

*Konstruktvalidierung von sozialen Einstellungen  
aus impliziten und expliziten  
Einstellungsmessungen*

Dissertation

zur Erlangung des akademischen Grades

doctor philosophiae (Dr. phil.)

vorgelegt dem Rat der Fakultät für Sozial- und Verhaltenswissenschaften  
der Friedrich-Schiller-Universität Jena

von Dipl.-Psych. Nicole Kämpfe

geboren am 16. 04. 1979 in Mühlhausen/ Thür.

Gutachter:

1. Prof. Dr. Rainer Riemann

2. Prof. Dr. Peter Noack

Tag des Kolloquiums: 01.11.2005

Danke!

Ganz herzlich danke ich Professor Dr. Rainer Riemann für seinen,  
in jeglicher Hinsicht motivierenden Beistand und seine fachliche Unterstützung.

~

Ebenfalls herzlich danken möchte ich Professor Dr. Peter Noack für seine  
freundliche Unterstützung und die hilfreichen Hinweise.

~

Darüber hinaus danke ich meiner lieben Kollegin Frau Dr. Kristin Mitte  
für die fruchtbare und heitere Zusammenarbeit, für ihren praktischen Beistand und nicht  
zuletzt für die seelische und moralische Unterstützung ihrerseits.

~

Nicht vergessen möchte ich all die fleißigen Versuchsleiter und Kodierer, ohne deren  
Unterstützung der empirische Teil vorliegender Arbeit undenkbar gewesen wäre.

~

Mein besonderer Dank gilt Mathias, der immer genug Geduld, Ruhe und Optimismus für  
Zwei hatte und mir dadurch sehr viel Kraft gab.

# GLIEDERUNGSVERZEICHNIS

1	Grundlagen der Einstellungsforschung	1-1
1.1	Begriffsbestimmung	1-1
1.2	Historischer Überblick zur Einstellungsforschung	1-2
1.3	Duale Prozesstheorien der Einstellungen	1-4
1.4	Begriffsklärung explizit - implizit	1-5
1.5	Direkte Messung von Einstellungen – Explizite Einstellung	1-7
1.6	Indirekte Messung von Einstellungen – Implizite Einstellung	1-9
1.6.1	Implicit Association Test (IAT)	1-12
1.6.1.1	Entwicklung des IAT	1-12
1.6.1.2	Aufbau des IAT	1-13
1.6.1.3	Empirische Validitätsnachweise für den IAT	1-14
1.6.1.4	Validitätsbereich des IAT	1-15
1.6.2	Extrinsic Affective Simon Task (EAST)	1-17
1.6.2.1	Entwicklung des EAST	1-17
1.6.2.2	Aufbau des EAST	1-17
1.6.2.3	Empirische Validitätsnachweise und Validitätsbereich des EAST	1-19
1.6.3	Zusammenfassende Betrachtung – Reliabilität und Konvergenz impliziter Maße	1-20
2	Bedeutung impliziter & expliziter Einstellungen	2-24
2.1	Empirische Befundlage	2-24
2.1.1	Konzeptualisierungsansätze für implizite und explizite Einstellungen	2-25
2.1.1.1	Lügendetektorperspektive soziale Einstellungen	2-26
2.1.1.2	Perspektive unabhängiger Bewertungen	2-17
2.1.2	Diskussion der verschiedenen Konzeptualisierungsansätze	2-29
2.2	Moderatoren der Beziehung impliziter & expliziter Einstellungen	2-32
2.2.1	Allgemeine Moderatoren der Implizit-Explizit-Korrespondenz	2-33
2.2.2	Spezifische Moderatoren der Implizit-Explizit-Korrespondenz: Die Motivationen zur Vorurteilstkontrolle	2-34
2.3	Implizite & explizite Motive	2-36
2.4	Das nomologische Netzwerk sozialer Einstellungen	2-37
2.4.1	Nomologisches Netzwerk expliziter Einstellungen	2-37
2.4.1.1	Right-Wing-Authoritarianism	2-38
2.4.1.2	Soziale Dominanzorientierung	2-39
2.4.1.3	Konservatismus	2-39
2.4.1.4	Ambiguitätsintoleranz & Konstrukte kognitiver Komplexität	2-40
2.4.2	Nomologisches Netzwerk impliziter Einstellungen	2-41
3	Fokus vorliegender Arbeit	3-43
3.1	Implizite Motivmessung mit dem Thematischen Apperzeptionstest	3-44
3.2	Die betrachteten Einstellungsobjekte	3-45
3.2.1	Bedeutung der ausgewählten Stereotypen	3-46
3.2.1.1	Übergewichtige Menschen	4-46
3.2.1.2	Homosexuelle Menschen	3-47
3.2.2	Einstellungen zu den ausgewählten Stereotypen	3-48
3.2.2.1	Übergewichtige Menschen	3-48
3.2.2.2	Homosexuelle Menschen	3-51
3.3	Der empirische Teil vorliegender Arbeit	3-52
3.3.1	Aufbau	3-52
3.3.2	Zeichenlegende	3-53

4	Studie I : Auswirkungen einer fixierten Stimulusabfolge auf die Reliabilität und Validität von IAT und EAST	4-55
4.1	Theoretische Grundlagen und Anliegen von Studie I	4-55
4.2	Methode	4-57
4.2.1	Stichprobe	4-57
4.2.2	Material	4-57
4.2.2.1	Implizite Maße	4-57
4.2.2.2	Explizite Maße	4-61
4.2.3	Durchführung	4-62
4.3	Datenbehandlung	4-62
4.4	Ergebnisse	4-63
4.4.1	Explizite Einstellungen	4-63
4.4.2	Implizite Einstellungen	4-65
4.4.2.1	Evaluation der impliziten Maße	4-66
4.4.2.2	Ergebnisse IAT	4-67
4.4.2.3	Ergebnisse EAST	4-70
4.4.2.4	Zusammenhänge zwischen IAT und EAST	4-79
4.4.3	Gesamtbetrachtung von IAT, EAST und Fragebogenmaßen	4-80
4.4.4	Moderation durch Motivation zur Vorurteilsfreiheit	4-82
4.5	Diskussion	4-84
5	Studie II: Entwicklung einer offenen Assoziativmessung zur Erfassung expliziter & impliziter Einstellungsanteile	5-89
5.1	Theoretische Grundlagen und Anliegen von Studie II	5-89
5.2	Entwicklung der offenen Assoziativmessung	5-90
5.2.1	Bildmaterial	5-90
5.2.2	Aufbau der Kodierschemata	5-92
5.3	Methode	5-96
5.3.1	Stichprobe	5-96
5.3.2	Material	5-98
5.3.3	Durchführung	5-99
5.3.4	Kodierung	5-99
5.4	Auswertung	5-100
5.4.1	Allgemeines Vorgehen	5-100
5.4.2	Hypothesen	5-102
5.5	Ergebnisse	5-103
5.5.1	Offene Assoziativmessung	5-103
5.5.1.1	Basisanalyse	5-103
5.5.1.2	Zusammenhänge der kodierten Variablen	5-105
5.5.2	Fragebogen	5-107
5.5.2.1	Basisanalyse	5-107
5.5.2.2	Zusammenhänge der Fragebogenmaße	5-108
5.5.3	Gegenüberstellung der offenen Assoziativmessung mit expliziten Einstellungen	5-109
5.5.3.1	Einstellungsobjekt Übergewicht	5-110
5.5.3.2	Einstellungsobjekt Homosexualität	5-112
5.5.3.3	Befunde für die Gesamtstichprobe	5-115
5.6	Expertenratings	5-117
5.6.1	Expertenrating I	5-117
5.6.2	Expertenrating II	5-119
5.7	Diskussion	5-119

6	Studie III: Konstruktvalidierung von Einstellungsmaßen aus expliziter, impliziter und offener Assoziativmessung	6-123
6.1	Theoretische Grundlagen und Anliegen von Studie III	6-123
6.2	Methode	6-126
6.2.1	Stichprobe	6-126
6.2.2	Material	6-127
6.2.2.1	Implizite Messung	6-127
6.2.2.2	Explizite Messung	6-131
6.2.2.3	Offene Assoziativmessung	6-134
6.2.3	Durchführung	6-135
6.2.4	Auswertung	6-136
6.3	Ergebnisse I: Basisanalysen	6-138
6.3.1	Implizite Messungen	6-138
6.3.1.1	IAT Übergewicht	6-138
6.3.1.2	IAT Homosexualität	6-140
6.3.1.3	EAST Übergewicht	6-142
6.3.1.4	EAST Homosexualität	6-146
6.3.1.5	Mittelwertsvergleiche für IAT und EAST zu T1 und T2	6-151
6.3.1.6	Zusammenhänge der impliziten Einstellungsmaße	6-154
6.3.2	Selbstberichtsmessung	6-156
6.3.2.1	Vorurteils- und Motivationsskalen	6-156
6.3.2.2	Gefühlsthermometer	6-158
6.3.2.3	Prädiktormäße I	6-159
6.3.2.4	Prädiktormäße II	6-160
6.3.2.5	Interne Zusammenhänge der Fragebogenmaße	6-161
6.3.3	Offene Assoziativmessung	6-174
6.3.3.1	Deskriptive Statistiken der Geschichtenvariablen	6-174
6.3.3.2	Zusammenhänge der Geschichtenmaße	6-175
6.3.4	Zusammenhangsanalyse der Einstellungsmaße	6-177
6.3.4.1	Zusammenhänge expliziter und impliziter Einstellungen	6-177
6.3.4.2	Zusammenhänge expliziter, impliziter und assoziativer Einstellungsmaße	6-179
6.3.4.3	Einstellungsobjekt Übergewicht	6-180
6.3.4.4	Einstellungsobjekt Homosexualität	6-181
6.3.4.5	Gesamtbetrachtung	6-183
6.4	Ergebnisse II: Latente Multitrait-Multimethod-Analyse	6-184
6.4.1	Voraussetzungen und Evaluation latenter Strukturgleichungsmodelle	6-184
6.4.2	Entscheidung über die Messmodelle	6-187
6.4.3	Grundlagen der latenten MTMM-Analyse	6-190
6.4.4	CT-C(M-1) Modellierungen	6-191
6.4.4.1	Grundlagen	6-191
6.4.4.2	Ergebnisse	6-192
6.4.4.3	Ergebnisse für Reaktionszeitdaten	6-196
6.4.5	Zwischenresümee	6-197
6.4.6	Modifikation und Respezifizierung MTMM-Modellierung	6-199
6.4.7	Gesamtresümee der MTMM Analyse	6-202
6.5	Ergebnisse III: Explizit-Implizit-Kontinuum der Einstellungen	6-203
6.5.1	Zusammenfassung der Basismodelle der Vorhersage	6-203
6.5.2	Gesamtmodelle der Vorhersage	6-204
6.5.2.1	Vorhersagemodell für direkte Messung, Geschichtenmessung und IAT	6-205
6.5.2.2	Vorhersagemodell für direkte Messung, Geschichtenmessung, IAT und EAST	6-206
6.5.3	Resümee zur Einstellungsprädiktion	6-207

6.6	Ergebnisse IV: Bedeutung individueller & antizipierter Einstellungen	6-207
6.6.1	Einstellungsobjekt Übergewicht	6-208
6.6.2	Einstellungsobjekt Homosexualität	6-209
6.7	Ergebnisse V: Moderatoranalysen Motivation zur Vorurteilsfreiheit	6-210
6.7.1	Moderatoreffekte der Beziehung zwischen impliziten und expliziten Einstellungen	6-211
6.7.1.1	Hierarchische Regression	6-211
6.7.1.2	Varianzanalyse	6-213
6.7.2	Moderatoreffekte der Beziehung zwischen offener Assoziativmessung und expliziten Einstellungen	6-215
6.8	Ergebnisse VI: Das nomologische Netzwerk von Einstellungen	6-217
6.8.1	Bedeutung expliziter Einstellungen	6-217
6.8.1.1	Korrelationsanalysen zum nomologischen Netzwerk expliziter Einstellungen	6-217
6.8.1.2	Regressionsanalysen zur Vorhersage expliziter Einstellungen	6-219
6.8.2	Bedeutung impliziter Einstellungen	6-221
6.8.2.1	Korrelationsanalysen zum nomologischen Netzwerk impliziter Einstellungen	6-221
6.8.2.2	Regressionsanalysen zur Vorhersage impliziter Einstellungen	6-224
6.8.3	Bedeutung der Variablen der offenen Assoziativmessung	6-224
6.8.3.1	Korrelationsanalysen zum nomologischen Netzwerk der Geschichtenvariablen	6-224
6.8.3.2	Regressionsanalysen zur Vorhersage des Vorurteilsgehalts der Geschichten	6-226
6.8.4	Resümee	6-227
6.9	Diskussion	6-228
6.9.1	Voraussetzung für die Überprüfung inhaltlicher Fragestellungen	6-228
6.9.2	Die Fragestellungen von Studie III	6-232
6.9.2.1	Fragestellung I: Konstruktvalidierung von Einstellungen	6-232
6.9.2.2	Fragestellung II: Implizit-Explizit-Dimension der Einstellungen	6-238
6.9.2.3	Fragestellung III: Die Rolle individueller & kultureller Einstellungen	6-241
6.9.2.4	Fragestellung IV: Die Moderatorwirkung von Motivation zur Vorurteilsfreiheit	6-247
6.9.2.5	Fragestellung V: Nomologisches Netzwerk expliziter & impliziter Einstellungen	6-251
6.9.3	Weitere Befunde	6-258
6.9.4	Zusammenfassung	6-259
7	Gesamtdiskussion	7-261
7.1	Konzeptualisierung von Einstellungen	7-263
7.2	Evaluation der impliziten Einstellungsmessungen	7-266
7.2.1	Interne Konsistenz	7-268
7.2.2	Stabilität	7-271
7.2.3	Aspekte der Reliabilitätsberechnung	7-273
7.3	Evaluation der offenen Assoziativmessung	7-274
7.4	Die Rolle der ausgewählten Einstellungsobjekte	7-278
7.5	Zusammenfassung und Ausblick	7-281
8	Literaturverzeichnis	8-282

## ZUSAMMENFASSUNG

Die Unterscheidung zwischen impliziten und expliziten Einstellungen (z.B. Greenwald & Banaji, 1995) und die Entwicklung des Impliziten Assoziationstest (IAT, Greenwald, McGhee & Schwartz, 1998) belebten die angewandte Einstellungsforschung in ungeahntem Ausmaß. Implizite Maße wie der IAT werden als Ausdruck der unreflektierten, unkontrollierbaren, spontanen und automatischen Kognitionsinhalte von Personen angesehen und expliziten Einstellungen, die mittel direkterer Methoden wie Selbstberichtsmessungen erfasst werden, im Hinblick auf den Grad der Verfälschungsanfälligkeit und differentielle Verhaltensvorhersage gegenübergestellt. Der Einsatz des IAT zur Messung automatischer Wertassoziationen konzentriert sich längst nicht allein mehr nur auf Einstellungen, sondern auf viele Bereiche der Allgemeinen-, Differentiellen-, Klinischen- und Sozialpsychologie. Neben dem IAT wurden eine Reihe weiterer indirekter Verfahren zur Messung impliziter Dispositionen entwickelt, doch nach wie vor ist ungeklärt, inwiefern diese Messmethoden gleichsam ein- und dasselbe implizite Konstrukt erfassen und inwieweit implizite Kognitionsinhalte aus einer gemeinsamen Grundgesamtheit mit expliziten Einstellungen stammen, die als reflektierter, elaborierter und kontrollierbarer charakterisiert sind.

Mit vorliegender Arbeit wird eine Untersuchung dieser Konzeptionalisierung von expliziten und impliziten Einstellungen zu zwei sozialen Stereotypen über eine Konstruktvalidierung konvergenter und diskriminanter Validität verschiedener Einstellungsmaßen vorgenommen. Zu diesem Zweck werden direkte und indirekte Messmethoden einbezogen. Selbstberichtskaalen und Gefühlsthermometermessungen dienen der Erfassung expliziter Einstellungen. Zwei Reaktionszeitparadigmen werden zur Messung impliziter Wertassoziationen gegenüber Homosexualität bzw. Übergewicht eingesetzt, und zwar der *Implizite Assoziationstest* und der *Extrinsic Affective Simon Task* (De Houwer, 2003). Darüber hinaus wird eine *offene Assoziativmessung* in Anlehnung an die projektive Methodik des Thematischen Apperzeptionstest (Murray, 1943) zur Messung von reflektierten und unreflektierten Einstellungsinhalten zu Homosexualität und Übergewicht entwickelt. Damit wird der Komplexität von Einstellungen in Form von motivationalen, kognitiven und affektiven Komponenten Rechnung getragen, wobei die Messverfahren verschiedene konzeptuelle Ebenen (direkt vs. indirekt) abbilden. Das Aufgebot dieses komplexen Methodenarsenals zur Einstellungsmessung dient der Validierung der Vorurteilkonzeption über eine latente Multitrait-Multimethod-Analyse und liefert darüber hinaus wichtige Informationen zur Evaluation der nach wie vor sehr kritisch diskutierten impliziten Maße.



Zur Untersuchung der Bedeutung der verschiedenen Einstellungsmessungen werden Variablen des nomologischen Netzwerkes von Einstellungen einbezogen, und zwar persönlichkeitsnahe kognitive Konstrukte wie *Ambiguitätsintoleranz* oder *Need for Cognition*, motivationale Konstrukte wie *Soziale Erwünschtheit* oder die *Motivation zur Vorurteilsfreiheit* und relativ etablierte soziale Einstellungsprädiktoren wie *Right-Wing-Autoritarismus*, *Soziale Dominanzorientierung*, *Konservatismus* und *Modern Racism*. Neben der Gegenüberstellung der nomologischen Netzwerke expliziter und impliziter Einstellungen, ermöglicht speziell die Berücksichtigung von vorurteilsrelevanter Motivation (z.B. Motivation zur Vorurteilsfreiheit) die Untersuchung von Moderatoreffekten in Bezug auf die Beziehung zwischen impliziten und expliziten Einstellungsmaßen. Ein weiterer Fokus liegt auf der Dissoziation von *individuellen* und *kulturellen* expliziten Einstellungen liegt.

Im Rahmen vorliegender Dissertation werden drei empirische Untersuchungen der Autorin. Diese sind dazu angelegt worden, einen wichtigen Beitrag zur Erweiterung des Kenntnisstandes in der differentiellen Einstellungsmessung zu leisten.

**– Theoretische Grundlagen –**

# 1 Grundlagen der Einstellungsforschung

## 1.1 Grundlegende Begriffsbestimmungen

Als grundlegendste Begriffe der Einstellungsforschung sind „Einstellung“, „Vorurteil“ und „Stereotyp“ anzuführen und zugleich voneinander zu unterscheiden, denn diese insbesondere sozialpsychologisch geprägten Begriffe sind sehr eng miteinander verknüpft: *Einstellungen* und *Stereotype* bilden die Grundlage von *Vorurteilen*. Die drei Begriffe bezeichnen individuelle und soziale Phänomene von ungeminderter gesellschaftlicher Relevanz. Die Definition von *Einstellungen* stellt in aller Regel eine Verknüpfung von Einstellungsobjekt mit affektiver, kognitiver und verhaltensrelevanter Wertung dar (Eagly & Chaiken, 1993). Da Einstellungen zentrale Bedeutung in aktueller Arbeit tragen, werden einige ausgewählte Definitionsansätze vorgestellt:

*Einstellungen sind individuelle Besonderheiten in der Bewertung spezifischer Objekte der Wahrnehmung oder Vorstellung (Asendorpf, 2004)*

*Einstellungen sind Assoziationen zwischen Einstellungsobjekten und Bewertungen dieser Objekte (Fazio & Roskos-Ewoldson, 1994)*

*Einstellungen sind die gelernte, relativ stabile Tendenz, auf Menschen, Konzepte und Ereignisse wertend zu reagieren (Zimbardo & Gerrig, 2004)*

*Einstellungen sind im Gedächtnis gespeicherte, überdauernde Bewertungen verschiedener Aspekte der sozialen Umwelt (Judd, Drake & Downing, 1991)*

*Eine Einstellung ist ein generelles und länger anhaltendes positives oder negatives Gefühl gegenüber einer Person, Objekt oder Thema. (Petty & Cacioppo, 1981)*

*Einstellungen sind die Gesamtbewertung eines psychologischen Objektes auf Attributdimensionen wie gut-schlecht, schädlich-nützlich, angenehm-unangenehm und sympathisch-unsympathisch (Petty, Wegener and Fabrigar, 1997)*

Stereotype beziehen sich dagegen spezielle auf soziale Gruppen. Unter einem Stereotyp wird die Überzeugung von Menschen verstanden, dass sich Mitglieder sozialer Gruppen durch bestimmten Eigenschaften oder Charakteristika auszeichnen. Stereotype sind demnach generalisierende Vereinfachungen der sozialen Komplexität, die sich durch einen hohen Wiedererkennungswert auszeichnen (Baron and Byrne, 2003). Stereotype bilden die wertneutrale kognitive Komponente eines Vorurteils (Bierhoff, 1998). Vorurteile selbst werden als „negative Stereotype, die sich auf Gruppen von Menschen beziehen“ verstanden (Zimbardo, 1995) bzw. als „eine auf falsche od. unvollständige Information gegründete, generalisierte feindselige od. negative Einstellung einer spezifischen Gruppe gegenüber“ (Aronson, Wilson & Akert, 2004).

