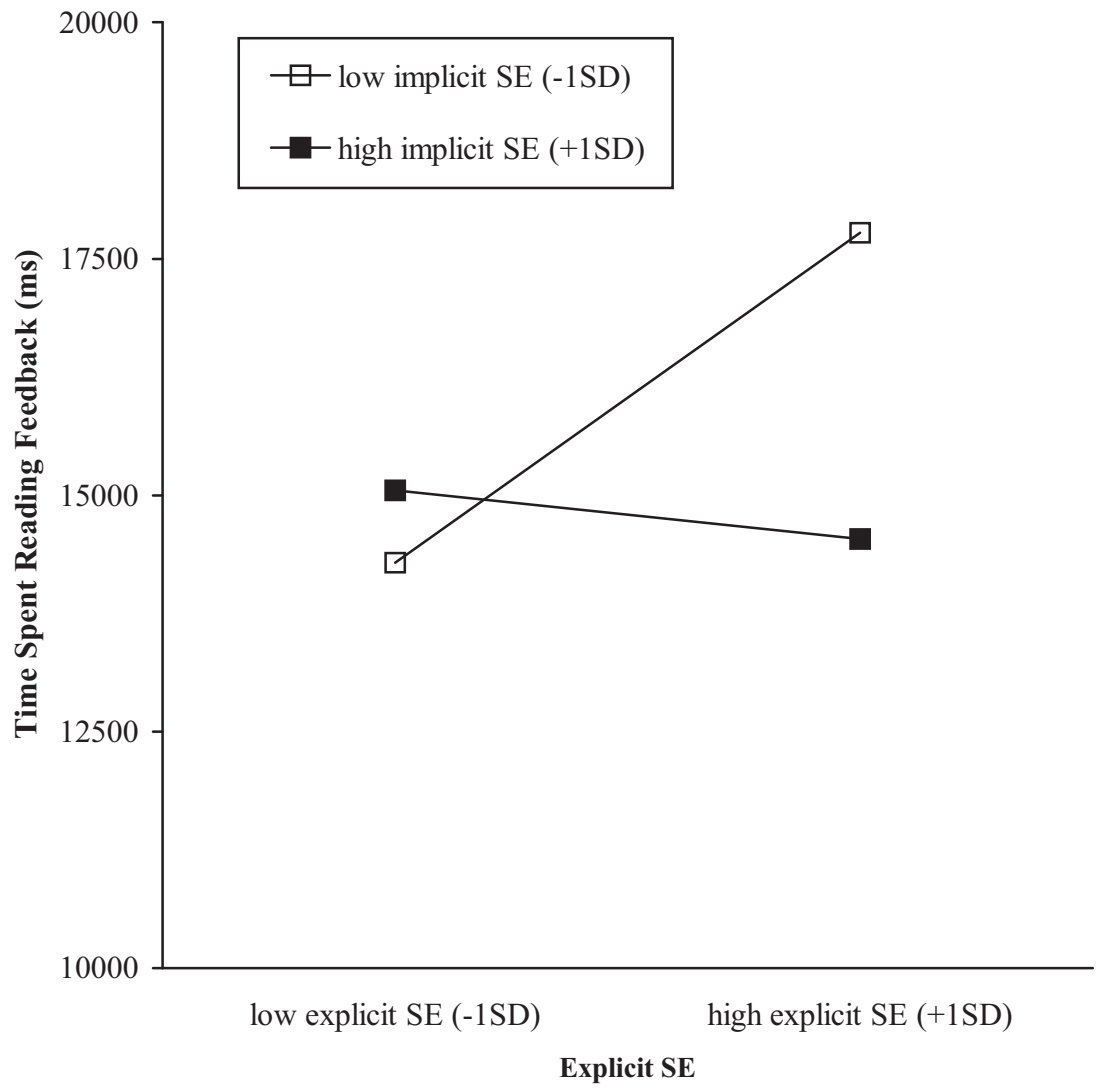