## 1.2 Historischer Überblick zur Einstellungsforschung

Das Konzept der Einstellungen ist eng verbunden mit den Konzepten Affekt und Valenz und somit mit den grundlegendsten psychologischen Dimensionen von Gut vs. Schlecht, Positiv vs. Negativ, Annäherung vs. Vermeidung (Musch & Klauer, 2003). Die Autoren betonen die Bedeutung des Forschungsfeldes (S. 1):

*„Accordingly, there has been a dramatic increase in interest in evaluative processes in the late 20th century. Research on the nature of evaluative processes is now one of the most rapidly growing endeavours of psychology and provides a unifying focus for researchers working in a variety of disciplines such as social, cognitive, and personality psychology.“*

Innerhalb der Sozialwissenschaften ist die Untersuchung von Einstellungen, evaluativen Bewertungen und Vorurteilen als einer der traditionsreichsten Forschungsbereiche zu charakterisieren. Basierend auf Duckitt (1992) unterscheidet Dovidio (2001) drei Wellen der Vorurteilsforschung. In der ersten Welle (1920 – 1950), deren frühe Väter zweifellos Reich und Fromm mit eher psychodynamischer Perspektive sind, wurden Vorurteile in aller Regel eher psychopathologisch betrachtet. In diesem Rahmen wurden die Messung und Beschreibung von Vorurteilen sowie die Untersuchung der Quellen und Veränderungsbedingungen von Vorurteilen besonders betont. Allmählich begann die Verortung vorurteilsvoller Einstellungen weniger in krankhaften sozialen oder individuellen Strukturen, denn vielmehr in der *normalen* Persönlichkeit. Aus dieser zweiten Welle stammen Ansätze, die sich mit den Bedingungen wahrer Vorurteilsfreiheit beschäftigen, da Vorurteile als weithin existentes Gesellschaftsphänomen anerkannt wurden.

Die Übergänge der von Dovidio deklarierten Phasen sind fließend. So ist auch schwer auszumachen, wann schließlich die dritte Welle der multidimensionalen Betrachtung von Vorurteilen begann. Aus der in den 80er Jahren zu beobachtenden Betonung kognitiver Ansätze zur Untersuchung von Einstellungen, Stereotypen, Vorurteilen und *automatischer* evaluativer Bewertungen, resultierte schließlich eine höchst brisante konzeptuelle Weiterentwicklung des Einstellungsbegriffes. Diese Weiterentwicklung betrifft die Differenzierung zwischen traditionellen „*expliziten*“ und „*impliziten*“ Kognitionen bzw. Einstellungen (zur Begriffsklärung siehe Abschnitt 1.4). In das Zeitfenster einer eher kognitiv ausgerichteten Perspektive der Einstellungs-, Stereotypen- und Vorurteilsforschung ist dementsprechend auch die Entwicklung konzeptuell neuer Messverfahren für Einstellungen einzuordnen, woraus sich wiederum neue Fragen erhoben, die beständig neuen Nährboden für ein nicht nachlassendes Forschungsinteresse bieten.

Mit Arbeiten, die einer experimentellen Untersuchung automatischer Prozesse in der Informationsverarbeitung gewidmet waren, erfuhr die Tradition der bislang vornehmlich *direkten* Einstellungsforschung eine inhaltliche Bereicherung und zugleich eine neue theoretische Ausrichtung (z.B. Bargh, 1982; Fazio, Sanbonmatsu, Powell & Kardes, 1986; Greenwald & Banaji, 1995). Das 1986 von Fazio

et al. eingeführte Paradigma des *Affektiven Primings* kann als früher Impuls eines etwa zehn Jahre später erst schlagartig einsetzenden Forschungsbooms zu impliziten Einstellungen und Vorurteilen betrachtet werden. Die Veröffentlichung einer Arbeit von Greenwald & Banaji (1995) zur Unterscheidung expliziter und impliziter Kognitionen und die Entwicklung des *Impliziten Assoziationstest* (IAT) von Greenwald, McGhee & Schwartz (1998, siehe Kapitel 1.6.1) zogen eine Welle an empirischen Untersuchungen nach sich, die psychologische Konstrukte wie Einstellungen nicht mehr nur auf traditionellem *direktem*, sondern nunmehr auch auf *indirektem* Weg zu erfassen suchten (siehe Abschnitt 1.5 und 1.6). Insofern ist der IAT gewissermaßen als „Pate“ für eine neue, facettenreiche Herangehensweise an die Erfassung von Einstellungen zu bezeichnen. Die Publikation des IAT liegt nunmehr sieben Jahre zurück. Die Auseinandersetzung mit Verfahren ist ungebrochen vielfältig, jedoch auch kritisch und führte zu Fort- und Weiterentwicklungen von Verfahren, mit denen implizite Dispositionen über Reaktionszeit- und Fehlerratenparadigmen erfasst werden können. Inzwischen existiert neben dem Affektiven Priming und dem IAT eine Bandbreite an Paradigmen zur Messung impliziter Kognitionen, zu denen z.B. der *Extrinsic Affective Simon Task* (EAST, De Houwer 2003), der *Go/No-Go-Task* (GNAT, Nosek & Banaji, 2001) oder der *Single-Target IAT* (STIAT, Wigboldus, van Knippenberg & Holland, 2001) zählen.

Diese indirekte Messmethodik wurde anfangs vornehmlich zur Erfassung von impliziten Einstellungen und Vorurteilen eingesetzt, jedoch relativ schnell ausgeweitet auf die Erfassung von *Persönlichkeit und Selbstkonzept* (z.B. Schmukle & Egloff, 2004; Teige, Schnabel, Banse & Asendorpf, 2004), *psychopathologische Aspekte* wie Selbstwert und Angst (z.B. de Jong, 2002; Egloff & Schmukle, 2002; Greenwald & Farnham, 2000; Greenwald, Banaji & Rudman, 2002; vgl. De Houwer, 2002) oder auch *Konsumentenverhalten* (z.B. Karpinski & Hilton, 2001; Maison, Greenwald, Bruin, 2001). Eher ungewöhnlich an der gesamten Erfolgsgeschichte impliziter Einstellungen ist das Verhältnis von Theorie und Empirie. Entgegen einer entsprechenden Parallelentwicklung handelt es sich hier um einen insbesondere *empirisch* vorangetriebenen Forschungsbereich, dessen theoretischen Unterboden es immer noch an Substanz mangelt (Fazio & Olson, 2003). Der in kürzester Zeit beachtlich herangewachsene Bestand an Publikationen zu impliziten Einstellungsmaßen umfasst hauptsächlich experimentelle Laborstudien, die in ihrer Gesamtheit viele neue Fragen zur theoretischen Konzeptionalisierung von expliziten und impliziten Einstellungen aufwerfen. Jedoch sind integrative bzw. konzeptionelle Arbeiten bis auf wenige Ausnahmen (z.B. De Houwer, 2003; Fazio & Olson, 2003; Rudman, 2004) kaum verfügbar. Nichtsdestotrotz existieren Modelle zur Funktionalität und Verhaltensrelevanz von Einstellungskomponenten, die eine empirische Überprüfung erlauben und somit eine theoretische Grundsteinlegung der Einstellungskonzeption ermöglichen. Dazu gehören duale Prozesstheorien wie das MODE-Modell von Fazio (1990, siehe auch Fazio & Towles-Schwen, 1999) oder die Unterscheidung zwischen impulsiven und reflexiven Komponenten sozialen Verhaltens von Strack & Deutsch (2004).

### 1.3 Duale Prozesstheorien der Einstellungen

Vor etwa fünfzehn Jahren stellte Fazio (1990) das so genannte *MODE*-Modell der Verhaltenssteuerung vor, das als Grundlage der modernen Einstellungsforschung betrachtet werden kann. Unter "MODE" versteht der Autor "*Motivation and Opportunity as DEterminants of Processing*" ("...of whether the attitude-to-behavior process is primarily spontaneous or deliberative in nature"; z.B. Fazio & Olson, 2004). *Motivation* und *Möglichkeit* sind hier die entscheidenden Determinanten verhaltenssteuernder Einstellungen, wobei zwischen automatischen (impliziten) und elaborierten (expliziten) Einstellungsinhalten unterschieden wird. Unter der Bedingung ausreichender Motivation und/oder Kapazität zur Informationsverarbeitung postuliert das *MODE*-Modell das *sorgfältige Überdenken* der Handlung inklusive der Folgen, die Ausbildung einer Verhaltensintention und – wenn keine externen Störeinflüsse die Handlung verhindern – das entsprechende Verhalten. Bei unzureichender Motiviertheit und/oder stark dezimierten kognitiven Ressourcen, z.B. aufgrund von Zeitdruck, Überlastung oder Unsicherheit, werden einfach zugängliche, relevante Einstellungen *automatisch* aktiviert. Je nach subjektiver Wertigkeit des Objektes erfolgt eine selektive Aufmerksamkeitsfokussierung und entsprechendes Annäherungs- bzw. Vermeidungsverhalten.

Das *MODE*-Modell beschreibt Verhalten nunmehr nicht prozessrein, sondern als Ergebnis spontaner *und* überlegter Komponenten, wobei Umweltfaktoren wie soziale Normen, Regeln und Vorschriften beeinflussen die Verhaltenstendenzen beeinflussen. Der *Grad* der Automatizität wird als Funktion von Bewusstheit, Intentionalität, Effizienz und Kontrollierbarkeit betrachtet (Bargh, 1994; Wegner & Bargh, 1998). Fazio und Kollegen sehen die postulierte Interaktion automatisch aktivierter Assoziationen und bewusster Motivation insbesondere bei Rassenstereotypen und Vorurteilen gegeben. Empirische Bestätigung für die Annahmen des im *MODE*-Modells liefern Sanbonmtsu & Fazio (1990) sowie Schuette & Fazio (1995).

Eine ähnliche Konzeption zur Einstellungs-Verhaltens-Korrespondenz stellt das *Reflective-Impulsive Model* von Strack und Deutsch (2004) dar, das eine Trennung zwischen interagierenden *reflexiven* und *impulsiven* Determinanten sozialen Verhaltens vornimmt. Die Autoren nehmen an, dass diese auf verschiedenen Bewusstheitsebenen operierenden Systeme maßgeblich menschliches Verhalten steuern. Wissens- und wertebezogenes Verhalten wird dabei durch das postulierte „reflective system“ bedingt, während von dem „impulsive system“ angenommen wird, Verhalten über Assoziationen und Motivation zu steuern. In Kongruenz zum *MODE*-Modell wird Verhalten als Funktion der Informationsverarbeitung, die auf verschiedenen Ebenen des Bewusstseins abläuft, betrachtet.

An dieser Stelle wird deutlich, dass Einstellungen nicht als unidimensionales Konstrukt zu konzipieren sind, sondern komplexe, auf unterschiedlichen Ebenen der Informationsverarbeitung operierende Dispositionen darstellen. Für die Vorhersage von Verhalten, das als Funktion bewusster und unbe-

wusster, reflektierter und spontaner Komponenten von Einstellung und Motivation zu betrachten ist, kommt der adäquate *Messung* dieser Komponenten eine große Bedeutung zu. Die moderne Einstellungsforschung ist entsprechend geprägt vom Bestreben, valide Messinstrumente zu entwickeln, um Kognitionen zu erfassen, die sowohl expliziter, als auch impliziter Natur sind und eine unterschiedliche Bedeutung für Verhalten, wie z.B. Diskriminierung sozialer Minderheiten, haben. Grundsätzlich mangelt es jedoch nach wie vor an einer soliden konzeptionellen Einbettung von expliziten und impliziten Einstellungen.

## 1.4 Begriffsklärung explizit - implizit

Klassische Theorien zur Struktur von Einstellungen gehen in der Regel von einer Dreikomponententeilung aus (z.B. Eagly & Chaiken, 1993; Rosenberg, Hovland & McGuire, 1960). Fazio & Olson (2003; 139) kommentieren:

*“In this view, the attitude is an unobservable psychological construct which can manifest itself in relevant beliefs, feelings, and behavioral components.”*

Es bleibt noch zu klären, inwiefern *modernere* Einstellungskonzeptionen, wie z.B. die Unterscheidung zwischen expliziten und impliziten Einstellungen, in den klassischen Theorien zu verorten sind. Bezeichnender Weise erfolgte die Verwendung der Begriffe „explizit“ und „implizit“ in der wissenschaftlichen Fachsprache lange Zeit relativ unreflektiert; sie dient auch nach wie vor insbesondere zur Deklaration der Resultat indirekter Einstellungsmessung, z.B. gemäß IAT, und deren Trennung von den traditionellen Selbstberichtsmessungen. Hieraus lassen jedoch noch keine hinreichenden Definitionskriterien ableiten, weshalb vor allem in jüngerer Zeit einige Vorschläge zur Definition und Konzeptionalisierung von impliziten und expliziten Einstellungen unterbreitet wurden (z.B. De Houwer, 2003b; Fazio & Olson, 2003). Die innerhalb der vorliegenden Arbeit wohl am häufigsten verwendeten Begriffe der expliziten und impliziten Einstellungen gehen auf eine Arbeit von Greenwald & Banaji (1995) zurück. Mit der Einführung der Bezeichnung „*implicit cognition*“ definieren die Autoren gleichsam eine Einstellungskomponente, die sich maßgeblich von dem unterscheidet, was bis dato als Einstellung aufgefasst wurde. Ihre im folgenden viel zitierte Arbeit lieferte zugleich einen fruchtbaren Anstoß für die allgemeine Einstellungsforschung. Unter einer impliziten Kognition verstehen die Autoren allgemein (S. 5):

*„...the introspectively unidentified (or inaccurately identified) trace of past experience that mediates reaction.“*

Erst allmählich wurden der Implizit-Explizit-Konzeption einige neue theoretische Bausteine beigelegt, während die empirische Befundlage rasant und beständig erweitert wurde. Ungeachtet der Tatsache

also, dass der Begriff „implizite Einstellung“ dem Leser seitdem erschienener psychologischer Publikationen an vielen Stellen begegnet, fehlt eine fundierte Klärung dessen, was die Attribute „explizit“ und „implizit“ bedeuten und wie sie eindeutig voneinander zu unterscheiden sind. Daher beginnen einige Autoren nun vermehrt, sich mit der Definition der Begriffe und ihrer theoretischen Trennbarkeit eingehender zu befassen, um die Lücke zwischen empirischer Untersuchung und theoretischer Fundierung zu minimieren (vgl. Olson & Fazio, 2003; De Houwer, 2003b, in press). Dabei treten bereits erste Schwierigkeiten bei einer genaueren Definition des Begriffes „implizit“ auf. Zwischen expliziten und impliziten Kognitionen bzw. Maßen wird nunmehr anhand folgender Annahmen getrennt:

	<b>Explizite Kognitionen /Maße</b>	<b>Implizite Kognitionen /Maße</b>
Antwortformat:	<i>verbalisierbar</i>	<i>nicht verbalisierbar</i>
Einsicht in die Messintention:	<i>vorhanden</i>	<i>nicht vorhanden</i>
Bewusstheitsgrad des gemessenen Konstruktes:	<i>eher bewusst</i>	<i>eher unbewusst</i>
Grad der willentlichen Steuerung:	<i>intendiert</i>	<i>unintendiert</i>
Verfälschungstendenzen:	<i>verfälschungsanfällig</i>	<i>unbeeinflusst</i>
kognitive Ressourcen:	<i>hoher Aufwand</i>	<i>geringer Aufwand</i>

**Abbildung 1-1**, Gegenüberstellung der begrifflichen Implikationen von „explizit“ und „implizit“ (vgl. Bargh, 1989, 1994)

Im Gegensatz zu einstellungsbezogenen Selbstberichtsmessungen (siehe unten) zeichnen sich implizite Einstellungsmaße also durch minimale bis fehlende Bewusstheit, eingeschränkten bewussten Zugriff, fehlende willentliche Kontrollierbarkeit und ergo auch stark reduzierte Verfälschbarkeit der Maße aus. Der Begriff „implizit“ kennzeichnet automatisierte Kognitionen, die kaum kognitive Ressourcen benötigen, spontanen Charakter tragen und nicht steuerbar sind. Inhaltlich ist der Begriff „implizit“ vergleichbar mit dem Konzept „automatischer“ Prozesse, die in der Hauptsache folgende Merkmale einschließen: Sie sind unbeabsichtigt, zielunabhängig, werden unwillkürlich und unbemerkt in Gegenwart eines Schlüsselreizes aktiviert, verbrauchen keine kognitive Kapazität, sind autonom, operieren außerhalb bewusster Kontrolle. Insofern teilen implizite Maße und automatische Effekte oder Prozesse eine Reihe an Eigenschaften, wobei jedoch „implizit“ hauptsächlich „unbewusst“ meint, während „automatisch“ sehr viel mehr Eigenschaften mit einschließt (vgl. De Houwer, in press). Diese Attribute werden Maßen, die aus indirekten Messungen wie z.B. dem Affektiven Priming, dem IAT oder dem EAST resultieren, mit der Bezeichnung „implizit“ gleichermaßen zugeschrieben, obwohl die notwendigen Bedingungen nicht unbedingt gegeben sein müssen. Während beispielsweise durch *subliminale* Primedarbietung beim Affektiven Priming nachweislich keine Einsicht in die Messung gegeben ist, berichten viele Probanden, die den IAT durchgeführt haben, zumindest den Inhalt der Messung, nicht



selten auch die Messintention und mehr noch sogar den eigenen Reaktionsvorteil bei kongruenter Stimuluskombination zu erkennen (Monteith, Voils & Ashburn-Nardo, 2001, siehe unten). De Houwer (in press) plädiert deshalb für einen konditionalen Ansatz des Terminus „implizit“, der erfordern würde, die sich mit dem Begriff verbindenden Implikationen für die jeweils eingesetzte indirekte Messmethodik tatsächlich nachzuprüfen. De Houwer rechtfertigt seine Forderung wie folgt (S. 4):

*“Despite the fact that many measures have been given the label „implicit measure“, surprisingly little research has been done to verify whether these measures indeed have the functional properties of implicit measures.”*

Entsprechende Untersuchungen, die eine Validierung gebräuchlicher indirekter Verfahren ermöglichen und eine Einschätzung, *inwieweit* diese tatsächlich implizite Maße liefern, fehlen noch. Anstelle einer Art Absolutheitsanspruch an ein bipolares Implizit-Explizit-Konzept, wäre es hilfreicher, Aussagen darüber treffen zu können, in welchen Aspekten ein Messergebnis wahrhaft implizit ist und in welcher Hinsicht womöglich nicht. Fazio & Olson (2003) trennen strikt zwischen Eigenschaften der Messung und Eigenschaften des Maßes. Sie raten allgemein dazu, direkte bzw. indirekte Messverfahren als „explizit“ bzw. „implizit“ zu bezeichnen, da es mangels empirischer Prüfung nicht zulässig ist, Annahmen auf die aus den Messungen resultierenden *Maße* zu übertragen und diese mit „explizit“ bzw. „implizit“ auszuweisen. Im Verlauf der theoretischen Grundlagenbetrachtung aktueller Arbeit wird deutlich, weshalb eine Konstruktvalidierung expliziter und impliziter Einstellungen von großer Bedeutung für die Einstellungsforschung ist.

## 1.5 Direkte Messung von Einstellungen – Explizite Einstellung

Traditionell bedient sich die Vorurteilsforschung hauptsächlich Selbstberichtskalen in Fragebogenerhebungen. Fragebogenskalen sind durch ein geschlossenes Antwortformat charakterisiert. Für die Validität eines Fragebogenitems stellt dessen sinngemäße Verständlichkeit eine Grundvoraussetzung dar. Notwendigerweise muss also das Thema einer Selbstberichtsmessung erkennbar sein, nicht selten erschließt sich daraus auch die Messintention. Aufgrund dieses explizit verbalen Charakters von Fragebogenskalen erfordert diese Methode der Messung eine gelingende Introspektion sowie Verbalisierung, also die *bewusste* Steuerung seitens des Probanden. Diese Voraussetzungen sind in vielen Fällen nicht oder nicht vollständig gegeben, was sowohl auf die Befragungsinhalte (z.B. Erinnerungen an die Erfahrungen im Elternhaus), als auch auf interindividuell variierende Faktoren wie kognitive Kapazität oder verbale Intelligenz zurückgeführt werden kann. Aufgrund der Abhängigkeit von willentlicher Steuerbarkeit muss die Validität von Fragebogenskalen – nicht nur zur Einstellungsmessung – als eingeschränkt bewertet werden. Darüber hinaus resultiert aus der Kontrollierbarkeit

von Selbstberichtmaßen jedoch eine weitere Problematik: Insbesondere bei der Befragung zu brisanten, sozial sensitiven und/ oder politisierten Einstellungen bietet die Einsicht in die Messintention gemeinsam mit der Kontrollierbarkeit der eigenen Antworten nahezu uneingeschränkte Möglichkeiten zur (sozial erwünschten) Selbstdarstellung (*Impression Management*; Paulhus & Reid, 1991; Tedeschi, Schlenker & Bonoma, 1971; Tetlock, 1981; vgl. Rudman, Ashmore & Gary, 2001).

Dieser Aspekt von Selbstberichtmessung ist seit langem bekannt und wird nicht erst in den jüngeren Generationen von Einstellungsforschern kritisiert. Zur Messung solcher Verfälschungstendenzen wie sozialer Erwünschtheit sind verschiedene, so genannte Kontrollskalen entwickelt worden (z.B. Edwards, 1957; Crowne & Marlowe, 1960; Stöber, 1999). Eine wirkliche Kontrolle ermöglichen diese Messungen dabei jedoch nicht; es konnte nämlich gezeigt werden, dass nicht nur verschiedene Methoden zur Vermeidung von Verfälschungen im Selbstbericht als unbefriedigend einzustufen sind, sondern sich außerdem eine statistische Kontrolle von sozialer Erwünschtheit (Partialkorrelationen) zumeist auf Kosten der Validität der gemessenen Konstrukte auswirkt (Banse & Gawronski, 2003). Soziale Erwünschtheit spiegelt sich natürlich nicht nur in einer Einstellungsmessung, sondern auch in realem, aber *kontrollierbarem* sozialem Verhalten wider, weshalb sozial erwünschte Selbstberichtsmessungen durchaus eine Vorhersagekraft für Verhalten besitzen. Da aber Verhalten in seiner Komplexität eine Funktion *verschiedener* Faktoren ist (vgl. Abschnitt 1.3) und deshalb durch explizite Einstellungsmaße immer nur anteilig vorhergesagt werden könnte und weil dessen ungeachtet seit je her nach einem Zugang zu den „wahren Einstellungen“ von Personen gesucht wird, sind Fragebogenskalen zur Einstellungsmessung nicht hinreichend befriedigend.

Davon abgesehen ist nicht prinzipiell davon auszugehen, dass Einstellungen vollständig dem Bewusstsein zugänglich und verbalisierbar sind, um Verhalten steuern zu können (z.B. Greenwald und Kollegen, Devine und Kollegen). Einstellungen oder einfache Bewertungsassoziationen operieren sowohl auf bewussten wie auch auf unbewussten Verarbeitungsebenen (z.B. Strack & Deutsch, 2004, Abschnitt 1.3). Hieraus resultieren unterschiedliche Konsequenzen nicht nur für die Messung von Einstellungen, sondern auch für beobachtbares Verhalten, wie Greenwald & Banaji (1995) betonen. Die Autoren bringen spontane, automatische Reaktionen (z.B. nonverbales Verhalten) mit introspektiv schwer- oder unzugänglichen Kognitionsinhalten in Verbindung. Mit Selbstberichtsmessungen werden jedoch in der Hauptsache bewusste und verbalisierbare Einstellungsanteile erfasst, da unbewusste Kognitionen unbekannt und unzugänglich sind und in Fragebogenskalen keinen Niederschlag finden können. Dementsprechend muss der Gültigkeitsbereich von Fragebogenmessung auf die Erfassung expliziter Einstellungsinhalte begrenzt sein. Fragebogenskalen sind zwar in Anwendung und Auswertung in aller Regel ökonomisch und nach den Kriterien der Klassischen Testtheorie mindestens zufrieden stellend reliabel, diese Messmethode erscheint jedoch wie dargestellt nicht ausreichend, komplexe Einstellungsanteile mit dem Ziel der Verhaltensvorhersage zu erfassen.

Modernere Ansätze der Einstellungs- und Vorurteilsforschung berücksichtigen die Befunde gegenläufiger Vorurteilslevel gemäß Selbstberichten und Verhaltensmaßen (vgl. Crosby, Bromley & Saxe, 1980) und vollziehen eine Trennung zwischen „alten“ und „modernen“ Vorurteilen (z.B. Katz & Hass, 1988; McConahay, 1986). Diesen Ansätzen liegt die Ansicht zugrunde, dass alle Menschen Vorurteile haben, die sie jedoch nicht zum Ausdruck bringen *wollen*. Mit Banaji & Greenwald (1995) ist diese Perspektive um die Komponente des Nicht-Ausdrücken-*Könnens* erweitert worden. Nunmehr wird also für die Diskrepanz zwischen unterschiedlich durchschaubaren und kontrollierbaren Einstellungsmaßen nicht allein eine motivierte Unterschlagung negativer Einstellungen verantwortlich gemacht, sondern zudem ein Mangel an Einsicht in die betreffende Einstellung.

Der Informationsgehalt direkter Einstellungsmessungen über Fragebogenskalen ist somit weniger als obsolet, denn vielmehr als *begrenzt* zu bezeichnen. Selbstberichte können lediglich die Resultate gelungener Introspektion und in irgendeiner Form motivierter Verbalisierung widerspiegeln. Es stellt sich also die Frage nach dem konkreten Gültigkeitsbereich von Selbstberichten zur Einstellungsmessung, wobei zu berücksichtigen bleibt, dass es Beispiele für Selbstberichte gibt, die über die Abfrage von konkreten Verhaltensindikatoren Rückschlüsse auf Einstellungen oder Persönlichkeitseigenschaften erlauben (z.B. das Minnesota Multidimensional Personality Inventory, MMPI). Hierbei sollte der Einblick, den eine Person in die zu messenden Konstrukte hat, sogar möglichst gering sein. Schwachpunkte weisen derartige Selbstberichtmaße insofern aber nach wie vor auf, als dass beispielsweise Mikroverhalten, welches ebenfalls relevant für die Messung von Einstellungen oder Persönlichkeitseigenschaften sein kann, schwer erkennbar und verbalisierbar bleibt.

Gemäß den dualen Prozesstheorien wie dem MODE-Modell kann eine präzise Vorhersage von Verhalten bzw. die Gültigkeit von Einstellungsmaßen für ein breiteres Verhaltensspektrum nur über die Einbeziehung bewusster und unbewusster Kognitionsinhalte gelingen.

## **1.6 Indirekte Messung von Einstellungen – Implizite Einstellung**

In Reaktion auf die berechtigte Kritik an Selbstberichtsmessungen fiel bereits Mitte der Sechziger Jahre großes Augenmerk und Interesse auf die Entwicklung undurchsichtigerer, nicht-reaktiver Messmethoden (vgl. Webb, 2000). Innerhalb der Einstellungsforschung war diese methodische Orientierung vorerst jedoch nicht viel mehr als ein Liebäugeln. Ein breiter Einsatz indirekter Verfahren ist bis in die Mitte der Achtziger Jahre nicht zu verzeichnen. Erst allmählich breitete sich ein Bewusstsein für die spontane, automatisierte und unbewusste Informationsverarbeitung aus (z.B. Jacoby, Lindsay & Toth, 1987). Paradigmen indirekter Einstellungsmessung greifen insbesondere die im vorangegangenen Abschnitt angeführten Begrenztheiten der traditionellen direkten Selbstberichtsmessung auf. Indirekte

Messungen zeichnen sich dementsprechend durch minimale Durchschaubarkeit und Kontrollierbarkeit aus und zielen auf die Erfassung hauptsächlich unreflektierter, unverfälschter Kognitionsinhalte ab. Dies ermöglicht die mehr oder weniger verhinderte Einsicht in die Messintention sowie das Antwortformat, welches in aller Regel spontanen, nonverbalen Charakter trägt. Gemäß den Modellen von Fazio (1986, 1999) und Strack & Deutsch (2004) sollten Messungen, die keine Gelegenheit zur Reflexion oder Kontrolle über die Antworten gewähren, eine Erfassung automatischer Assoziationen und Kognitionen ermöglichen. Die Ergebnisse solcher Messungen werden als „implizite Einstellungen“ bezeichnet.

Es gibt mittlerweile eine Bandbreite verschiedener indirekter Messverfahren zur Erfassung impliziter Einstellungen. Einige dieser indirekten Messmethoden basieren auf der Auswertung einfacher Verhaltensindikatoren wie Reaktionszeiten, Fehlerraten und Gedächtnisleistungen. Dazu gehören z.B. der IAT, der EAST oder das Affektive Priming. Hierbei werden im Gedächtnis repräsentierte Inhalte nicht über verbale Abfrage im Selbstbericht, sondern mittels computergestützter Experimentalparadigmen erfasst. Es gibt zudem eine Reihe anderer indirekter Verfahren zur Einstellungsmessung, die zum Teil bereits schon lange publiziert sind, jedoch wenig verbreitet, kaum standardisiert und nicht immer eindeutig zu interpretieren. Dazu können physiologische Marker, wie die Hautleitfähigkeit, die Herzrate oder die Pupillengröße (Hess, 1965), aber auch projektive Verfahren, wie z.B. die Vervollständigung von Wortfragmenten (z.B. Son Hing, Li & Zanna, 2002) gezählt werden.

In dieser Arbeit werden jedoch ausschließlich *gerichtet zu interpretierende* implizite Einstellungsmaße auf der Basis von Reaktionszeiten und Fehlerraten betrachtet, von denen angenommen wird, auf kognitiven Assoziationsstärken zu beruhen. Entscheidend ist hierbei die Stärke der Assoziation zwischen Einstellungsobjekten und Valenzkonzepten wie *positiv vs. negativ* oder *angenehm vs. unangenehm*. Das interessierende implizite Maß ist in vielen Fällen *relativer* Natur – es resultiert aus der Kontrastierung der Wertigkeitsassoziationen zum fokussierten Einstellungsobjekte (z.B. Homosexualität) mit den Wertigkeitsassoziationen zu einem Vergleichsobjekt (z.B. Heterosexualität). Implizite Maße dieser Art sind also Differenzmaße, womit sich einige methodische Schwierigkeiten verbinden (siehe Abschnitt 1.6.3). Die den impliziten Maßen tatsächlich zugrunde gelegten Prozesse sind noch zu weiten Teilen unerforscht. Dies ist u.a. auf die mangelhafte theoretische Fundierung und die immer noch unzureichende empirische Grundlagenforschung zurückzuführen (wie z.B. Rothermund & Wentura, 2001, 2004; Steffens, 2004). Insofern besteht keine hinreichende Klarheit über die Funktionsweise populärer indirekter Messverfahren, wie dem IAT. Abgesehen von den ungeklärten theoretischen Grundlagen impliziter Einstellungsmessung, zeichnet sich auch auf empirischer Ebene eine äußerst kontroverse Befundlage hinsichtlich der Bedeutung impliziter Maße ab (siehe unten). Zur Gewinnung gesicherter Erkenntnisse ist eine sowohl theoretisch, als auch empirisch kritische Auseinandersetzung mit diesen Messverfahren zu fordern (De Houwer, 2003b).

Inzwischen sind zur Messung von impliziten Einstellungen, Stereotypen, Vorurteilen, sogar Selbstwert und Selbstkonzept mittels Reaktionszeiten und Fehlerraten eine Reihe verschiedener indirekter Messinstrumente verfügbar (De Houwer, 2003b; Fazio & Olson, 2003). Die Vielfaltigkeit innerhalb dieser Verfahrensgruppe ist enorm. Für einen Überblick werden an dieser Stelle ausschließlich die zentralen Verfahren berücksichtigt, in die jeweils unterschiedliche Variationen eingeordnet werden könnten. Eine entsprechende Systematisierung reaktionszeitbasierter impliziter Messungen nimmt De Houwer, 2003b vor und auch Voß, Rothermund & Wentura (2003) vergleichen unterschiedliche Latenzmessverfahren miteinander. Die verschiedenen indirekten Reaktionszeitparadigmen können in vier Hauptkategorien unterschieden werden:

- *Affektives Priming* (Fazio, Sanbonmatsu, Powell & Kardes, 1986)
- *Affective Simon Task* (De Houwer & Eelen, 1998)
  - *Extrinsic Affective Simon Task* (De Houwer, 2003a)
- *Emotional Stroop Task* (Pratto, 1994; Pratto & John, 1991)
- *Impliziter Assoziationstest* (IAT, Greenwald, McGhee & Schwartz, 1998)
  - *Go/No-Go Task* (Nosek & Banaji, 2001)
  - *Single-Target-IAT* (Wigboldus, van Knippenberg & Holland, 2001).

Nach eingehender Betrachtung der kritischen Befundlage impliziter Einstellungsmaße wurden für die empirische Untersuchung anliegender Fragestellungen der relativ etablierte *Implizite Assoziationstest* (IAT) und der *Extrinsic Affective Simon Task* (EAST) als relativ unerforschtes, jedoch viel versprechendes Alternativmaß ausgewählt. Beide Verfahren folgen einem relativ standardisierten Ablauf, wobei der IAT dem enorm breit gefächerten Einsatz und den umfassenden Publikationen nach als etabliert bezeichnet werden kann. Der IAT ist zwar alles andere, als ein unanfechtbares Verfahren (vgl. Abschnitt 1.6.1.3), jedoch liegt eine Bandbreite an Validitätsnachweisen über Zusammenhänge mit expliziten Einstellungs- und Verhaltensmaßen vor. Der EAST, eines der neuesten impliziten Maße, wurde in Reaktion auf einige Nachteile des IAT entwickelt und hat u.a. einen entscheidenden Vorteil gegenüber dem IAT, der prinzipiell *relative*, keine *unabhängigen* Maße liefert: Im EAST können unabhängige Effekte für die Zielkonzeptkategorien betrachtet werden (vgl. De Houwer, 2003; Nosek, Greenwald & Banaji, 2005). Einige empirische Studien zum EAST liegen bereits vor, dies kann jedoch bislang nur als Orientierung betrachtet werden. IAT und EAST sollten in ihrer Funktionsweise, der Erfassung *kognitiver Assoziationsstärken*, miteinander vergleichbar sein (De Houwer, 2003a). Das Affektive Priming, das seinerseits eng verwandt ist mit der Emotional Stroop Task (z.B. Pratto, 1994, John & Pratto, 1991) und der Affective Simon Task (De Houwer & Eelen, 1998; vgl. Wentura & Rothermund, 2003) basiert dagegen auf *Bahnungs- und/oder Hemmungseffekte* (siehe De Houwer, 2003a; Voß, Rothermund & Wentura, 2003).

Der Einsatz von IAT *und* EAST bietet die Möglichkeit der Gegenüberstellung zweier theoretisch impliziter Maße, wobei der IAT als Vergleichsstandard betrachtet werden kann, da der Kenntnisstand zum EAST noch immer dürftig ist. Insofern kann der EAST noch nicht alleinige Basis eines validen impliziten Einstellungsmaßes sein. Einerseits ermöglicht jedoch der Vergleich des EAST mit einem relativ etablierten Verfahren wie dem IAT eine Evaluation des Verfahrens. Andererseits können spezifische Vorzüge des EAST genutzt werden. Die Auswahl dieser beiden verwandten Verfahren dient des Weiterem dem Ziel, die methodische Varianz *innerhalb* der impliziten Maße so gering wie möglich zu halten, da bereits gezeigt werden konnte, dass nicht nur die Beziehung zwischen impliziten und expliziten Maßen, sondern auch die verschiedener impliziter Maße ausgesprochen instabil ist (Fazio & Olson, 2003; Olson & Fazio, 2003).

## 1.6.1 Implicit Association Test (IAT)

### 1.6.1.1 *Entwicklung des IAT*

Der *Implicit Association Test* (IAT, Greenwald, McGhee & Schwartz, 1998) ist das derzeit wohl populärste experimentelle Messverfahren zur Erfassung impliziter Kognitionen. Diese Popularität lässt sich u.a. an der Zahl wissenschaftlicher Untersuchungen zum IAT ablesen, die bereits so kurze Zeit nach Veröffentlichung des Paradigmas einen beachtlichen Umfang angenommen haben. Darunter finden sich zahlreiche empirische Untersuchungen, die auf die Validität des IAT nicht nur zur Messung impliziter Einstellungen fokussieren, ebenso wie Arbeiten, die zum kritischen Diskurs beitragen. Zudem wurde kurze Zeit nach seiner Publikation durch Greenwald und Kollegen, der insbesondere aus Praktikabilitäts-, Kosten- und Aufwandsgründen attraktive Einsatz des IAT via Internet ermöglicht, was zur weiteren Verbreitung des Verfahrens führte. Hierdurch wurden enorme Stichprobenumfänge ermöglicht (vgl. Nosek, 2003), wobei ein Infragestellen der Güte der Ergebnisse aufgrund der mangelnden Kontrolle über die Messbedingungen berechtigt erscheint. Nicht zuletzt belebte die Auseinandersetzung mit dem IAT die Entwicklung alternativer indirekter Messverfahren. Den außerordentlichen Nachhall verdankt der IAT seinem relativ einfachen Aufbau, was sowohl in der Programmierung, als auch der Auswertung von Vorteil ist, seiner leichten Adaptationsfähigkeit auf verschiedene Fragestellungen, den vergleichsweise vertretbaren psychometrischen Eigenschaften, relativ großen Effektstärken und den verschiedentlich erbrachten Validitätsnachweisen. Mit der Veröffentlichung des IAT wurde die bereits in den Achtzigerjahren aufgekommene, jedoch weniger stringent verfolgte Auseinandersetzung um die Konzeptionalisierung insbesondere von (sozialen) Einstellungen wieder aufgegriffen und fruchtbringend vorangetrieben.

## 1.6.1.2 Aufbau des IAT

Im IAT werden Stimuli aus 2 verschiedenen Konzepten mit jeweils 2 kontrastierenden Kategorien präsentiert. Es handelt sich zum Einen um ein Zielkonzept mit zwei Kategorien, wie z.B. Homosexuell vs. Heterosexuell und zum anderen einem Attributkonzept mit den entgegengesetzten Kategorie wie positiv vs. negativ oder angenehm vs. unangenehm. Die Stimuli werden nach einem bestimmten Aufbau am Monitor präsentiert (vgl. Abb. 1-2). Den Probanden stehen zwei Antworttasten zur Verfügung, die Stimuli aus maximal vier verschiedenen Kategorien so schnell und so korrekt wie möglich unter die ausgewiesenen Kategorienlabel einzuordnen. Der IAT umfasst also für Stimuli des Zielkonzepts eine *Klassifizierungsaufgabe* und für Stimuli des Attributkonzepts eine *evaluative Entscheidungsaufgabe*. Versuchspersonen lernen im Verlauf der Messung, auf die Stimuli jeder Aufgabe je nach festgelegter Zuordnung mit entweder linkem oder rechtem Tastendruck zu reagieren. Nach der Einzelübung der Aufgaben werden Klassifizierungs- und evaluative Entscheidungsaufgabe gemischt, so dass mit zwei Antworttasten für die Reaktion auf nunmehr vier Stimuluskategorien (2 je Aufgabe) jeweils zwei Kategorien von Stimuli auf einer Taste liegen.

Stimuli	Kategorien des Zielkonzepts (Bsp.)		Homosexuelle vs. heterosexuelle Pärchen			
	Kategorien des Valenzkonzepts (Bsp.)		negative vs. positive Stimuli			
Aufbau	Block	Trials	Zweck	Aufgabe	linke Taste [A]	rechte Taste [5]
	1	20	Einzelübung	Attributdiskriminierung <i>evaluative Entscheidungsaufgabe</i>	negativ	positiv
	2	20	Einzelübung	Zielkonzeptdiskriminierung <i>Klassifizierungsaufgabe</i>	Homosexuell	Heterosexuell
	3	24	gemischter Übungsblock 1	kombinierte Diskriminierung	negativ × Homosexuell	positiv × Heterosexuell
	4	48	Experimentalblock 1	<i>kompatible Tastenbelegung</i>		
	5	20	Einzelübung	umgekehrte Zielkonzeptdiskriminierung	Heterosexuell	Homosexuell
	6	24	gemischter Übungsblock 2	kombinierte Diskriminierung	negativ × Heterosexuell	positiv × Homosexuell
7	48	Experimentalblock 2	<i>inkompatible Tastenbelegung</i>			
IAT-Effekt	<b>Reaktionszeiten und Fehlerraten</b>		gemischte Übungsblöcke:	Mittelwert (Block6) – Mittelwert (Block3)		
			Experimentalblöcke:	Mittelwert (Block7) – Mittelwert (Block4)		

Abbildung 1-2, Aufbau des IAT am Beispiel eines Homosexualitäts-IAT; grau schattiert: kritische Messblöcke

Den Kern des Impliziten Assoziationstest bildet die Doppelbelegung der Antworttasten, aus der kongruente und inkongruente Stimuluskombinationen resultieren. Für einen IAT, mit dem automatische Wertassoziationen z.B. gegenüber Homosexualität zur Einstellungs- bzw. Vorurteilmessung erfasst werden, stellt die Kombination von Stimuli für Homosexualität (z.B. gleichgeschlechtliche Paarbilder) mit negativ-wertigen Stimuli (z.B. positive Gesichtsausdrücke) auf einer Antworttaste die kompatible Tastenbelegung dar. Entsprechend ist auch die Kombination der kontrastierenden Kategorie Heterosexualität mit positiv-wertigen Stimuli auf der anderen Antworttaste kompatibel. Die inkompatible Tastenbelegung entsteht durch die umgekehrten Stimuluskombinationen, also die Beantwortung der Homosexualitäts- und positiven Stimuli mit der einen Taste und die der Heterosexualitäts- und negativ-Stimuli mit der anderen Taste. Blockweise variiert wird im Verlauf des IAT nun diese Kompatibilität der Stimuluskategorien, die einer gemeinsamen Taste zugeordnet sind.