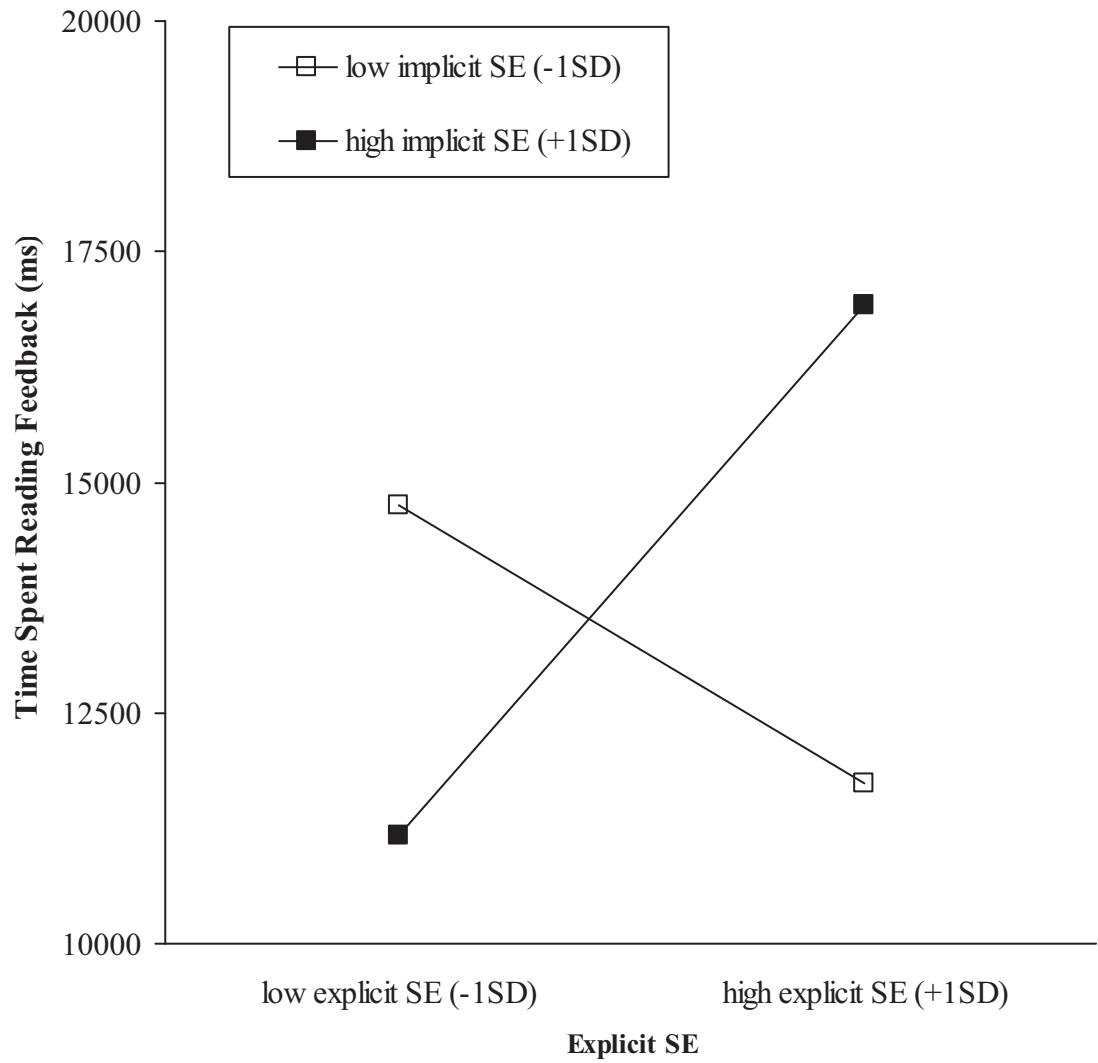