Der subjektive Grad der Kompatibilität bzw. Inkompatibilität hat spezifische Auswirkungen auf die Schnelligkeit und Exaktheit der Beantwortung der Stimuli, was als Maß der Assoziationsstärke interpretiert wird. Eng miteinander assoziierte Konzepte sollten leichter mit einer Taste zu beantworten sein, als weniger stark assoziierte Konzepte. Nach De Houwer (2003a) ist der IAT-Effekt auf eine automatische Aktivierung von Antworten, die mit der Valenz des dargebotenen Stimulus *kompatibel* sind, zurückzuführen. Somit sollten Probanden mit negativeren Assoziationen zu z.B. Homosexualität die kompatiblen Blöcke mit kürzeren Reaktionszeiten und weniger Fehlern absolvieren bzw. die inkompatiblen Blöcke mit höheren Reaktionszeiten und -Fehlern, als Probanden mit positiveren Einstellungen. Da der eigentliche IAT-Effekt die Differenz der Mittelwerte (Reaktionszeiten, Fehler) aus inkompatiblen und kompatiblen Block darstellt, zeichnen sich Probanden mit negativeren Einstellungen durch entsprechend höhere IAT-Effekte aus.

### 1.6.1.3 *Empirische Validitätsnachweise für den IAT*

Es liegt eine überwältigend große Zahl an Untersuchungen vor, die die Validität des IAT zur indirekten Erfassung unterschiedlicher Dispositionen stützen, nichtsdestotrotz liegen auch diverse kritische Befunde zu IAT vor (siehe nächsten Abschnitt). Grundsätzlicher Tenor der empirischen Studien ist jedoch eine relative Bedeutung des IAT für selbstberichtete Einstellungen und/ oder Verhalten dokumentiert, wobei eindeutige Kriterien für die Relevanz impliziter Maße wie dem IAT nicht unbedingt an Korrelationen zu Selbstberichtmaßen abgeleitet werden können, da je nach Konzeptionalisierungsansatz von Einstellungen Zusammenhänge bzw. *keine* Zusammenhänge zwischen expliziten und impliziten Einstellungen erwartet werden (siehe Kapitel 2). Die Bedeutung des IAT zeichnet sich vielmehr dort ab, wo eine inkrementelle Validität zur Verhaltensvorhersage nachgewiesen wird. Entsprechend



positive Befunde zur Vorhersage unkontrollierten, spontanen Verhaltens berichten Asendorpf, Banse & Mücke (2002, Schüchternheit), Bosson, Swann & Pennebaker (2000, Selbstwert), Egloff & Schmuckle (2002, Ängstlichkeit) und auch Phelps et al. (2000), die zeigen, dass der IAT in Beziehung zu neurophysiologischen Prozessen steht.

Um den bis derzeit auszumachenden Einsatzbereich des IAT zu umreißen, sind zunächst diverse Einstellungsobjekte anzuführen, wie *Nationalitäten* (z.B. DasGupta & Greenwald, 2001; Gawronski, 2002; Kühnen et al., 2001; McConnell & Leibold, 2001; Neumann & Seibt, 2001; Ottaway, Hayden & Oakes, 2001), *Alte Menschen, altersbezogene Einstellungen und -Identität* (Hummert, Garstka & O'Brien, 2002; Karpinski & Hilton, 2001; Rudman, Greenwald & Mellott., 1999), *Männer vs. Frauen* (Aidman & Carroll, 2003; Skowronski & Lawrence, 2001; Steffens & Plewe, 2001), *Eigen- vs. Fremdgruppenbias bei Minderheiten* (Rudman, Feinberg & Fairchild, 2002) und *Religiöse Gruppen* (Greenwald & Mellott., 1999; Rowatt, Franklin & Cotton, 2005). Des Weiteren wurde der IAT zur Messung persönlichkeitsnaher geprüft, wie *Selbstwert und Selbstkonzept* (Aidman & Carroll, 2003; Greenwald & Farnham, 2000; Greenwald et al., 2002), *Schüchternheit* (Asendorpf, Banse & Mücke, 2002) und *Ängstlichkeit bzw. Furcht* (Egloff & Schmuckle, 2002; Egloff, Wilhelm & Neubauer, 2002; Teachman & Woody, 2003). Daneben lassen sich IAT-Studien anführen zu *Konsumentenpräferenzen* (Karpinski & Hilton, 2001; Maison, Greenwald & Bruin, 2001) und *stigmatisiertem vs. nicht stigmatisiertem Verhalten*, wie Rauchen und Vegetarismus (Swanson, Rudman & Greenwald, 2001).

Von besonderer inhaltlicher Relevanz für vorliegende Arbeit sind die Untersuchungen von Banse, Seise & Zerbes (2001) und Seise, Banse & Neyer (2002). Die Autoren erfassen implizite Vorurteile gegenüber *Homosexualität* via IAT und können erwartungskonform zeigen, dass sich heterosexuelle von homosexuellen Menschen signifikant in ihren impliziten Einstellungen gegenüber Homosexualität unterscheiden. Ebenso von spezieller Bedeutung ist die Untersuchung von Teachman et al. (2003), in der der IAT zur Messung eines impliziten *Anti-Fat Bias* im amerikanischen Sprachraum eingesetzt wurde. Darüber hinaus konnte die Validität des IAT für Einstellung von Übergewichtigen zu hochkalorischen Lebensmitteln nachgewiesen werden (Roefs & Jansen, 2002, siehe Kapitel 3.2.2.1).

#### 1.6.1.4 Validitätsbereich des IAT

Trotz der umfassenden und in der Mehrheit viel versprechenden Befundlage, ist die Diskussion um die Verfahrensvalidität des IAT eines der kontroversesten Themen innerhalb der impliziten Einstellungsmessung. Da der IAT die jeweilige assoziative Stärke zwischen den beiden globalen Kategorien des Zielkonzepts (z.B. Heterosexuelle vs. Homosexuelle) und den beiden Polen der Attributdimension (angenehm vs. unangenehm) erfassen soll, spiegelt der IAT-Effekt Bewertungsassoziationen auf Kategorien-Level, also globalem Niveau wider. Offensichtlich misst der IAT also weniger *individuelle*, son-

dern eher gesellschaftliche oder *kulturelle* Bewertungsassoziationen (Karpinski & Hilton, 2001; Nosek, Banaji & Greenwald, 2002).

Allerdings ist die Funktionsweise des IAT über eine Erfassung von *Bewertungsassoziationen* generell umstritten (Brendl, Markman & Messner, 2001; Rothermund & Wentura, 2001, 2004) und der Forschungsbedarf zu den Prozessen, die dem IAT-Maß unterliegen, ungestillt. Es liegen verschiedene alternative Erklärungsansätze für den IAT-Effekt vor, die jedoch ihrerseits bislang ebenso wenig wie das Postulat von Greenwald & Kollegen bestätigt werden konnten (z.B. Brendl, Markman & Messner, 2001; Fazio & Olson, 2003; Gawronski, 2002; Kawakami & Dovidio, 2001; Mierke & Klauer, 2001; Rothermund & Wentura, 2001). Steffens, Lichau & Still (2004) untersuchen die Prozesse, die dem IAT-Effekt zugrunde liegen und kommen unter anderem zu dem Schluss, dass ein individueller IAT-Wert kein eindeutig interpretierbares Kriterium abgibt. Es bleibt entsprechend unklar, ob die Assoziationen von vorurteilsrelevantem Zielkonzept, (alte Menschen) und „unangenehm“ auf bewertete Oberflächeneigenschaften der Stimuli zurückzuführen sind (z.B. eine Abneigung gegen alte Namen) oder tatsächlich auf eine negativere Einstellung gegenüber alten Menschen. Steffens & Plewe (2001) bezeichnen Effekte *stimulusimmanenter* Valenzunterschiede als „Cross-Category-Associations“. Hierbei handelt es sich um a priori existente Konzept-Attribut-Assoziationen, die auftreten, wenn die Stimuli der beiden gegenübergestellten Konzepte (z.B. Blumen-Insekten) bereits mit angenehmen bzw. unangenehmen Attributen besetzt sind (z.B. Rose, Veilchen... vs. Ohrenkneifer, Spinne...). Ein entsprechender IAT-Effekt würde in diesem Falle keine individuellen Präferenzen widerspiegeln, sondern allgemeine Konzepteigenschaften. Es ist daher unbedingt auf eine Ausgewogenheit der Stimuli hinsichtlich Valenz und – insbesondere bei bildhaften Stimuli – hinsichtlich Attraktivität und Sympathie zu achten.

Allgemein problematisch im IAT ist die notwendige Kontrastierung des Zielkonzepts. So werden in einem IAT-Homosexualität bspw. keine reinen wertbezogenen Assoziationen zu Homosexualität erfasst, sondern zusätzlich die wertbezogenen Assoziationen zur Kontrastkategorie Heterosexualität. Diese Kontrastierung von Einstellungsobjekten führt entsprechend nur zu *relativen* Effekten. Demgemäß konnten Mitchell, Nosek & Banaji (2003) zeigen, dass IAT-Effekte von den jeweiligen Kontextbedingungen wie der Vergleichskategorie stark beeinflusst werden. Dazu kommt, dass es unter Umständen problematisch sein kann, überhaupt eine sinnvolle Vergleichskategorie für das fokussierte Einstellungsobjekt aufzustellen, z.B. wenn implizite Kognitionen zu komplexeren Einstellungsinhalten von Interesse sind (z.B. Farbige Frauen, Weiße Männer; siehe Mitchell, Nosek & Banaji, 2003) oder nicht klar ist, *welche* der vielen möglichen Kontrastkategorien die am meisten angemessene ist (z.B. für Rauchen, siehe Swanson, Rudman, Greenwald, 2001).

Da der IAT-Effekt ein Differenzmaß ist, unterliegt das Maß möglicherweise psychometrischen Dilemmata (Buchner & Brandt, 2003; May & Hittner, 2003; Williams & Zimmerman, 1996, Abschnitt 1.6.3).

## 1.6.2 Extrinsic Affective Simon Task (EAST)

### 1.6.2.1 *Entwicklung des EAST*

Während der IAT aufgrund seiner relativen Simplizität in Programmierung und Auswertung, der erzielten hohen Effektstärken, einer im Vergleich zu anderen Verfahren befriedigenden Reliabilität und einer Vielzahl validitätsunterstützender Befunde in der empirischen Forschung nach wie vor das Verfahren der Wahl darstellt, sind immer noch hinreichend viele Kritikpunkte anzuführen (insbesondere die Alternativerklärungen zum IAT-Effekt: Umkodieren der Aufgabe aufgrund von Salienz-Asymmetrien oder Oberflächeneigenschaften). Diese Tatsache inspirierte Jan De Houwer (2003a) eine „modifizierte Version“ des IAT zu entwickeln: Den *Extrinsic Affective Simon Task* (EAST), so benannt aufgrund der strukturellen Ähnlichkeit der Zielkonzepttrials zur Affective Simon Task (AST, De Houwer & Eelen, 1998). Mit dem EAST ist somit ein weiteres indirektes Messverfahren für implizite Einstellungen und Vorurteile auf der Basis von Antwortlatenzen und Fehlerraten entwickelt worden, das jedoch noch zu wenig evaluiert ist, um als hinreichend verlässlich bezeichnet werden zu können.

Die Modifikation im EAST greift einen großen Nachteil des IAT-Effekts auf, der aus einem Vergleich zwischen zwei verschiedenen Aufgaben, nämlich der des inkompatiblen mit der des kompatiblen Blocks, resultiert. Der EAST ermöglicht aufgrund seines Aufbaus dagegen den Vergleich von Trials innerhalb *einer* Aufgabe durch die Relevanz eines leicht manipulierbaren Oberflächenmerkmals für die Beantwortung der Target-Stimuli und eine extrinsische Valenzzuweisung der Antworttasten.

### 1.6.2.2 *Aufbau des EAST*

Im EAST werden kann *jeder* einstellungsrelevante Stimulus als seine eigene Kontrolle fungieren, da die Stimuli der Zielkonzeptkategorien (z.B. Homosexuell vs. Heterosexuell) aufgrund nicht-evaluativer Merkmale, wie z.B. die Oberflächenfarben blau vs. grün, oder bei verbalen Stimuli zwei semantische Kategorien wie Verb vs. Substantiv, kategorisiert werden sollen (De Houwer, 2003b). Probanden müssen im EAST prinzipiell zwei Stimulusarten unterscheiden, auf die sie jeweils nach unterschiedlichen Stimulusmerkmalen reagieren sollen. Ist der Stimulus beispielsweise ein bläulich oder grünlich eingefärbtes Bild, so soll je nach *Farbe* mit rechtem (z.B. bei bläulicher Farbe) oder linkem Tastendruck (z.B. bei grünlicher Farbe) reagiert werden. Ist der Stimulus dagegen beispielsweise ein Schwarzweißfoto, so soll auf der Basis der *Stimulusvalenz* reagiert werden. Dies stellt eine evaluative Entscheidungsaufgabe vergleichbar derer im IAT dar. Bei angenehmen Bildinhalten (z.B. lachende Gesichter) soll die eine- und bei unangenehmen Bildinhalten (z.B. wütende oder traurige Gesichter) die andere Taste gedrückt werden.

Die Tastenbelegung für die Gesichter und die Farbzuzuordnung bleibt über das gesamte Experiment hinweg unverändert. Es wird angenommen, dass die beiden Antworttasten durch die konstante Valenzzuordnung die Bedeutung einer „Gut-Taste“ und einer „Schlecht-Taste“ annehmen (vgl. Abb. 1-3). Variiert wird im EAST-Experiment lediglich das für die Aufgabe relevante Oberflächenmerkmal der Stimuli aus den beiden Zielkonzeptkategorien (z.B. die Farbe). Jeder einzelne Stimulus wird also sowohl in bläulicher, als auch in grünlicher Farbe präsentiert. Wenn davon auszugehen ist, dass die Antworttasten durch Üben der evaluativen Entscheidungsaufgabe tatsächlich extrinsisch mit jeweils einer Wertigkeit besetzt sind, können die Zielkonzeptstimuli in sich selbst sowohl kompatibel, als auch inkompatibel sein. Gegeben, dass jeweils bläuliche Zielkonzeptstimuli und angenehme Valenzstimuli auf einer Taste liegen, sowie grünliche Zielkonzeptstimuli und unangenehme Valenzstimuli, so sind bläulich gefärbte Stimuli für Heterosexualität und entsprechend grünlich gefärbte Stimuli für Homosexualität *kompatibel*, während bläuliche Stimuli für Homosexualität und grünliche für Heterosexualität *inkompatibel* sind. Somit kann *jeder* einstellungsrelevante Stimulus als seine eigene Kontrolle fungieren (vgl. De Houwer, 2003a).

<b>Stimuli</b>	<b>Kategorien des Zielkonzepts</b>		Homosexuelle und. heterosexuelle Pärchen in bläulicher bzw. grünlicher Farbe			
	<b>Kategorien des Valenzkonzepts</b>		schwarzweiße negative vs. positive Stimuli			
<b>Aufbau</b>	<b>Block Trials</b>	<b>Zweck</b>	<b>Aufgabe</b>	<b>linke Taste [A]</b>	<b>rechte Taste [5]</b>	
	<b>1</b>	24	Einzelübung	Diskriminierung der Attributkonzeptstimuli nach Bildinhalt (Valenz)	negativ	positiv
	<b>2</b>	24	Einzelübung	Diskriminierung der Zielkonzeptstimuli nach Oberflächenmerkmal (Farbe)	grünlich	bläulich
	<b>3</b>	32	gemischter Übungsblock	kombinierte Diskriminierung nach Bildinhalt (Valenz) bzw. Oberflächenmerkmal (Farbe)	negativ × grünlich	positiv × bläulich
	<b>4..</b>	53...	Experimentalblöcke...			
<b>EAST-Effekt</b>	<b>Reaktionszeiten und Fehlerraten</b>	gemischter Übungsblock: Experimentalblock 1: Experimentalblock 2:		inkompatible – kompatible Zielkonzeptstimuli		

**Abbildung 1-3**, Aufbau des EAST am Beispiel eines Homosexualitäts-EAST

### 1.6.2.3 Empirische Validitätsnachweise und Validitätsbereich des EAST

Zur Validität des EAST liegen noch zu wenige Befunde vor, um eine allgemeine Aussage über die Güte des Verfahrens treffen zu können. Nichtsdestotrotz ist das EAST-Paradigma vielversprechend und zog seit der Publikation im Jahre 2003 bereits eine Reihe an empirischen Untersuchungen nach sich, die jedoch bislang unveröffentlicht sind (z.B. De Houwer, 2005; Ellwart, Rinck & Becker, 2005; Perugini, O’Gorman & Connor, 2005; Schmitz, Teige, Voss & Klauer, 2005; Stahl & Degner, 2005). Bislang publiziert sind erst drei empirische Studien, in denen der EAST eingesetzt wurde (De Houwer, 2003a; De Houwer, Crombez, Koster & De Beul, 2004; Teige et al., 2004). Darüber hinaus sind nach Wissensstand der Autorin weitere zwei Arbeiten zur Publikation angenommen (De Raedt, Schacht & De Houwer, 2004; Huijding, de Jong & ter Hart, 2003).

In seiner Originalarbeit zum EAST weist De Houwer (2003a) die kritische Interaktion zwischen Stimulusart und extrinsischer Valenz für die Reaktionszeiten sowie die Fehlerdaten nach. Entsprechend der extrinsischen Valenz der Antworttasten resultierten kompatibel gefärbte Zielkonzeptwörter in längeren Reaktionszeiten und höheren Fehlerraten im Vergleich zu Wörtern, deren Färbung extrinsisch kompatibel war. In dieser Studie wurden jedoch keine interindividuellen Unterschiede im EAST-Effekt und deren Bedeutung untersucht. Ihrer Untersuchung zu alkoholbezogenen Assoziationen gemäß berichten De Houwer, et al. (2004) jedoch statistisch unbedeutsame Interaktionen zwischen Stimulusart und extrinsischer Valenz. Darüber hinaus bestehen in dieser Studie keine Korrelationen zwischen dem EAST und dem IAT. Teige et al. (2004) misslingt die Adaptation des EAST zur simultanen Messung von selbstkonzeptbezogenen Konstrukten wie Schüchternheit, Ängstlichkeit und Ärger. Die EAST-Scores weisen weder eine akzeptable interne Konsistenz auf (siehe unten), noch bestehen Außenkorrelationen zu Selbstberichtmaßen oder dem IAT. Möglicherweise eignet sich das Verfahren also nicht zur Messung relativ valenzneutraler Traits. Da der EAST zur Messung affektiver Assoziationen entwickelt wurde, kann dieses Ergebnis nicht als Beweis *gegen* die allgemeine Tauglichkeit des Verfahrens betrachtet werden, sondern zeigt lediglich die Grenzen der Validität auf. Als ein interessanter Nebenbefund der Teige-et-al-Studie ist allerdings der beobachtete emotionale Stroop-Effekt anzuführen. Die Probanden waren allgemein langsamer in ihren Reaktionen auf *negative* Stimuli. Daran wird einerseits deutlich, dass verschiedene Prozesse am EAST-Effekt beteiligt sein könnten und andererseits, dass das Verfahren ein Potential zur Messung affektiver Assoziationen birgt.

Die wenigen bislang veröffentlichten Studien zum EAST weisen relativ einhellig auf eine äußerst unbefriedigende Reliabilität des EAST-Effekts hin (z.B. De Houwer & De Bruycker, 2005; Ellwart, Rinck & Becker, 2005). Dies muss jedoch keine unabwendbare Charakteristik des Verfahrens sein. Teige et al. (2004) führen für die geringe Reliabilität des EAST in ihrer Untersuchung vier Gründe an:

- a) Mangelnde Repräsentativität der ausgewählten Stimuli
- b) Geringe interindividuelle Varianz in den Einstellungen der Probanden zum Einstellungsobjekt
- c) Prozedural variierende Element wie z.B. randomisierte Stimulusabfolge oder Tastenbelegung
- d) Geringes Aggregationsniveau der EAST-Scores (zu wenig Trials)

Im Gegensatz zum IAT werden im EAST die gemittelten Einstellungen zu den Kategoriebeispielen, nicht zur globalen Kategorie selbst erfasst. Wenn die verwendeten Stimuli nicht einheitlich genug sind, um das Zielkonzept ausreichend gut zu repräsentieren, ist der resultierende EAST-Effekt möglicherweise nicht interpretierbar. Mehr noch als für den IAT ist hier demnach die sorgfältige Auswahl der einzelnen Stimuli von zentraler Bedeutung. Zudem konnte gezeigt werden, dass die Individualität spezifischer Mitglieder stigmatisierter Gruppen stark divergierende Einstellungen auslösen kann (siehe Blair, 2002), weshalb insbesondere bei der Verwendung von Bildstimuli große Sorgfalt auf die Repräsentativität des verwendeten Materials zu legen ist. Aufgrund der Sensibilität des EAST auf Kategorienbeispiele ist das Verfahren besser als der IAT dazu angelegt, individuelle anstelle sozial geteilter Vorurteile zu messen, da der IAT Assoziationen vor allem auf der Ebene der globale Kategorienlabel misst. Repräsentative Stimuli einer Zielkonzeptkategorie für die interindividuelle Unterschiede in den evozierten Assoziationen zu erwarten sind, eine festgelegte Stimulussequenz und eine ausreichende Anzahl von Trials je Block könnten die Reliabilitätsproblematik eventuell lösen (vgl. Banse, 2001; De Houwer, 2003a; Frings & Wentura, 2003).

### 1.6.3 Zusammenfassende Betrachtung – Reliabilität und Konvergenz impliziter Maße

Ein häufig berichteter Befund sind keine oder geringe Zusammenhänge zwischen verschiedenen impliziten Maßen (z.B. Brauer, Wasel & Niedenthal, 2000; Fazio & Towles-Schwen, 1999; Sherman, Presson, Chassin & Rose, 1999; Teige et al., 2004). Es kann eine Reihe möglicher Ursachen für die mangelnde Korrespondenz der impliziten Maße aufgeführt werden:

Zuvorderst kritisch zu beurteilen ist die relativ geringe Reliabilität der impliziten Maße, wobei hier deutlich zwischen verschiedenen Paradigmen zu unterscheiden ist. Zudem liegt in vielen Arbeiten der Fokus stärker auf Gruppenmittelwerten, als auf interindividuellen Unterschieden, was möglicherweise erklärt, weshalb Reliabilitäten nur selten dokumentiert sind. Vergleichsweise hohe interne Konsistenzen werden für den IAT berichtet, wobei auch hier eine enorme Variationsbreite auffällt. Hofmann, Gawronski, Gschwendner, Le & Schmitt (in press) finden in ihrer Meta-Analyse bei 32% der 147 betrachteten IAT-Studien Reliabilitätsbefunde und berechnen eine mittlere interne Konsistenz bzw.

Split-Half-Reliabilität von .80 (39 Studien). Für die Test-Retest-Reliabilität des IAT ermitteln sie dagegen ein Wert von .54 (8 Studien). Die Reliabilitäten anderer impliziter Maße liegen deutlich unter denen, die für die IAT berichtet werden. Banse (2001) dokumentiert das Cronbach's Alpha für verschiedene Priming-Effekte, die von .00 bis .57 variieren. Cunningham, Preacher & Banaji (2001) untersuchen verschiedene Reaktionszeitmaße und finden für das Response-Window Priming eine interne Konsistenz von .64 sowie eine relativ hohe Stabilität von .68.

Die wenigen Reliabilitätsbefunde für den EAST deuten eine äußerst geringere interne Konsistenz der Messung an. De Houwer (2003a) berichtet Reliabilitäten von -.25 bis .55 (Mittelwert: .25) für Fehler- und Reaktionszeiteffekte des EAST, führt jedoch inhaltliche und prozedurale Ursachen für die geringen Reliabilitäten an. Zum einen können für ein Blumen-Insekten-Zielkonzept keine großen interindividuellen Unterschiede erwartet werden, dagegen findet De Houwer (2003a) höhere Reliabilitäten von .48 und .44 (Reaktionszeiten und Fehlereffekte) für den Selbstwert-EAST. Diese sind für eine Untersuchung interindividueller Differenten jedoch immer noch nicht ausreichend hoch. De Houwer führt die randomisierte Stimulusdarbietung und die geringe Anzahl an verrechneten Trials für die EAST-Effekte als weitere mögliche Ursachen an. Für letzteren Einflussfaktor der Reliabilität liefern Hofmann et al. (in press) eine Bestätigung. Sie finden eine substantiell positive Korrelation zwischen der Anzahl an Trials und der Reliabilität für den IAT ( $r=.51$ ). Teige et al. (2004) räumen jedoch auch inhaltliche Ursachen für die geringen internen Konsistenzen von -.01 (Schüchternheit, Fehlerraten), bis .24 (Ängstlichkeit, Reaktionszeiten). Wie bereits angeführt ist der EAST zur Messung valenzbezogener Assoziationen konzipiert, woraus Validitätseinschränkungen für z.B. eine Selbstkonzeptmessung, resultieren können.

Die Reliabilitätsbefunde resümierend bleiben relativ geringe interne Konsistenzen, Split-Half- und Retest-Reliabilitäten für implizite Maße festzuhalten (siehe auch Banse, Seise, Zerbes, 2001; Boniecki, Jacks & Zuwerink, 2002; Bosson, Swann & Pennebaker, 2000; Cunningham, Preacher & Banaji, 2001; Fazio & Olson, 2002; Gawronski, 2002; Kawakami & Dovidio, 2001). Dies liefert zumindest eine Teilerklärung für die geringen Effektstärken der impliziten Maße und die allenfalls geringen Korrelationen zwischen den Kennwerten *verschiedener* indirekter Verfahren (siehe Fazio & Olson, 2003; Olson & Fazio, 2003). Hierbei ist jedoch zunächst zu hinterfragen, *welche* Reliabilitätsmethode für die Effektmaße der Reaktionszeiten und Fehlerraten als adäquates Kriterium betrachtet werden kann. Darüber hinaus sind implizite Maße Differenzmaße, für die unter Umständen die aus der Veränderungsmessung bekannte Reliabilitätsproblematik berücksichtigt werden muss (vgl. Meier & Perrig, 2000; Williams & Zimmermann, 1996). Die klassischen und auch hier angewandten Reliabilitätsberechnungen setzen die Annahmen der Klassischen Testtheorie voraus, wie beispielsweise gleiche Varianzen und unkorrelierte Messfehler.

Beides muss für die impliziten Reaktionszeitparadigmen nicht unbedingt gültig sein, was relativ offensichtlich erscheint. Mit der Split-Half-Methodik werden beispielsweise Ermüdungserscheinungen vernachlässigt, die über die Messung hinweg unvermeidlich sind, nicht unbedingt linearer Natur sein müssen und sich dementsprechend sich stark reliabilitätsmindernd auf reaktionszeit- sowie fehlerbasierte Differenzeffekte auswirken können. Retest-Reliabilitäten dagegen erfordern mindestens zweimaliges Messen, was zum einen sehr aufwendig ist und zum anderen die Problematik von Lerneffekten berücksichtigen muss. Zudem sind Retest-Reliabilitäten in der Regel nicht so hoch wie interne Konsistenzen (Banse, Seise, Zerbes, 2001). Dass implizite Maße einer starken Messfehlerbehaftung unterliegen, die sich zusammenhangsmindernd auswirkt, zeigen Cunningham, Preacher & Banaji (2001), die in einer *latenten* Variablenanalyse deutlich höhere Zusammenhänge zwischen zwei verschiedenen IAT-Varianten und einem semantischen Priming-Paradigma erzielen, als auf der Ebene manifeste Pearson-Korrelationen. Neben der allgemeinen Reliabilitätsproblematik von Differenzenscores kann auch eine Ursache für die mangelnde Reliabilität indirekter Verfahren wie dem IAT und EAST die standardmäßige Darbietung der Stimuli in randomisierter Reihenfolge sein (vgl. Banse, 2001; De Houwer, 2003a; Teige et al., 2004, siehe Studie I).

Ein weiterer möglicher Grund für die mangelnde Korrespondenz impliziter Maße ist in ihren unterschiedlichen Funktionsweisen zu suchen. Wie bereits angeführt unterscheiden sich die indirekten Verfahren diesbezüglich teilweise beträchtlich voneinander. Während IAT und EAST beispielsweise auf der Erfassung kognitiver Assoziationsstärken basieren sollen, werden für das Affektive Priming sogar drei mögliche Funktionsprinzipien diskutiert: Die dabei am wenigsten bestätigte Hypothese ist die der Aktivierungsausbreitung im semantischen Netzwerk. Nahe liegend dagegen ist eine Beteiligung Stroop-ähnlicher Mechanismen von Antwortkonflikten infolge einer Hemmung bei affektiver Inkongruenz (vgl. De Houwer, Hermans, Rothermund & Wentura, 2002; Voß, Rothermund & Wentura, 2003). Klauer & Musch (2003) postulieren zusätzlich affective-matching Mechanismen, die bei Plausibilität positive Reaktionen (GO!) und bei Implausibilität negative Reaktionen (NO GO!) fördern. Entsprechend spiegelt der Priming-Effekt die Höhe des Kongruenzeffektes, also den quantitativen Vorteil kongruenter relativ zu inkongruenten Trials, wider. Des Weiteren wurde oben angeführt, dass der IAT kognitive Bewertungsassoziationen auf *Kategorienebene* erfasst. Hierin unterscheidet sich der IAT von den meisten anderen, eher stimulusbasierten impliziten Maßen. Aufgrund dieser eher globalen Assoziationsmessung kann der IAT-Effekt, insbesondere bei inhaltlich relativ breit gefächerten Stereotypen oder starken Divergenzen zwischen persönlichen und kulturellen Stereotypen, deutlich vom EAST-Effekt abweichen, für den die *einzelnen* Stimuli eine größere Rolle spielen. Dass dieser methodische Unterschied von Bedeutung ist, unterstützen Arbeiten, denen zufolge eine Fokussierung auf *Subtypen* der Analyse *globaler* Kategorienlabel vorzuziehen ist, da letztere zu viele divergierende und möglicherweise sogar orthogonale Cluster vereint (z.B. Devine & Baker, 1991).



De Houwer (2003b) schlägt verschiedene Adaptionen vor, um die Vergleichbarkeit indirekter Messungen zu erhöhen. Es wird jedoch auch deutlich, dass sowohl die Inhalte der untersuchten Stereotype als auch das Stimulusmaterial zur Messung besonderer Aufmerksamkeit in der Versuchsplanung bedarf. Die mangelnde Konvergenz der impliziten Maße kann, unter der Voraussetzung hinreichend gegebener Validität, also auf verschiedene Ursachen zurückgeführt werden, wobei hauptsächlich die allgemein geringe Reliabilität sowie die abweichenden Funktionsweisen der impliziten Maße anzuführen sind. Insbesondere das Reliabilitätsproblem könnte zudem ursächlich für die geringen Zusammenhänge zwischen impliziten und expliziten Maßen sein (Meier & Perrig, 2000; Buchner & Wippich, 2000; Buchner & Brandt, 2003). Diesbezüglich sind Auswertungsmethoden zu präferieren, die eine Kontrolle des Messfehlers ermöglichen, was jedoch im Hinblick auf die notwendigen Stichprobenumfänge keine gängige experimentalpsychologische Praxis darstellt. Vorrang vor der Verfolgung angewandter Fragestellungen mit impliziten Verfahren muss noch immer die Sicherstellung ihrer Zuverlässigkeit und Gültigkeit haben. Dieser Problematik ist Studie I gewidmet (siehe Kapitel 4).

## 2 Bedeutung impliziter & expliziter Einstellungen

### 2.1 Empirische Befundlage

Da die impliziten Maße als eine neue, vorteilhaftere Herangehensweise an die Messung von Einstellungen und anderen Dispositionen eingeführt worden sind, wären prinzipiell positive Beziehungen zwischen impliziten und entsprechenden expliziten Maßen zu erwarten. Die empirischen Befunde sprechen jedoch eine andere, immer noch nicht hinreichend interpretierbare Sprache: Eine überwältigende Anzahl an Arbeiten berichtet keine oder nur sehr geringe Zusammenhänge zwischen expliziten und impliziten Messungen (Banse & Gawronski, 2003; Devine, 1989; Dovidio, Kawakami, Johnson, Johnson & Howard, 1997; Gaertner & McLaughlin, 1983; Greenwald & Banaji, 1995; Greenwald, McGhee & Schwartz, 1998; Hofmann et al., in press; vgl. Fazio & Olson, 2003). Einige andere Arbeiten weisen dagegen positive Beziehungen nach (Kawakami, Dion & Dovidio, 1998; Lepore & Brown, 1997; McConnell & Liebold, 2001; Wittenbrink, Judd & Park, 1997, 2001).

In einer Metaanalyse zur Beziehung zwischen impliziten (IAT) und expliziten Maßen finden Hofmann et al. (in press) entsprechend eine mittlere minderungskorrigierte Korrelation von  $r=.24$  (Range:  $-.25$  bis  $.60$ ) zwischen. Dementsprechend ist zu resümieren, dass Zusammenhänge nicht stabil nachzuweisen sind und die Beziehung zwischen impliziten und expliziten Messwerten bestenfalls gering ist. Zudem variiert die Korrespondenz in Abhängigkeit von den Einstellungsobjekten, Kontextvariablen und verschiedenen anderen Faktoren (Ashburn-Nardo, Voils & Monteith, 2001; Dasgupta & Greenwald, 2001; Kawakami, Spears & Dovidio, 2002; Kawakami, Young & Dovidio, 2002; Kawakami, Dovidio & Dijksterhuis, 2003; Lowery, Hardin & Sinclair, 2001; Mitchell, Nosek & Banaji, 2003; Nosek, 2003; Pratto & Shih, 2000; Rudman, Ashmore & Gary, 2001; Wittenbrink, Judd & Park, 2001). Ekehammar, Akrami & Araya (2003) finden in drei Experimenten mit einem Primingparadigma stabile Geschlechtsunterschiede für implizite und explizite Ablehnung schwedischer Immigranten: Während Frauen stärkere implizite Vorurteile gegenüber Immigranten aufweisen als Männer, zeigen diese höhere explizite Vorurteile als Frauen. Diese Ergebnisse könnten jedoch beispielsweise auf die höhere Sensibilität in der Bildverarbeitung unter weiblichen Probanden zurückzuführen sein (Anooshian & Seibert, 1996; McGivern et al., 1998). In Übereinstimmung zur „expliziten“ Angstforschung finden Egloff & Schmukle (2004) bei Frauen eine ausgeprägtere Ängstlichkeit in expliziten Maßen. Dies gilt auch für die impliziten Maße IAT und Emotional Stroop Task.

Es konnte zudem gezeigt werden, dass implizite und explizite Einstellungen oder Motive das Verhalten in unterschiedlicher Weise beeinflussen (z.B. Zajonc, 1980). Während explizite, also reflektierte, Einstellungen vor allem bewusstes, geplantes, willentliches Verhalten vorhersagen können, stehen implizite, also automatische, Einstellungen zu spontanem, schwer kontrollierbarem und unfreiwilligem

Verhalten in Beziehung (z.B. Amodio, Harmon-Jones & Devine, 2003; Czopp, Monteith, Zimmerman & Lynam, 2004; Dovidio et al., 1997; Neumann, Hülßenbeck & Seibt, 2004; vgl. auch Fazio & Olson, 2003). An dieser Stelle wird der Bezug zum MODE-Modell von Fazio (1990) nochmals deutlich. Während explizite Einstellungen die steuerbaren Komponenten im Modell repräsentieren, können die impliziten Einstellungen mit den automatischen Komponenten in Verbindung gebracht werden. Eine Berücksichtigung beider Komponenten sollte die größtmögliche Verhaltensaufklärung erbringen.

### 2.1.1 Konzeptualisierungsansätze für implizite und explizite Einstellungen

Die Dissoziation verschiedener Maße einer Disposition ist nach Dovidio & Fazio (1992) insbesondere bei sozial sensitiven Themen jedoch nicht verwunderlich, sondern aufgrund sozialer Erwünschtheits-tendenzen sogar zu erwarten. Die interindividuell variierende Tendenz, sich sozial erwünscht zu verhalten, beeinflusst vornehmlich *direkte* Messungen, z.B. Selbstberichte in Fragebögen, und kann sich stark mindernd auf die Korrespondenz zu impliziten Maßen auswirken, da diese relativ wenig Spielraum für kontrollierte Verfälschungen bieten (z.B. Banse, Seise & Zerbis, 2001; Steffens, 2004). Unter Umständen stellen die neuen indirekten Verfahren nunmehr jedoch keine alternativen Messmethoden für bisher hauptsächlich über Selbstbericht erfasste Dispositionen dar, sondern erfassen *distinkte* Konstrukte. Brauer, Wasel & Niedenthal (2000) resümieren diesen Sachbestand:

*“Different tasks have been designed to measure the two processes. Thus, although social psychologists have remained interested in the same phenomenon over the years—that is, in prejudice—in devising different kinds of assessment techniques they may no longer be measuring the same construct.”*

Hier zeichnet sich enormer Klärungsbedarf ab. Die Hauptfrage ist also nunmehr, ob die indirekten Verfahren zur Messungen impliziter Kognitionen einen optimierten und verzerrungsfreien Zugang zu verhaltenswirksamen Komponenten einer zugrunde liegenden Einstellungsdisposition darstellen (*Same Construct Approach*, Wittenbrink, et al. 1997) oder ob es sich bei impliziten und expliziten Kognitionen um verschiedene Konstrukte handelt (*Dissociation Approach*, Devine 1989; vgl. Brauer, Wasel & Niedenthal, 2000). Um diese Frage hinreichend zu klären, kann eine empirische Konstruktvalidierung von Einstellungen nützlich sein, bei der eine notwendige Bedingung für Konstruktvalidität die Konvergenz verschiedener Methoden zur Erfassung einer Einstellungsdisposition ist. Zentrales Augenmerk liegt demnach auf einer adäquaten Messung von Einstellungen, wobei hier die Funktionsweise der verschiedenen indirekten Verfahren, die Interpretation insbesondere impliziter Messwerte und die Beziehung zwischen impliziten und expliziten Maßen, auch in Bezug auf moderierende Drittvariablen, von hoher Relevanz sind.

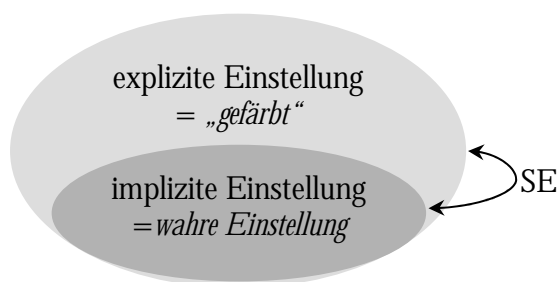
Die Entwicklung im Bereich der Einstellungsforschung wirft notwendigerweise neue Fragen auf und trotz der fruchtbaren Forschungsergebnisse rekurriert die Einstellungsforschung nach wie vor auf vielen Spekulationen. Fazio & Olson (2003, 304) erheben angesichts der Unsicherheit über die verschiedenen Einstellungsmaße traditioneller direkter oder neuerer indirekter Erfassung die Frage nach der „wahren“ Einstellung:

*„Will the real attitude please stand up?“*

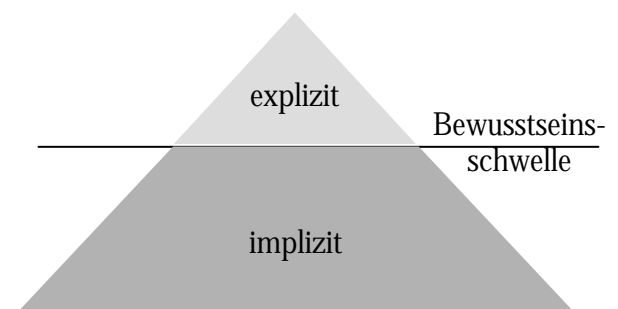
Hierin verbirgt sich die uralte Suche nach dem wahren Kern von Einstellungen, die, nachdem Phänomene wie Soziale Erwünschtheit bei der Messung psychologischer Dispositionen Berücksichtigung fanden, seit Generation von Einstellungsforschern verfolgt wird. Die Hoffnung, dieser wahre Kern könnte mittels impliziter Messungen erfassbar werden, findet jedoch keine ausreichende Bestätigung. Zur Beziehung zwischen expliziten und impliziten Einstellungen stellen Nosek & Banaji (2002) die zwei in der Fachwelt diskutierten Positionen einander gegenüber: Die „Lügendetektorperspektive“ und die „Perspektive unabhängiger Bewertung“, die im folgenden diskutiert werden.

#### 2.1.1.1 Lügendetektorperspektive soziale Einstellungen

Aus der Lügendetektor-Perspektive werden implizite Einstellungen als „True Score“ und explizite- als zu Selbstpräsentationszwecken verfärbte Einstellungen betrachtet (Abb. 2-1). Bei Karpinski & Hilton (2001) findet sich die Lügendetektorperspektive repräsentiert als „Eisbergmodell“ (Abb. 2-2), wobei explizite Einstellungen auf der Ebene bewusster Kontrolle und implizite Einstellungen auf der Ebene unbewusster, unkontrollierbarer Kognitionen angeordnet werden.