Figure 3



Dipl.-Psych.
Almut Rudolph

Persönliche Daten

Geburtstag	30. Juni 1980
Geburtsort	Frankfurt/Oder
Familienstand	ledig

Anschrift

dienstlich	Technische Universität Chemnitz Institut für Psychologie Persönlichkeitspsychologie und Diagnostik Wilhelm-Raabe-Straße 43 09120 Chemnitz Tel.: 0371-531 35781 E-Mail: almut.rudolph@phil.tu-chemnitz.de
------------	---

privat	Stadlerstr. 7 09126 Chemnitz
--------	---------------------------------

Ausbildung

- seit 10/2004 **Promotionsstudium**
Institut für Psychologie, Technische Universität Chemnitz
- seit 10/2004 **Ausbildung zur Psychologischen Psychotherapeutin (VT)**
Zentrum für Psychotherapie Chemnitz
- 10/2004 **Diplom in Psychologie** (Note 1,4)
Technische Universität Chemnitz und University of
Oklahoma, USA
Diplomarbeit: Implizite und Explizite Selbstwertschätzung in
Beziehung zur Selbstkonzeptstruktur, Erleben und Verhalten
in Belastungssituationen.
- 07/1999 **Abitur** (Note 1,7)
Johann-Gottfried-Seume-Gymnasium Grimma

Berufliche Tätigkeiten

- seit 04/2005 **Psychologische Psychotherapeutin (i.A.)**
Zentrum für Psychotherapie Chemnitz
- seit 10/2004 **Wissenschaftliche Mitarbeiterin**
Professur Persönlichkeitspsychologie und Diagnostik, Institut
für Psychologie, Technische Universität Chemnitz
Lehrveranstaltungen in Testtheorie und Testanwendung
- 04/2005 – 09/2006 **Diplom-Psychologin**
Klinikum Chemnitz, Klinik für Psychiatrie,
Verhaltensmedizin und Psychosomatik
- 07/2003 – 10/2003 **Forschungsassistentin**
Prof. Dr. Carolin Showers, Department of Psychology,
University of Oklahoma, USA
- 10/2000 – 09/2004 **Studentische Hilfskraft**
Persönlichkeitspsychologie und Diagnostik, Allgemeine
Psychologie und Arbeitspsychologie, Klinische Psychologie,
Technische Universität Chemnitz

Workshops und Fortbildungen

- 09/2007 Doktorandenworkshop der Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik (Manfred Schmitt, Hannelore Weber)
Universität Koblenz-Landau
- 03/2006 Freiburger Methodenwoche (Markus Wirtz, Rainer Leonhart)
Methodenzentrum RFV, Freiburg-Bad Säckingen
- 07/2005 Assessing adaptive and maladaptive personality traits across the life span: Conceptual issues and cross-cultural and cross-language perspectives (Filip de Fruyt)
5th Conference on the International Test Commission, Brussels
- 12/2004 Einführung in die Rasch-Analyse mit WINMIRA und WINSTEPS (Maren Böcker, Markus Wirtz)
Methodenzentrum RFV, Freiburg-Bad Säckingen
- 11/2004 Implicit Measures (Aiden Gregg)
Technische Universität Chemnitz

Wissenschaftliche Interessen und Schwerpunkte

Erfassung und Korrelate impliziter Selbstwertschätzung

Selbstwertschätzung als Eigenschafts- und Zustandsvariable

Narzissmus und Selbstkonzept bei Straftätern

Selbstkonzept und Selbstwert in der Psychotherapie

Buch- und Zeitschriftenartikel

Rudolph, A., Schröder-Abé, M., Riketta, M. & Schütz, A. (2009). Easier when done than said! Implicit self-esteem predicts observed or spontaneous behaviour, but not self-reported or controlled behaviour. Manuscript submitted for publication.

Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. P., & Sedikides, C. (2008). Through a Glass, Less Darkly? Reassessing Convergent and Discriminant Validity in Measures of Implicit Self-Esteem. *European Journal of Psychological Assessment*, 24, 273-281.

Rudolph, A., Schröder-Abé, M. & Schütz, A. (2008). Ein Impliziter Assoziationstest zur Erfassung von Selbstwertschätzung. In W. Sarges & D. Scheffer (Hrsg.). *Innovative Ansätze in der Eignungsdiagnostik* (S. 197-206). Göttingen: Hogrefe.