**Abbildung 2-1,** Lügendetektorperspektive auf die Korrespondenz impliziter und expliziter Einstellungen; SE = Soziale Erwünschtheit (Nosek & Banaji, 2002)



**Abbildung 2-2,** Eisbergmodell expliziter und impliziter Einstellungen (Karpinski & Hilton, 2001)

Modelle dieser Konzeptualisierung sind unter dem „Same Construct Approach“ zusammen zu fassen (Wittenbrink et al., 1997), in dem von *einem gemeinsamen* Einstellungskonstrukt ausgegangen wird und

insofern auch von der grundsätzlichen Existenz positiver Zusammenhänge zwischen expliziten und impliziten Einstellungen. Die Befundlage fehlender oder geringer Korrelationen zwischen beider Kategorien von Maßen wird auf die Auswirkungen bewusster, willentlicher Steuerung von expliziten Messungen bzw. auf Artefakte der Messmethoden für explizite und implizite Einstellungen zurückgeführt (Brauer, Wasel & Niedenthal, 2000; Nosek & Banaji, 2002).

Empirische Untermauerung für den Ein-Konstrukt-Ansatz liefern sowohl Untersuchungen, die auf manifester Ebene signifikant positive Korrelationen zwischen impliziten und expliziten Maßen finden (z.B. Frings & Wentura, 2003), als auch Arbeiten, die diese Zusammenhänge mit latenter Modellierung aufdecken (z.B. Cunningham, Preacher & Banaji, 2001), sowie Studien, in denen Moderatoreffekte durch beispielsweise Soziale Erwünschtheit auf die Beziehung expliziter und impliziter Messwerte oder nachgewiesen werden (z.B. Hoffmann, Gschwendner & Schmitt, 2003).

### 2.1.1.2 Perspektive unabhängiger Bewertungen

Dem Ein-Konstrukt-Ansatz entgegen steht die *Perspektive unabhängiger Bewertung*, bei der von einer unabhängigen Wirksamkeit unterschiedlicher Ebenen der Informationsverarbeitung ausgegangen wird, weshalb eine Dissoziation in zwei distinkte Konstrukte explizite und implizite Einstellungen nahe liegt (Abb. 2-3). Dieser Ansatz schreibt impliziten und expliziten Bewertungsmodi einen jeweils eigenständigen Bedeutungsanteil zu<sup>1</sup> und postuliert keine speziellen Faktoren (wie z.B. Soziale Erwünschtheit), die eine Beziehung zwischen beiden Konstrukten erklären könnten. Die Perspektive unabhängiger Bewertung korrespondiert mit dem Modell zweier distinkter, aber interagierender Determinanten sozialen Verhaltens von Strack und Deutsch (2004, siehe oben).



**Abbildung 2-3,** *Perspektive unabhängiger Bewertung auf die Beziehung impliziter und expliziter Einstellungen (z.B. Devine, 1989; Wilosn, Lindsey & Schooler, 2000)*

<sup>1</sup> Dabei könnte es sich sowohl um verschiedene Ebenen von Bewusstheit, als auch um unterschiedliche kognitive oder affektive Komponenten handeln (Banaji, 2004)

Modelle der Perspektive unabhängiger Bewertungen repräsentieren einen so genannten „Dissociation Approach“ impliziter und expliziter Maße. Devine (1989) zum Beispiel postuliert in ihrem Dissoziationsmodell lediglich einen gemeinsamen Ausgangspunkt von Einstellungen bzw. Stereotypen, und zwar die sozialisierten, also *kulturellen Stereotype*. Diese besitzen für alle Mitglieder einer Gesellschaft Gültigkeit, werden frühzeitig gelernt, und zwar noch vor der Entwicklung der Kritikfähigkeit (vgl. Prohansky, 1966), internalisiert und automatisiert. Kulturelle Stereotype werden leicht, geradezu automatisch aktiviert beim Kontakt mit dem Einstellungsobjekt (z.B. einem Farbigen). Dieses tradierte Wissen repräsentiert bei einem Erwachsenen die *impliziten* Einstellungen. Aufgrund individueller Lerngeschichten und daraus resultierender persönlicher Überzeugungen kann sich eine Spaltung zwischen den frühzeitig sozialisierten und den später entwickelten, individualisierten Einstellungen vollziehen, in der Art, dass neue Erfahrungen *explizite* Einstellungen formen, die von den ursprünglichen impliziten Einstellungen abweichen. Diese Abweichung ist eine Funktion von Lernerfahrung und Motivation, sie bedeutet einen Konflikt zwischen kulturellen Stereotypen und persönlichen Überzeugungen und erfordert bei Personen, die eine solche Dissoziation erleben, energieaufwendige kognitive Strategien, um den Automatismen der kulturellen Stereotype zu begegnen. Die persönlichen Überzeugungen werden als motivierte individualisierte, gut elaborierte und dem Bewusstsein zugängliche Kognitionen verstanden (vgl. Abschnitt 1.4).

In erwachsenen Personen finden sich Devine zufolge nach sowohl kulturelle, als auch individuelle Einstellungen, die voneinander abweichen können („low-prejudice persons“), aber nicht zwangsläufig müssen („high-prejudice persons“). Fehlende oder instabile Zusammenhänge zwischen expliziten und impliziten Maßen werden im Dissoziationsansatz mit den interindividuellen Unterschieden in dieser Abweichung erklärt und ähnlich dem Ein-Konstrukt-Ansatz zu Lasten einer Variation in den *expliziten* Einstellungen gelegt. Das heißt, während vorurteilsvolle Personen eine breite Überlappung zwischen kulturellen Stereotypen und ihren persönlichen Überzeugungen repräsentieren, wonach starke positive Zusammenhänge zwischen entsprechenden Maßen zu erwarten sind, unterscheiden sich vorurteilsarme Personen durch willentlich modifizierte Überzeugungen, die in starkem Kontrast zum kulturellen Stereotyp stehen können. Hier handelt es sich also um einen Moderatoreffekt der Motivation zur Vorurteilsfreiheit (siehe Abschnitt 2.2.2). Ein wichtiges Unterscheidungsmerkmal des Dissoziationsmodells von Moderatoransätzen des Ein-Konstrukt-Ansatzes ist die *motiviert* Ausprägung einer neuen, bewusst gesteuerten, nicht automatisierten und verbalisierbaren Einstellungsdisposition.

In ähnlicher Weise argumentieren Wilson, Lindsey & Schooler (2000), die in ihrem „Model of Dual Attitudes“ implizite Einstellungen wie folgt definieren:

*“Evaluations that (a) have an unknown origin (i.e., people are unaware of the basis of their evaluation); (b) are activated automatically; and (c) influence implicit responses, namely, uncontrollable responses and one people do not view as an expression of their attitude and thus do not attempt to control.”*

Die Autoren betrachten implizite Einstellungen als ehemalige explizite Einstellungen, die von neuen expliziten Einstellungen abgelöst, jedoch nicht überschrieben wurden. Vergangene Einstellungen verbleiben als latente und automatisch operierende implizite Einstellungen und ähneln insofern Eigenschaftsdispositionen. Explizite und implizite Einstellungen können miteinander korrespondieren, im Falle dualer, also divergierender Einstellungen, wird die implizite Einstellung automatisch aktiviert, während die explizite Einstellung Kapazität und Motivation verlangt (vgl. MODE-Modell, Abschnitt 1.3). Auch bei Aktivierung der expliziten Einstellungen, beeinflussen die impliziten Einstellungen vorrangig unreflektierte bzw. unkontrollierbare Verhaltensweisen. Das Modell dualer Einstellungen postuliert weiterhin eine relative Veränderungsanfälligkeit expliziter im Vergleich zu impliziten Einstellungen, die als eigenschaftsähnliche Kognitionsinhalte relativ stabil sind oder sich allenfalls langsam verändern. Empirische Bestätigung für das Dissoziationsmodell liefern Untersuchungen, die keine Zusammenhänge zwischen expliziten und impliziten Einstellungsmaßen aufzeigen (z.B. Karpinski & Hilton, 2001) bzw. zeigen können, dass sich motivationale Subsamples durch entsprechend heterogene Beziehungsmuster auszeichnen (z.B. Plant & Devine, 1998).

### 2.1.2 Diskussion der verschiedenen Konzeptualisierungsansätze

Die beiden gegenübergestellten Konzeptualisierungsansätze von Einstellungen setzen mehr oder weniger eine Art Allgemeingültigkeit für implizite und explizite Einstellungen voraus, entweder in Form eines gemeinsamen Konstruktes oder in Form distinkter Dispositionen. Einige empirische Untersuchungen können jedoch zeigen, dass sowohl Charakteristika des Einstellungsobjektes selbst, als auch definierte Kontextvariablen belangreich für Einstellungsmaße sind und insofern auch Bedeutung für die Korrespondenz expliziter und impliziter Messung haben (Nosek & Banaji, 2001, 2002). Mitchell, Nosek & Banaji (2003) veranschaulichen beispielsweise, wie implizite Einstellungen in Abhängigkeit von der kontextuellen Einbettung der Messung variieren können. Sie bestätigen die Hypothese, dass farbige Frauen bei Kontrastierung mit weißen Frauen negativ bewertet werden (Kontext Rasse), während die Kontrastierung mit farbigen Männern zu einer positiven Bewertung farbiger Frauen führt (Kontext Geschlecht), da Frauen im allgemeinen leicht positiver bewertet werden als Männer (Carpenter & Banaji, 2000; Eagly & Mladinic, 1989). Weiße Männer dagegen werden in der Kontextbedingung Rasse besser bewertet, als in der Kontextbedingung Geschlecht.

Diese Ergebnisse unterstreichen die Komplexität von Einstellungen, die bei impliziter und expliziter Messung berücksichtigt werden muss.

In einer aktuellen Arbeit postulieren Nosek & Hansen (2004) Einstellungen als nicht-statische Konstrukte, die unmittelbar von der jeweiligen Messmethode abhängig sind. Unterschiedliche Messmethoden sollten demnach nicht als mehr oder weniger valide bewertet werden, sondern repräsentieren verschiedener Einstellungsaspekte. Dies macht zum einen die Notwendigkeit eines mehrdimensionalen Methodenansatzes zur Messung von Einstellungen bzw. Vorurteilen deutlich und untermalt zum anderen erneut die Suche nach der Grundgesamtheit verschiedener Einstellungsmaße. Fazio & Olson (2003, 8) beurteilen die Verfolgung verschiedener Erklärungsansätze zur Beziehung zwischen impliziten und expliziten Einstellungsmaßen:

*“In our view the variability regarding the correspondence between implicit and explicit measures indicates that discussion of whether a relation exists is not very productive. We already know enough to be able to say that the question has no simple answer. That is, the answer is ‘it depends’.”*

Anstelle der Untersuchung der absoluten Frage nach der Existenz einer Explizit-Implizit-Beziehung, raten die Autoren zu einer weiteren Untersuchung von Faktoren, die den Zusammenhang erklären. Entsprechend subsumieren Fazio & Olson (2003, 304) die angesichts des aktuellen Kenntnisstands zukünftig notwendigen Untersuchungsfragen:

*„When, under what conditions, and for what kind of people, are implicit and explicit measures related?“*

Solche Einflussfaktoren auf die Korrespondenz impliziter und expliziter Einstellungen können messthematischen (Einstellungsobjekt), personenbezogenen (Motivation, Persönlichkeit), situativen (Erhebungskontext) oder rein methodischen bzw. statistischen Ursprungs sein. So führen Wittenbrink, Judd & Park (1997) v.a. methodische Gründe wie die Stimulusauswahl und das Prozedere für die unterschiedlichen Zusammenhangsergebnisse an. Karpinski & Hilton (2001) diskutieren darüber hinaus die Folgen einer Einschränkung des Variabilitätsranges der Einstellungsmaße, die sich unmittelbar auf mögliche Zusammenhänge mit anderen Variablen auswirkt. Für explizite Vorurteilsmaße ist in spezifischen Stichproben (z.B. Studenten, vgl. Six, 1996) eine allgemein sehr liberale Ausprägungen zu beobachten. Die Höhe einer Korrelation zwischen zwei Variablen, von denen mindestens eine einen sehr geringen Range aufweist, ist natürlicherweise begrenzt. Die Möglichkeit eingeschränkter Variabilität muss demnach in der Auswertung berücksichtigt werden.

In jüngerer Zeit wird vermehrtes Augenmerk auf den Einfluss verschiedener personaler und situativer Moderatorvariablen gelegt. Ihrer Metaanalyse zufolge subsumieren Hofmann et al. (in press, S. 18):

*„...the overall IAT-explicit-relationship can be diagnosed as heterogenous“.*



Auf einem hohen Abstraktionslevel bestehen offenbar Beziehungen zwischen expliziten und impliziten Messungen. Die Autoren schätzen dabei, dass etwa ein Drittel der Varianz in den Beziehungen zwischen dem IAT und Selbstberichtmaßen auf Moderatorvariablen zurückzuführen ist (vgl. Kapitel 2.2). In einem kurzen Überblicksartikel vollzieht Rudman (2004) eine Verortung der Quellen impliziter Einstellungen und führt zu Beginn der Betrachtung an, dass, obwohl die bisherige Befundlage noch keine endgültigen Schlüsse zulässt, von unterschiedlichen Quellen automatischer und kontrollierter Bewertung auszugehen ist und daher eine Konzeptualisierung verschiedener Konstrukte nahe liegt. Er führt eine stärkere Beeinflussbarkeit impliziter Einstellungen bzw. die gegensätzliche Wirkung auf explizite und implizite Einstellungen durch vier Faktoren an (Abb. 2-4), womit der Autor die theoretische und empirisch untermauerte Distinktheit der Konstrukte verdeutlicht. In seinem Resümee schreibt Rudman (2004, 81) schließlich:

*“For one thing, implicit and explicit attitudes sometimes correlate well. This is particularly true when attitude objects are noncontroversial (...), but even measures of implicit and explicit attitudes can be dissociated even for noncontroversial objects (...). Thus the challenge for future research is to uncover the variables that determine when implicit and explicit attitudes converge.”*

Quellen	Implizite Einstellungen	Explizite Einstellungen
<b>Frühe Erfahrungen:</b> z.B. Devine (1989); Banaji & Greenwald (1995)	repräsentieren frühkindliche, vergessene, nicht mehr bewusste Erfahrungen → Indirektes (implizites), vorverbales Lernen	repräsentieren kürzliche, zugängliche Erfahrungen → Direktes (explizites), verbales Lernen
<b>Emotionale Erfahrungen:</b> z.B. Amodio, et al. (2004); Phelps et al. (2000)	Stehen in Beziehung zur Aktivität der Gehirnareale für emotionale Kontrolle (v.a. Amygdala) und zu emotionsbasierten Prädiktoren → emotional gesteuert	Stehen in Beziehung zur Aktivität Gehirnarealen für bewusste, kontrollierte Informationsverarbeitung und zu kognitiven Prädiktoren → kontrolliert, kognitiv gesteuert
<b>Kulturelles Milieu:</b> z.B. Livingston (2002); Nosek, et al. (2002); Rudman, et al. (2002)	Korrespondieren mit sozialem Status: z.B. zum kulturellen Stereotyp konforme Eigengruppenabwertung in Minderheiten (z.B. arme o. übergewichtige Menschen)	Stehen nicht oder in entgegen gesetzter Beziehung zu sozialem Status: Keine Abwertung bzw. sogar Aufwertung der Eigengruppen bei Mitgliedern von Minderheiten
<b>Prinzipien kognitiver Konsistenz:</b> Greenwald et al. (2002)	Folgen den Prinzipien kognitiver Konsistenz	Folgen nicht zwangsläufig den Prinzipien kognitiver Konsistenz

**Abbildung 2-4,** Quellen impliziter und expliziter Einstellungen nach Rudman (2004)

Gemäß der Reflexion der verfügbaren Literatur, bleiben nun mehrere Schlüsse zu ziehen: Grundsätzlich besteht eine enorme methodische Variabilität allein innerhalb der impliziten Vorurteilmessungen, welche gemeinsam mit oft nicht berücksichtigten Individual-, Prozedur- und Kontextfaktoren zumindest partiell eine Erklärung für divergierenden Befunde liefern kann. Neben diesen messmethodischen Eigenheiten der Verfahren können noch weitere Ursachen für die instabilen Zusammenhänge zwischen impliziten und expliziten Einstellungen in Betracht gezogen werden. Ein Weg, wertvolle Informationen für eine Konzeptualisierung expliziter und impliziter Einstellungsmaße bereitzustellen, ist eine Konstruktvalidierung, die es ermöglicht Trait- und Methodenspezifische Varianz unter Berücksichtigung des Messfehlers auf ihre Bedeutung hin zu suchen. Gerade, da immer wieder auf die geringen Reliabilitäten impliziter Maße hingewiesen wird, ist eine solche Untersuchung viel versprechend. Auf die Bedeutung einer Konstruktvalidierung impliziter Maße weisen u.a. auch Cunningham, Preacher & Banaji (2001, 164) hin, indem sie das entscheidende Missverhältnis im Einsatz indirekter Messverfahren kritisieren:

*„Despite the wide usage of response-latency measures of implicit attitudes, their construct validity remains largely untested.“*

Im folgenden werden einige der Faktoren besprochen, für die ein moderierender Einfluss auf den Zusammenhang expliziter und impliziter Messung nachgewiesen werden konnte.

## **2.2 Moderatoren der Beziehung impliziter & expliziter Einstellungen**

Inzwischen ist die Bedeutung von Moderatoreffekten in der Korrespondenz impliziter und expliziter Einstellungen nicht mehr anzuzweifeln. In ihrer, hier bereits viel zitierten Metaanalyse finden Hofmann et al. (in press) deutliche Hinweise auf eine moderierende Wirkung der Beziehung expliziter und impliziter Einstellungen für verschiedene Variablen. Allgemeine Kandidaten zur Moderation der Implizit-Explizit-Korrespondenz sind z.B. der Grad der Selbstpräsentation oder bewussten Adjustierung, die Fähigkeit zur Introspektion, die Stärke und Dimensionalität der betrachteten Einstellung und die wahrgenommene Diskrepanz zwischen den eigenen Einstellungen und denen der Bezugsgruppe (vgl. Gschwendner-Lukas, Hofmann & Schmitt, 2004; Nosek, 2003). Einige dieser allgemeinen Moderatoren werden zunächst kurz vorgestellt. Im Anschluss erfolgt eine ausführlichere Betrachtung der speziell für soziale Einstellungen und Vorurteile relevanten Moderatorvariablen *Motivation zur Vorurteilkontrolle*, die zu großer Prominenz für die explizit-implizit Korrespondenz gelangte.

### 2.2.1 Allgemeine Moderatoren der Implizit-Explizit-Korrespondenz

In verschiedenen Fällen stellt bereits das Geschlecht eine Moderatorvariable der Implizit-Explizit-Korrespondenz dar. Pelham et al. (2005) zeigen in sechs Studien, dass für Frauen eine engere Beziehung zwischen implizitem und explizitem Selbstwert besteht, als für Männer. Sie führen diesen Befund auf die allgemeine Sozialisation von Frauen zurück, eine höhere Sensibilität für ihre Gefühle und Intuitionen auszubilden. Daraus sollte für Frauen eine ausgeprägtere Fähigkeit zur *Introspektion* resultieren – eine Fähigkeit, die nach Hofmann, Gschwendner und Schmitt (2003) die Implizit-Explizit-Beziehung positiv moderiert. Je höher die Introspektion einer Person, umso stärker sind beispielsweise ihre Einstellungen elaboriert und desto näher ist ein Selbstbericht der Messungen unreflektierter, spontaner Kognitionen. Konstrukte, die individuelle Introspektionsneigung/ bzw. -fähigkeit widerspiegeln sind z.B. *Private Selbstaufmerksamkeit* (Fenigstein, Schreier & Buss, 1975) oder das *Einstellungswissen* (Hofman, Gschwendner und Schmitt, 2003). Für die Korrelation zwischen expliziter und impliziter Ängstlichkeiten finden Egloff & Schmukle (2004) ebenso eine stärkere Korrespondenz unter Frauen im Vergleich zu Männern. Auch für dieses Resultat sind geschlechtsspezifische Unterschiede in der Introspektion angstbezogener Kognitionen als Ursache denkbar.