Rudolph, A., Schütz, A. & Schröder-Abé, M. (2007). Selbstkonzept, Selbstwert und Selbstwertregulation. In S. Herpertz, F. Caspar & C. Mundt (Hrsg.). *Störungsorientierte Psychotherapie* (S. 207-226). München: Elsevier Verlag, Urban & Fischer.

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

Schröder-Abé, M., Rudolph, A., Wiesner, A. & Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *International Journal of Psychology*, 42, 174-183.

Rudolph, A., Schröder, M. & Schütz, A. (2006). Ein Impliziter Assoziationstest zur Erfassung von Selbstwertschätzung. In T. M. Ortner, R. T. Proyer & K. D. Kubinger (Hrsg.). *Theorie und Praxis Objektiver Persönlichkeitstests* (S. 153-163). Bern: Verlag Hans Huber.

Vorträge (chronologisch geordnet)

Rudolph, A., Schröder-Abé, M., & Schütz, A. (2008, July). *And the Winner is...! Reassessing Reliability and Validity in Measures of Implicit Self-Esteem*. Paper presented at the 14th European Conference on Personality, Tartu, Estonia.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2008, July). *Negative Effects of Discrepancies between Explicit and Implicit Self-Esteem in Interpersonal Relationships*. Paper presented at the 14th European Conference on Personality, Tartu, Estonia.

Schütz, A., Schröder-Abé, M., Horch, G., & Rudolph, A. (2007, September). *Implizite Selbstwertschätzung und Partnerschaft*. Vortrag auf der 9. Arbeitstagung der Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, Wien.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007, September). *Discrepancies between implicit and explicit self-esteem*. Paper presented at the 11. Tagung der Fachgruppe Sozialpsychologie, Freiburg.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007, July). *Discrepancies between implicit and explicit self-esteem predict poor mental health and problematic emotion regulation*. Paper presented at the Biennial Meeting of the International Society for the Study of Individual Differences, Gießen.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007, May). *Implicit self-esteem: Assessment options and discrepancies with explicit self-esteem*. Paper presented at the 9th European Conference on Psychological Assessment, Thessaloniki, Greece.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007, März). *Diskrepanzen zwischen impliziter und expliziter Selbstwertschätzung: Zusammenhänge mit Defensivität und Gesundheitsindikatoren*. Vortrag auf der 49. Tagung experimentell arbeitender Psychologen (TeaP), Trier.

Rudolph, A., Schröder-Abé, M., & Schütz, A. (2006, September). *Selbstwertschätzung und psychische Gesundheit*. Vortrag im Rahmen des gemeinsamen Kongress der Deutschen Gesellschaft für Medizinische Psychologie (DGMP) und der Deutschen Gesellschaft für Medizinische Soziologie (DGMS), Leipzig.

Rudolph, A., Schröder-Abé, M., & Schütz, A. (2006, September). *Reaktionszeitbasierte Testverfahren in der Persönlichkeitsdiagnostik. Ein Impliziter Assoziationstest zur Erfassung von Selbstwertschätzung*. Vortrag im Rahmen der Tagung Innovative Ansätze in der Eignungsdiagnostik, Hamburg.

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2006, September). *Diskrepanzen zwischen impliziter und expliziter Selbstwertschätzung: Zusammenhänge mit Defensivität und Gesundheitsindikatoren*. Vortrag im Rahmen des 45. Kongress der Deutschen Gesellschaft für Psychologie, Nürnberg.

Schröder-Abé, M., Vater, A., Rudolph, A., & Schütz, A. (2006, September). *Implicit self-esteem in patients with Major Depression and Borderline Personality Disorder*. Paper presented at the 53rd British Psychological Society Social Psychology Section Annual Conference, Birmingham, UK.

Rudolph, A., Schröder, M., & Schütz, A. (2006, July). *An Implicit Association Test to Measure Self-Esteem*. Paper presented at the 5th Conference of the International Test Commission, Brussels.

Schröder, M., Rudolph, A., & Schütz, A. (2006, July). *Health Outcomes of Discrepancies between Implicit and Explicit Self-Esteem*. Paper presented at the 13th European Conference on Personality, Athens, Greece.