Es könnten jedoch auch spezifische Eigenschaften der gemessenen Disposition sein, die gemeinsam mit dem Geschlecht die unterschiedlichen Zusammenhangsmuster zwischen expliziten und impliziten Maßen bedingen (vgl. Nosek, 2003). Dies kann an einem einfachen Beispiel gezeigt werden: Es würde niemanden verwundern, wenn hinsichtlich der Einstellungen zu Fussballmannschaften eine höhere Korrespondenz zwischen impliziten und expliziten Messungen für Männer besteht (vgl. Degner & Wentura, 2005). In diesem Fall würde die unterschiedliche *Einstellungsstärke* der zentrale Aspekt des Einstellungsobjektes sein, welches die Stärke der Beziehung zwischen impliziten und expliziten Maßen beeinflusst. Bereits Fazio et al. (1986) sowie Bargh, Chaiken, Gøvdender & Pratto (1992) weisen auf die Bedeutung der Einstellungsstärke in Hinblick auf die *Zugänglichkeit* der Einstellung hin. Als Indikator der Zugänglichkeit bezeichnen sie die Latenz einer evaluativen Bewertung. Die Zugänglichkeit ist jedoch nur einer von zwei zentralen personalen Moderatoren im Modell von Hofmann, Gschwendner und Schmitt (2003). Eine gelungene Introspektion ist die Voraussetzung für eine bewusste Repräsentation der zugrunde liegenden Einstellungsdisposition. Erst wenn einer Person die bewusste Repräsentation über eine interessierende Einstellung vorliegt, kommt eine weitere personale Moderatorvariable der Implizit-Explizit-Korrespondenz in Spiel, die in diesem Rahmen unter vielerlei Namen diskutiert wird – Hofman, Gschwendner und Schmitt (2003) bezeichnen sie als Tendenz der *Ajustierung*, Klein, Snyder & Livingston (2004) als *Self-monitoring* und Nosek (2003) als *Selbstpräsentation*. Es gäbe noch eine Reihe anderer Begriffe für die bewusst verfärbte Selbstpräsentation einer Person bzw. ihrer

Einstellungen, weithin bekannt ist diese Moderatorvariable jedoch unter der Bezeichnung *Soziale Erwünschtheit*. Soziale Erwünschtheit sollt sich lediglich in *expliziten* Messungen niederschlagen.

Im speziellen Fall der Messung negativer Einstellungen wird die Tendenz zur Adjustierung direkter Einstellungsmessung mit der so genannten *Motivation zur Vorurteilstkontrolle* bzw. *-freiheit* beschrieben.

## 2.2.2 Spezifische Moderatoren der Implizit-Explizit-Korrespondenz:

### Die Motivationen zur Vorurteilstkontrolle

Ende der Achtziger, Anfang der Neunziger Jahre postulieren verschiedene Modelle zur sozialen Urteilsbildung interindividuelle Unterschiede im Rückgriff auf reflektierte versus heuristische Prozesse bei der Urteilsbildung und darüber hinaus unterschiedliche Ausmaße an Motivation, über die eigene Urteilsbildung zu reflektieren (z.B. Brewer, 1988; Snyder, 1992). Devine (1989) konnte zeigen, dass Stereotype zwar geteiltes Wissen darstellen, das im Gedächtnis eines jeden Individuums mehr oder weniger repräsentiert ist und bei einem relevanten Reiz automatisch aktiviert wird, dass sich Personen jedoch hinsichtlich der aktiven Unterdrückung stereotyper Informationsverarbeitung unterscheiden. Das MODE-Modell von Fazio (1990) bestätigend findet Devine (1989) unter kognitiver Belastung („cognitive load“) eine stereotypenkonsistente Bewertung ambivalenter Verhaltensweisen sowohl unter hoch-, als auch unter niedrig vorurteilsbehafteten Personen. Das bedeutet, wenn Personen keine Möglichkeit haben, ihren automatisch aktivierten, stereotypen Gedächtnisinhalten durch bewusstes Reflektieren zu begegnen, greifen sie gleichsam auf diese zurück. Devine kann also zeigen, dass sich Personen im Grad ihrer Nutzung stereotyper Informationen nur dann voneinander unterscheiden, wenn ausreichend kognitive Ressourcen vorhanden sind. Zugleich weist die Autorin den starken Einfluss von persönlichen Überzeugungen wie *Gleichheit* und einer *Ablehnung von Stereotypen* auf die automatische Informationsverarbeitung nach. Außerdem können verschiedene Arbeiten nachweisen, dass manche Personen sogar Schuldgefühle erleiden, wenn sie realisieren, dass sie Vorurteile haben und diese nicht ausreichend kontrollieren können (z.B. Devine, Monteith, Zuwerink & Elliot, 1991; Monteith, Devine & Zuwerink, 1993). Blair (2001) und Blair, Ma & Lenton (2001) veranschaulichen zudem die Effektivität verschiedener Strategien zur Kontrolle von Vorurteilen in expliziten und impliziten Maßen. Solche Strategien sind beispielsweise habituelle Unterdrückung, Vermeidung stereotypischer Informationen bzw. Signale oder auch das aktive Suchen nach stereotypenkonträren Informationen.

Fazio, Jackson, Dunton & Williams (1995) entwickeln zur Überprüfung ihres MODE-Modells die *Motivation to Control Prejudice Reactions Scale* (MCPR-Scale) und zeigen die moderierende Wirkung der Motivation zur Vorurteilstkontrolle auf die Korrespondenz impliziter und expliziter Einstellungen auf.

Den Autoren zufolge resultiert aus geringer Motivation zur Vorurteilstkontrolle eine stärkere Implizit-Explizit-Korrespondenz für die Einstellungen zu Afroamerikanern. Den Ergebnissen entsprechend hat das Ausmaß der Motiviertheit zur Vorurteilstkontrolle einer Person nur Einfluss auf ihre *expliziten*, also kontrollierbaren Einstellungen. Hoch vorurteilsfrei motivierte Personen haben positive explizite Einstellungen ungeachtet ihrer durchaus negativeren automatisierten impliziten Einstellungen (vgl. Dunton & Fazio, 1997). Die Diskrepanz zwischen impliziten und expliziten Vorurteilen ist also auf das Wechselspiel zwischen interindividuell variierender Motivation, die eigenen Vorurteile zu verbergen bzw. zu überkommen und dem unterschiedlichen Grad bewusster Steuerbarkeit von direkten und indirekten Messungen zurückzuführen. Die MCPR-Skala von Fazio und Kollegen ist jedoch nicht in der Lage, zwischen den *Quellen* der Motivation zur Vorurteilstkontrolle zu unterscheiden. Zum einen ist es denkbar, dass Personen ihre Vorurteile aufgrund politischer Korrektheitsstandards unterdrücken mit dem Ziel, sich sozial erwünscht zu verhalten. Zum anderen kann das individuelle Wertesystem eine intrinsische Quelle der Motivation, keine Vorurteile zu haben, sein.

Um hinsichtlich der Motivation zur Vorurteilsfreiheit zwischen sozialer Erwünschtheit und innerer Überzeugung unterscheiden zu können, entwickelten Plant & Devine (1998) die „*Motivation to Respond Without Prejudice Scales*“, die aus zwei Subskalen – der internalen (IMS) und der externalen Motivation (EMS) – besteht. Die beiden Teilskalen EMS und IMS, die am besten als Motivation zur Vorurteils*frei*heit zu bezeichnen wären, sind unkorreliert und repräsentieren demnach inhaltlich distinkte Motivationen. Die Autoren konnten verschiedentlich zeigen, dass hoch internal und zugleich gering external motivierte Probanden sowohl in expliziten als auch in impliziten Messungen, darüber hinaus selbst unter schwer kontrollierbaren Bedingungen (kognitiver Belastung) und sogar in psychophysiologischen Markern (z.B. Lidschlussreflex) weniger Vorurteile aufweisen (vgl. Amodio, Hamon-Jones & Devine, 2003; Devine, Plant, Amodio, Harmon-Jones & Vance, 2002). Devine et al. (2002) schreibt der Untersuchung von Selbstbestimmungs- und Selbstregulationsprozessen für die Einstellungs- und Vorurteilsforschung eine zentrale Bedeutung zu, da eine ausgeprägte Selbstregulation unabhängig von äußeren Faktoren zu erfolgreicher Werte- und Zielverfolgung analog zu den eigenen Standards befähigt. Demgemäß kann die internale Motivation zur Vorurteilsfreiheit, die Werte wie Gleichberechtigung und Universalismus widerspiegelt, eine wirksame Vorurteilsminimierung zur Folge haben (vgl. *Self-Determination Theory*, Deci & Ryan, 1985, 2000).

In Anlehnung an die MCPR-Skala von Fazio et al. (1995) entwickeln Banse & Gwaronski (2003) die deutschsprachige Skala „*Motivation zu vorurteilsfreiem Verhalten*“ (MVV-16). Sie zeigen in Übereinstimmung zu den Befunden von Fazio und Kollegen eine negative Moderatorwirkung von MVV für die Beziehung zwischen explizitem Sexismus und implizitem Geschlechtsstereotyp, die auf systematisch verringerten Vorurteilsausprägungen ausschließlich im *expliziten* Sexismus bei hoch motivierten Probanden zurückzuführen ist.

### 2.3 Implizite & explizite Motive

Bei der Messung von vorurteilsbezogenen Motiven, die gewissermaßen eine Brücke zwischen verschiedenen Einstellungsmessung schlagen sollen, wird selten beachtet, dass diese auch via Selbstbericht erfasst werden und somit allen oben aufgeführten Einschränkungen direkter Messmethoden unterliegen. Es verwundert insofern wenig, dass verschiedene direkter Vorurteilstrollmaße zu expliziten Einstellungen in Beziehung stehen, nicht jedoch zu impliziten. Inwiefern die Vorurteilstrollskala selbst zur Vorurteilstmessung werden kann, ist nicht bekannt, diese Möglichkeit zu berücksichtigen, liegt jedoch nahe.

Nach einer theoretischen Betrachtung der Forschungsbelege zur Motivdiagnostik resümieren Schmalt und Sokolowski (2000) entsprechend, dass Fragebogenmaße am wenigsten in der Lage seien, Motive realitätsnah zu messen, und ergo kaum mit tatsächlichem Verhalten in Verbindung stehen. Die Interdependenz von Motiven, Einstellungen und Verhalten spiegelt sich eben bereits auf messtechnischer Ebene wider: Vorurteilsfreies Verhalten muss nicht unbedingt erst im realen sozialen Kontext zum Tragen kommen, sondern kann bereits ein möglichst vorurteilsfreies Antwortverhalten sein. Das im Selbstbericht gemessene Motiv zu vorurteilsfreiem *Verhalten*, könnte sich so auch in dem Motiv, entsprechende Fragebögen vorurteilsfrei zu *beantworten* niederschlagen. Das kann sogar non-reaktive Skalen wie beispielsweise die *Modern Racism Scale* (MRS) von McConahay (1986) betreffen (vgl. Fazio, Jackson, Duntan & Williams, 1995).

McClelland, Koestner und Weinberger (1989) und Koestner, Weinberger & McClelland (1991) unterscheiden explizite von impliziten Motiven in ähnlicher Art und Weise, wie zwischen expliziten und impliziten Einstellungen getrennt wird (z.B. Greenwald & Banaji, 1995, siehe Abb. 1-1). Explizite Motive resultieren aus direkter Messung in Form von Selbstberichten, während implizite Motive nur indirekt erfasst werden können, z.B. mit dem *Thematischen Apperzeptionstest* (TAT, Murray, 1943), der von McClelland und Kollegen (1989) als „Fantasy Measure“ bezeichnet wird. Selbstberichtete, also explizite, Motive repräsentieren bewusste und external orientierte Motive, die durch soziale Erwünschtheit beeinflusst werden können. Implizite Motive stellen dagegen internale Standard dar, so genannte „Needs“ (vgl. Koestner, Weinberger & McClelland, 1991). Es herrscht dahingehend Konsens über die Differenzierbarkeit von Motiven, dass *explizite Motive* auf soziale Lernerfahrungen zurückgeführt werden, verstärkt auf soziale Appelle ansprechen, zudem eher mit dem kognitiven System in Verbindung stehen und eine enge Beziehung zum Selbstkonzept unterhalten. Im Gegensatz dazu sollten *implizite Motive* nicht bewusst, stärker an Emotionen geknüpft und eher genetisch determiniert sein, worin ihre Ansprechbarkeit auf angeborene Auslöser begründet liegt (vgl. Devine et al., 2002; Dovidio, et al., 1997; Kawakami & Dovidio, 2001; McClelland, 1995; Plant & Devine, 1998; Schmalt & Sokolowski, 2000; Wittenbink, Judd & Park, 1997).

Explizite und implizite Motive sind weitestgehend unkorreliert (z.B. McClelland, Koestner und Weinberger, 1989). Zudem stehen sie in unterschiedlicher Beziehung zu Verhalten (z.B. DeCharms, Morrison, Reitman & McClelland, 1955). Die empirische Befundlage zeigt, dass implizite Motive den expliziten Motiven insbesondere hinsichtlich der Vorhersage von Verhalten deutlich überlegen sind (z.B. Marlowe, 1959; McClelland & Boyatzis, 1982).

Trotz ihres geringen korrelativen Zusammenhanges und ihres unterschiedlichen Einflusses auf Kognitionen und Emotionen ist anzunehmen, dass implizite und explizite Motive gemäß unabhängiger Systeme miteinander interagieren (Schmalt & Sokolowski, 2000). Bei der Betrachtung von Verhalten, inklusive Antwortverhalten, ist die Beteiligung expliziter und impliziter Motive nur schwer voneinander zu trennen. Folglich – und damit wieder zum Ausgangspunkt der Betrachtung – erscheint der alleinige Einsatz *expliziter* Selbstberichte für eine valide Vorurteilmessung unzulänglich.

## 2.4 Das nomologische Netzwerk sozialer Einstellungen

### 2.4.1 Nomologisches Netzwerk expliziter Einstellungen

Das nomologische Netzwerk expliziter Einstellungen und Vorurteile ist bereits relativ gut untersucht. Eine Reihe interindividuell variierender Variablen konnten als stabile Korrelate sozialer Einstellungen und Prädiktoren von relevantem Verhalten etabliert werden. Für die Gegenüberstellung der nomologischen Netzwerke expliziter und impliziter Einstellungen werden in vorliegender Arbeit folgende Konstrukte ausgewählt und im Anschluss konzeptuell vorgestellt:

- *Right-Wing-Authoritarianism* (RWA, Altemeyer, 1981, 1988, 1996)
- *Soziale Dominanzorientierung* (SDO, Pratto, Sidanius, Stallworth, & Malle, 1994; Sidanius, Pratto & Bobo, 1994; Sidanius & Pratto, 1999)
- *Konservatismus* (z.B. Wilson, 1973; Riemann et al., 1993; Saucier, 2000)
- *Ambiguitätsintoleranz* (AIT, Frenkel-Brunswik, 1949, Dalbert, 1999)
- *Personal Need for Structure* (PNS, Neuberg & Newsom, 1993)
- *Need for Cognitive Closure* (NCC, Webster & Kruglanski, 1994)
- *Need for Cognition* (NFC, Caccioppo & Petty 1982)

Darüber hinaus werden einstellungsrelevante Selbstidentifikationsvariablen einbezogen, von denen eine spezifische Beziehung zu den ausgewählten Einstellungsobjekten angenommen werden kann. Diese werden jedoch nicht an dieser Stelle, sondern im Rahmen der betreffenden empirischen Untersuchungen diskutiert.

### 2.4.1.1 *Right-Wing-Authoritarianism*

Das Konstrukt des Autoritarismus geht bis zur Mitte des Zwanzigsten Jahrhunderts zurück, auf eine Forschergruppe der Frankfurter Schule um Theodor Adorno, die mit „The Authoritarian Personality“ ein bahnbrechendes Werk über die individuelle Psychologie des Faschismus und zugleich die Ergebnisse ihrer umfangreichen empirischen Untersuchungen, z.B. mit der F-Skala, veröffentlichten. Neuen Aufwind erhielt die Konzeption in den Achtziger Jahren durch den Kanadier Bob Altemeyer. Aus dem ursprünglich neun Facetten umfassenden Autoritarismus-Syndrom extrahiert Altemeyer die drei Komponenten Autoritäre Aggression, Autoritäre Submission und Konventionalismus als wesentlichen Kern und bezeichnet das daraus resultierende Konstrukt als *Right-Wing-Authoritarianism* (RWA).

Die Relevanz von RWA in der Einstellungs- und Vorurteilsforschung belegt eine große Anzahl empirischer Untersuchungen, die konsistent positive Zusammenhänge zwischen RWA und der Ablehnung sozialer Minderheiten berichten (z.B. Adorno et al., 1950; Altemeyer, 1981, 1988, 1994, 1996, 1998; Haddock, Zanna & Esses, 1993; Hassan & Khalique, 1987; Kirscht & Dillehay, 1967; Oesterreich, 1993; Rippl & Seipel, 1998; van Ijzendoorn, 1989; Zick, 1997). Für den Zusammenhang zwischen Autoritarismus und Vorurteilen gegenüber Homosexualität finden McFarland & Adelson (1996) eine Korrelation von .47, während Altemeyer (1998) einen noch stärkeren Zusammenhang von  $r=.61$  dokumentiert. Außerdem werden Beziehungen zwischen RWA und der Ablehnung von Farbigen bzw. Frauen aufgezeigt (z.B.  $r=.30$  bzw.  $r=.38$ ; Altemeyer, 1998). Als Spannweite der Korrelationsbeträge zwischen RWA und verschiedenen Vorurteilsmessungen geben Six, Wolfradt & Zick (2001)  $r=.39$  bis  $r=.66$  an. Nach Downing & Monaco (1986) weisen autoritäre Personen zudem einen höheren Ingroup-Bias bei stärkerer Ablehnung der Outgroup auf. Darüber hinaus stehen RWA und Konservatismus in positiver Beziehung zueinander (Cloetta, 1983; Eckhardt, 1991; McHoskey, 1996; Pentony et al., 2000; Wilson, 1973; Zick, 1997).

Autoritäre Personen zeichnen sich insbesondere Werte wie „Konformität“, „Tradition“, „Sicherheit“, „Macht“ und „Wohlfahrt“ aus, während sie „Selbstbestimmung“, „Stimulation“, „Universalismus“ und „Hedonismus“ weniger wichtig finden, als nicht autoritäre Personen (Altemeyer, 1998; Feather, 1996). Frindte, Funke & Jacob (1997, 1999) sowie Funke (1996) zeigen, dass insbesondere die Werte „Nationale Sicherheit“, „Soziale Macht“, „Autorität“, „Selbstdisziplin“, „Soziale Ordnung“, „Achtung vor der Tradition“, „familiäre Sicherheit“, „Selbstachtung“ und „Chancengleichheit“ zwischen hoch- und niedrig-autoritären Probanden trennen. Dementsprechend sollten hoch-autoritäre Personen weniger internal zur Vorurteilsfreiheit motiviert sein, da sich hierin eine universalistische Werthaltung widerspiegelt. Inwiefern autoritäre Personen jedoch external vorurteilsfrei motiviert sind, sollte eher mit der wahrgenommenen Gesellschaftsnorm zusammen hängen (z.B. Political Correctness).



### 2.4.1.2 Soziale Dominanzorientierung

Eine im Vergleich zu Autoritarismus relativ junge, jedoch mit RWA oft in einem Atemzug erwähnte Konzeption innerhalb der Einstellungs- und Vorurteilsforschung ist die *Soziale Dominanztheorie* (SDT) von Sidanius und Pratto (z.B. 1999). Die SDT beschreibt die Existenz sozialer Hierarchien, ihre Legitimierung und Aufrechterhaltung durch die Mitglieder einer Gesellschaft. Die Soziale Dominanzorientierung (SDO) bezeichnet entsprechend die individuelle Tendenz, die Gesellschaft als hierarchisch in soziale Gruppen gegliedert wahrzunehmen, die Gesellschaftsgruppen als unterschiedlich *wertig* zu betrachten, quasi in eine Art Rangverhältnisse zu setzen und entsprechend überlegene- den unterlegenen Gruppen gegenüberzustellen. Eine sozial dominante Person hat entsprechend das Ziel, möglichst einer dominierenden Gruppe anzugehören und deren Status zu sichern (z.B. Pratto, Sidanius & Bobo 1994; Sidanius & Pratto, 1999).

SDO steht wie RWA mit Vorurteilen in Beziehung (Sidanius & Pratto, 1999; Duckitt, 2001, Duckitt, Wagner, Du Plessis & Birum, 2002; Levin & Sidanius, 1999; Lippa & Arad, 1999; Sidanius, Pratto & Bobo, 1994; Six, Wolfradt & Zick, 2001; van Hiel & Mervielde, 2002). Altemeyer (1998) findet analog zu RWA positive Korrelationen zwischen SDO und der Ablehnung von Homosexuellen ( $r=.42$ ), Frauen ( $r=.49$ ) und Farbigen ( $r=.52$ ). In der Arbeit von McFarland & Adelson (1996) wird ein Zusammenhang von  $r=.50$  zwischen SDO und einem aggregierten Vorurteilsmaß berichtet, das negative Einstellungen gegenüber Farbigen, Frauen und Homosexuellen reflektiert. SDO steht ebenso in positiver Beziehung zu Konservatismus (Pratto et al., 1994) und in negativem Zusammenhang zu einem universalistischen und wohlfahrtsorientierten Wertesystem (Duriez & van Hiel, 2002; Heaven & Connors, 2001).