Schröder, M., Schütz, A., & Rudolph, A. (2005, June). *Psychometric Properties of German Language Implicit Self-Esteem Measures*. Paper presented at the 5th Workshop on Implicit Representations and Personality, Berlin.

Schröder, M., Schütz, A. & Rudolph, A. (2005, Juni). *Psychometric Properties of German Language Implicit Self-Esteem Measures*. Vortrag auf dem 5. DFG-Workshop "Implicit Representations and Personality", Berlin.

Poster (chronologisch geordnet)

Schröder-Abé, M., Rudolph, A., & Schütz, A. (2009, February). *When less is more: Negative outcomes of discrepancies between explicit and implicit self-esteem*. Poster presented at the 10th Annual Conference of the Society for Personality and Social Psychology, Tampa, Florida, USA.

Rudolph, A., Schütz, A., & Schröder-Abé, M. (2008, July). *Validation of a German Version of the State Self-Esteem Scale*. Poster presented at the XXIX. International Congress of Psychology, Berlin.

Schröder-Abé, M., Rudolph, A. & Schütz, A. (2008, July). *When less is more: Negative outcomes of discrepancies between explicit and implicit self-esteem*. Poster presented at the XXIX. International Congress of Psychology, Berlin.

Rudolph, A., Schütz, A., & Schröder-Abé, M. (2008, July). *Development and Validation of a German-Language Version of the State Self-Esteem Scale*. Poster presented at the 14th European Conference on Personality, Tartu, Estonia.

Rudolph, A., Schröder-Abé, M. & Schütz, A. (2007, September). *Ein IAT und ein ST-IAT zur Erfassung Impliziter Selbstwertschätzung*. Posterpräsentation auf der 9. Arbeitstagung der Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, Wien.

Rudolph, A., Schröder-Abé, M. & Schütz, A. (2007, July). *IAT and EAST as Measures of Implicit Self-Esteem*. Poster presented at the Biennial Meeting of the International Society for the Study of Individual Differences, Gießen.

Rudolph, A., Schröder, M. & Schütz, A. (2005, September). *Analyse einer EAST (Extrinsic Affective Simon Task) zur Erfassung impliziter Selbstwertschätzung*. Posterpräsentation auf der 8. Arbeitstagung der Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, Marburg.

Rudolph, A., Schütz, A., Showers, C. J. & Zeigler-Hill, V. (2005, September). *Sichere und fragile Selbstwertschätzung beim Umgang mit Stress: Beweise für die Dissoziation von spontanem und kontrolliertem Verhalten*. Posterpräsentation auf der 10. Tagung der Fachgruppe Sozialpsychologie, Jena.

Schröder, M., Schütz, A. & Rudolph, A. (2005, September). *Fragile Selbstwertschätzung und Selbstwertbedrohung: Vorhersage von Reaktionen auf Misserfolge im sozialen und Leistungsbereich durch implizite und explizite Selbstwertschätzung*. Posterpräsentation auf der 10. Tagung der Fachgruppe Sozialpsychologie, Jena.

Schütz, A., Schröder, M. & Rudolph, A. (2005, September). *Auf die ,andere' Kategorie kommt es an! Vergleich zweier Versionen des IAT*. Posterpräsentation auf der 8. Arbeitstagung der Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, Marburg.

Rudolph, A., Schütz, A., Showers, C. J. & Zeigler-Hill, V. (2005, Juli). Secure versus fragile self-esteem in coping with stress. Evidence for the dissociation of spontaneous and controlled reactions. Poster presented at the 14th General Meeting of the European Association of Experimental Social Psychology, Würzburg.

Qualifizierungsarbeiten

Rudolph, A. (2004). Implizite und Explizite Selbstwertschätzung in Beziehung zur Selbstkonzeptstruktur, Erleben und Verhalten in Belastungssituationen. Unveröffentlichte Diplomarbeit. TU Chemnitz.

Erklärung

Ich erkläre, die vorliegende Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Hilfsmittel verwendet.

Almut Rudolph

Chemnitz, 10.06.2009