### 2.4.1.3 Konservatismus

Konservatismus stellt eine Art Grundhaltung dar, von der ausgehen verschiedene andere Einstellungen beeinflusst werden. Kennzeichen einer konservativen Grundhaltung sind nach Muller (1997) die Befürwortung von Autorität, was den Zusammenhang zu RWA erklärt, die Legitimität von Ungleichheit und die Notwendigkeit von Eliten, was seinerseits die Brücke zu SDO ist, sowie die Familie als wichtigste Sozialisationsinstanz und den Staat als Ordnungsfaktor anzusehen. Konservatismus repräsentiert bei faktorenanalytischen Untersuchungen an Variablen-Sets sozialer Einstellungen häufig den ersten extrahierten Faktor (z.B. Eysenck, 1954; Riemann et al., 1993; Saucier, 2000; Schenk, 1980; Six, Wolfradt & Zick, 2001; Wilson, 1970). Konservatismus wurde als relativ stabiles Vorurteilkorrelat mit engen Beziehungen zu Autoritarismus und Sozialer Dominanzorientierung ausgewiesen (z.B. Duckitt et al., 2002; Eckhardt, 1991; McHoskey, 1996; Oesterreich, 1993; Pentony et al., 2000; Ray & Lovejoy, 1986; Wilson, 1973; Zick, 1997).

#### 2.4.1.4 Ambiguitätsintoleranz & Konstrukte kognitiver Komplexität

In der von Else Frenkel-Brunswik (1949) stammenden Ursprungskonzeption der „Tolerance of Ambiguity“ wird dieses Merkmal u.a. als Tendenz zu „Schwarz-Weiß-Lösungen“ und vorschnellem, eingeschränkten Urteilen, als Negierung der Realität und als Streben nach unzweideutiger Akzeptanz und Ablehnung anderer Menschen definiert. Unter Ambiguitätstoleranz/ -intoleranz ist ein interindividuell variierender Indikator kognitiver Eigenschaften zu verstehen (Reis, 1979; Ruch & Hehl, 1983) und speziell die Tendenz, mehrdeutige Situationen als angenehm bzw. unangenehm zu empfinden (Budner, 1962).

Die theoretischen Überlegungen von Else Frenkel-Brunswik, dass autoritäre und vorurteilsvolle Personen Mehrdeutigkeiten versuchen zu vermeiden, konnte mittlerweile vielfach empirisch bestätigt werden (Block & Block, 1951; Durrheim, 1998; Funke, 1996; Hassan & Khaliq, 1987; Kischkel, 1983; McGregor, Zanna, Holmes & Spencer, 2001; Ruch & Hehl, 1983; Sarma, 1973; Sidanius, 1978; Sinha & Hassan, 1975; Zacker, 1973). Durrheim (1998) findet eine signifikant negative Korrelation zwischen Ambiguitätstoleranz (AT) und Rassismus ( $r = -.29$ ; vgl. Sarma, 1973; Sinha & Hassan, 1975; Hassan & Khaliq; 1987). Ambiguitätsintoleranz korreliert dagegen positiv mit RWA (Durrheim, 1998; Funke, 1996) und Konservatismus (Kischkel, 1983; Ruch & Hehl, 1983; Sidanius, 1978).

Eng mit AIT verwandt und insbesondere für die Stereotypen- und Vorurteilsforschung relevant sind weitere Konstrukte kognitiver Komplexität, wie *Personal Need for Structure* (PNS, Neberg & Newsom, 1993), *Need for Cognitive Closure* (NCC, Webster & Kruglanski, 1994) oder *Need for Cognition* (NFC, Caccioppo & Petty 1982). Diese Konstrukte lassen bereits auf theoretischer Ebene deutliche Parallelen zum komplexitätsminimierenden Charakter stereotypischer Repräsentationen erkennen. Stereotypisches Kategorisieren geht einerseits zu Lasten der Vielfalt in der realen sozialen Welt, bedeutet jedoch andererseits einen enormen Vorteil für die gesamte Informationsverarbeitung und hilft, kognitive Überlastung und Ambivalenzen zu vermeiden (vgl. Neberg & Newsom, 1993; Stangor, 2000). Für die engen Beziehungen zwischen PNS, NCC sowie, NFC und der Ambiguitätsintoleranz liegen neben konzeptuellen Anhaltspunkten auch empirische Befunde vor (Leone, 1999).

Eine Disposition für *Need for Structure* zeichnet Personen dadurch aus, ihre (soziale) Welt in weniger komplexe Kategorien zu organisieren und diese Kategorien insbesondere auch in neuen oder mehrdeutigen Situationen zu nutzen, um Sicherheit zu erlangen (Neberg & Newsom, 1993). Ein hohes Maß an *Need for Cognitive Closure* bezeichnet das Bedürfnis nach schnellen, sicheren Antworten, Entscheidungen und Beurteilungen sowie die Abneigung gegen unvorhersagbare, unsichere oder ambige Konstellationen in der Welt (Kruglanski & Webster, 1996, 2000; Webster & Kruglanski, 1994). Personen mit hoher *Need for Cognition* zeichnen sich dagegen durch eine Vorliebe für komplexe Denkprozesse aus (Caccioppo & Petty 1982). Das Konzept wurde vor allem innerhalb der Forschung zu Über-

redung und Einstellungsänderung untersucht. Hier liegt eine negative Beziehung zwischen NFC und Stereotypisierung nahe.

Inhaltlich weisen insbesondere Konzepte der AIT, PNS und NCC große Ähnlichkeiten auf. Die daraus resultierenden messmethodischen Überlappungen führten zu starker Kritik z.B. am Need-for-Closure-Konstrukt (siehe Diskurs Kruglanski et al., 1997; Neuberg, Judice & West, 1997; Neuberg, West, Thompson & Judice, 1997). Teilweise liegen zwar gemeinsame Validierungen für NFS und NCC vor (Hänze, 2002, deutschsprachige Kurzskalen; Leone, 1999), dennoch ist die Konstruktvalidität, insbesondere die *diskriminante Validität* beider Variablen nicht ausreichend geklärt.

Weniger als inhaltlich bedeutsames Korrelat sozialer Einstellungen und Vorurteile, denn vielmehr als Kontrollvariable findet die *soziale Erwünschtheit* insbesondere bei sozial sensitiven Fragestellungen wie der Untersuchung von Vorurteilen Beachtung. Soziale Erwünschtheit kann als differentielle Variable betrachtet werden und bezeichnet insofern die Tendenz positiver Selbstdarstellung (vgl. Holtgraves, 2004). Da Vorurteile, bzw. deren offene Äußerung, nicht sozial erwünscht sind, sollte zwischen selbst berichteten Vorurteilen und sozialer Erwünschtheit ein negativer Zusammenhang bestehen.

#### 2.4.2 Nomologisches Netzwerk impliziter Einstellungen

Zum nomologischen Netzwerk impliziter Einstellungen ist noch sehr wenig bekannt. Die in Abbildung 2-4 zusammengefassten Quellen impliziter Einstellungen legen nahe, dass es sich bei den Korrelaten impliziter Maße möglicherweise um andere Konstrukte als im nomologischen Netzwerk expliziter Einstellungen handelt, bzw. dass hierfür zum Teil sogar gegensätzlich gerichtete Korrelationsmuster zu erwarten wären. Konkrete Untersuchungen zu eigenständigen Prädiktoren impliziter Einstellungen liegen jedoch ebenso wenige vor, wie empirische Befunde zu den Zusammenhängen zwischen impliziten Einstellungen und etablierten Prädiktoren expliziter Einstellungen liegen jedoch kaum vor. Ein großes Problem in diesem Rahmen stellt das ebenfalls explizite Messformat der im vorangegangenen Abschnitt beschriebenen Korrelate expliziter Einstellungen dar. Bei der Überprüfung von Korrelationen zwischen expliziten Einstellungskorrelaten, wie RWA, SDO oder Konservatismus, und impliziten Einstellungen ist aufgrund der unterschiedlichen Methodenvarianz mit stark verminderten Korrelationsbeträgen zu rechnen.

Bislang können zwei Arbeiten von Rowatt und Kollegen zum nomologischen Netzwerk impliziter Einstellungen angeführt werden. Rowatt, Franklin & Cotton (2005) untersuchen u.a. RWA, SDO und soziale Erwünschtheit (Balanced Inventory of Desirable Responding von Paulhus & Reid, 1991,

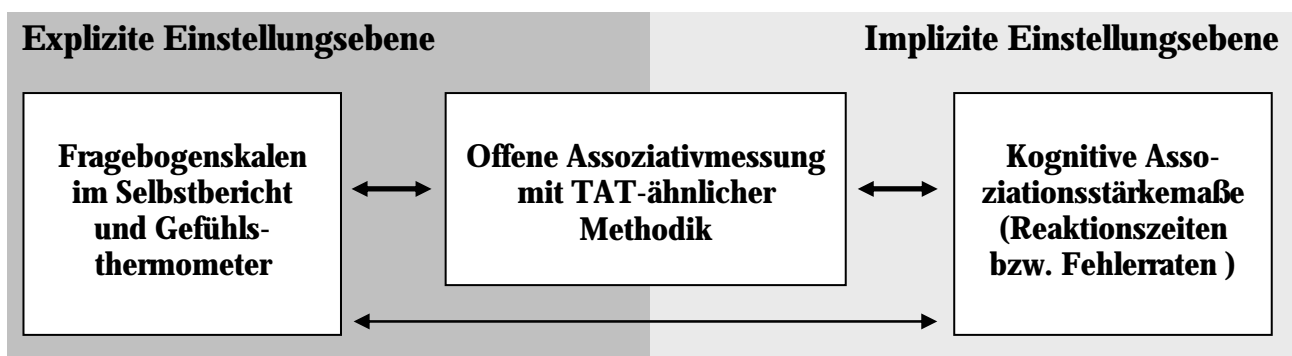
BIDR) als Korrelate expliziter und impliziter religiöser Einstellungen. Sie finden keinerlei Beziehungen zwischen den selbstberichteten Einstellungsprädiktoren und einem IAT, zur Erfassung der impliziten Assoziationsstärken von Christen vs. Moslems und den evaluativen Polen angenehm vs. unangenehm. Dagegen korreliert RWA mit der entsprechenden expliziten Messung zu  $r=.63$ . Der Zusammenhang zwischen SDO und expliziten Einstellungen ist mit  $r=.19$  schwach positiv und für die beiden Skalen des BIDR (Selbst- und Fremdtäuschung) werden keine signifikanten Korrelationen zu den expliziten Einstellungen beobachtet. In einer weiteren Untersuchung von Rowatt & Franklin (2004) zeichnet sich RWA dagegen als bester Prädiktor eines klassischen Rassen-IATs (Schwarze vs. Weiße, angenehm vs. unangenehm) aus.

Es bleibt ein prinzipiell wenig erleuchtetes und bislang noch recht heterogenes Bild festzuhalten, inwiefern klassische Einstellungsprädiktoren, also Korrelate selbstberichteter Einstellungen, in Beziehung zu impliziten Einstellungen stehen, wobei hier auch die gleichen messmethodischen Differenzen wie zwischen expliziten und impliziten Einstellungsmaßen bestehen.

### 3 Fokus vorliegender Arbeit

Die vorliegende Arbeit ist der Untersuchung der Konstruktvalidität sowie der Bedeutung expliziter und impliziter Einstellungen gegenüber zwei sozialen Minderheiten, und zwar homosexuellen und übergewichtigen Menschen, gewidmet. Eine Konstruktvalidierung im Sinne einer Multitrait-Multimethod-Analyse ist eine klassische Methode, konvergente und diskriminante Validität psychologischer Konstrukte zu untersuchen. Im vorliegenden Falle soll die Konstruktvalidität der aus expliziten und impliziten Einstellungsmessungen resultierenden Maßen überprüft werden, um zur Konzeptualisierungskontroverse um Einstellungen einen wichtigen Beitrag zu leisten. Zumeist werden direkte und indirekte Einstellungsmessungen entweder als zwei Seiten einer Medaille behandelt oder als die beiden Pole einer Implizit-Explizit-Dimension, wobei der „Zwischenraum“, das Kontinuum nicht näher beschrieben ist. Da der Grad der „Direktheit“ bzw. „Indirektheit“ einer Einstellungsmessung schwerlich als etwas Absolutes betrachtet werden kann (vgl. De Houwer, 2003b), sollte es Messungen geben, die in der Lage sind, explizite *und* implizite Einstellungsanteile zu erfassen und deren resultierende Maße einerseits zu selbstberichteten Fragebogenmaßen, andererseits zu impliziten Assoziationsmaßen wie dem IAT in Beziehung stehen (siehe Abb. 3-1). Für die Untersuchung zur Konzeptualisierung expliziter und impliziter Einstellungen wird in Anlehnung an Thematischen Apperzeptionstest (TAT, Murray, 1943) eine offene Assoziativmessung entwickelt, die sowohl direkte, als auch indirekte Messeigenschaften vereinigt. Damit werden zugleich die Möglichkeiten einer impliziten Motivmessung auch im Einstellungsbereich erforscht. Sollte es sich bei expliziten und impliziten Einstellungen um Aspekte *einer* zugrunde liegenden Disposition handeln, sollten die Fragebogenmaße, die Assoziationsstärkemaße und die offene Assoziationsmaße innerhalb eines Einstellungsinhaltes (z.B. Homosexualität) höher korrelieren, als dazwischen.

Im Folgenden werden die Grundlagen einer offenen Assoziativmessung zur Erfassung individueller Einstellungen gegenüber homosexuellen und übergewichtigen Menschen in Anlehnung an die Methodik des TAT umrissen. Im Anschluss daran erfolgt eine Betrachtung der empirischen Befunde zu den beiden ausgewählten Einstellungsobjekten.



**Abbildung 3-1.** Die hypothetische Explizit-Implizit-Dimension der Einstellungen

### 3.1 Implizite Motivmessung mit dem Thematischen Apperzeptionstest

Der im Jahre 1943 von Henry Murray publizierte Thematische Apperzeptionstest (TAT) ist eines der inzwischen meistgenutzten projektiven Verfahren in den USA (Ronan, Date & Weisbrod, 1995). Der TAT findet insbesondere im Bereich der klinischen Diagnostik Anwendung.

Das Verfahren ist charakterisiert durch spezielles Stimulusmaterial sowie Antwortformat. Als Stimulusmaterial dienen unvollständige, unstrukturierte oder mehrdeutige, in jedem Fall aber vielfältig interpretierbare Schwarzweiß-Zeichnungen sozialer Situationen, zu welchen der Proband jeweils eine Geschichte erfinden soll. Das resultierende qualitative Datenmaterial kann mittels Kodierschlüssel nach verschiedenen Gesichtspunkten ausgewertet werden. In der Originalfassung sind zwei Serien á zehn Bildtafeln zu bearbeiten. Die Versuchsperson ist angehalten zu jedem Bild eine Geschichte zu erzählen, welche aufgenommen und später transkribiert werden. Der TAT wird als operantes Messverfahren bezeichnet, da das Antwortformat spontanen und freien Charakter trägt (McClelland, Koestner und Weinberger, 1989; Scheffer, Kuhl & Eichstaedt, 2003; Zeldow & McAdams, 1993).

Es liegen bereits diverse Forschungsarbeiten vor, in denen kognitive Repräsentationen vom Selbst und von anderen mit dem TAT erfasst wurden (vgl. Rosenberg, Blatt & Oxman., 1994). Zeldow & McAdams (1993) weisen dieser Methodik ein Potential vor allem zur Messung von Motiven und fernherhin Schemata zu. Bereits an dieser Stelle fallen Parallelen zur Konzeptionalisierung von Stereotypen auf.

Verschiedene via TAT gemessene Abwehrmechanismen konnten als Indikator für Persönlichkeitsstile, den Entwicklungsstand und –Verhaltensbeobachtungen von (nicht-klinischen) Probanden in Verbindung gebracht werden (Cramer, 2001, 2002; Hibbard, Farmer & Wells, 1994). Auch kritische Lebensereignisse können sich in TAT-Geschichten widerspiegeln (Balk et al., 1998). Leigh, Westen & Barends (1992) validieren Interviewdaten und TAT-Geschichten für die Messung der Komplexität mentaler Repräsentationen von Personen. Woike & Polo (2001) können autobiografische Erinnerungen von eher einzelgängerischen vs. gemeinschaftlich orientierten Probanden mit entsprechend kongruenten impliziten Motiven im TAT (Leistung und Macht vs. Intimität) in Zusammenhang bringen. Darüber hinaus wurden diverse Hinweise auf die Güte eines TAT-Kodiersystem zur Problemlösekompetenz für klinische und nicht-klinische Probanden erbracht (Ronan, 1990; Ronan und Kollegen, 1993, 1995, 1996). Dollinger (2003) verwendete unter anderem TAT-Geschichten und Wortassoziationen, um die Bedeutung von *Need for Uniqueness* und *Need for Cognition* für Kreativität zu untersuchen. Diese Untersuchungen belegen das Potential narrativer Verfahren wie des TAT für die Erfassung *kognitiver* Repräsentationen und Variablen.

Insbesondere auch dem Bereich der Motivmessung (v.a. Leistungs-, Macht- und Anschlussmotiv) entstammen verschiedene Nachweise für die Validität des TAT (Kwon, Campbell & Williams, 2001;

McClelland, 1980; McClelland, Koesters & Weinberger, 1989; Schultheiss & Brunstein, 2001; Spangler, 1992). Darüber hinaus ist dem TAT die deutliche Überlegenheit bei der Erfassung impliziter Motive im Vergleich zum Fragebogen belegt worden (Schmalt & Sokolowski, 2000), da das Verfahren potentiell Zugang zu kognitiven Repräsentationen ermöglicht (McClelland, Koesters & Weinberger, 1989). Trash und Elliot (2002) bestätigen die moderierende Wirkung von Selbstbestimmung bei der Kongruenz von impliziten und expliziten Motiven auch für TAT-Messungen: Für selbst bestimmte Probanden finden sich positive Zusammenhänge zwischen mit TAT gemessenem und selbst berichtetem Leistungsmotiv (siehe Abschnitt 2.1.2.3.). Hierin begründet sich erneut der Wert einer TAT-analogen Messung zur Erfassung von sozialen Einstellungen für die hier verfolgten Fragestellungen.

Der TAT wird jedoch auch aus vielerlei Perspektiven kritisiert. Allem voran wird die geringe Reliabilität allgemein der projektiven Verfahren bemängelt (z.B. Sharkey & Ritzler, 1985). Zudem ist der TAT kein ausgesprochen ökonomisches Testverfahren. Für die meisten Fragestellungen ist das qualitative Datenmaterial nach festgelegten Kriterien zu kodieren, was zur Erhöhung der Verlässlichkeit auch nicht nur von einem Kodierer vorzunehmen ist. Der TAT ist eine kontextsensitive Messung (vgl. Kraiger, Hakel & Cornelius, 1984). Aus diesem Grund sind die Gütekriterien der Klassischen Testtheorie zur Evaluation psychometrischer Tests hier nicht unbedingt adäquat. Keiser & Prather (1990) führen in ihrem Review des Forschungsstandes zum TAT zudem die mangelnde Vergleichbarkeit des Stimulusmaterials als Ursache für die nach wie vor strittige Validität des TAT an (vgl. Hibbard, Mitchell & Porcerelli, 2001).

Die *Methodik* des TAT bedeutet trotz aller berechtigten Kritik ein nicht zu bestreitendes Potential zur Erfassung auch unbewusster Kognitionsinhalte; sie wird in vorliegender Arbeit genutzt, um ein weiteres Einstellungsmaß zu erhalten, das im Gegensatz zu Fragebogenskalen und impliziten Assoziationsstärkemaßen sowohl explizite, als auch implizite Anteile erfassen sollte. Da weder die originalen Bildkarten verwendet werden können, noch etablierte Kodierschemata für Leistungs-, Macht- und/ oder Affiliationsmotiv, wird die Messung nicht als TAT-analogue, sondern als „offene Assoziativmessung“ bezeichnet. Die Entwicklung der Messung wird in Studie II beschrieben (siehe Kapitel 5).

### **3.2 Die betrachteten Einstellungsobjekte**

Für eine Konstruktvalidierung gemäß einer Multitrait-Multimethod-Analyse bedarf es neben mehr als einer Messung (hier: Fragebogen, offene Assoziativmessung, IAT, EAST) auch mindestens zwei Messinhalte. Im vorliegenden Fall werden zwei soziale Einstellungsobjekte gewählt, die in unserer Gesellschaft und darüber hinaus nachweislich relevant für Vorurteile und Diskriminierung sind, und

zwar homosexuelle Menschen und übergewichtige Menschen. Insbesondere zum Stereotyp Übergewicht besteht immer noch ein relativ hoher Forschungsbedarf. Sowohl übergewichtige, als auch homosexuelle Menschen stellen eine gesellschaftliche Minderheit des öffentlichen Lebens dar, zu denen interindividuelle unterschiedliche Einstellungen zu erwarten ist.

### 3.2.1 Bedeutung der ausgewählten Stereotypen

#### 3.2.1.1 *Übergewichtige Menschen*

Übergewicht ist ein in unserer Gesellschaft stigmatisiertes Merkmal. Gemäß dem Leistungsprinzip sollen die Mitglieder der Gesellschaft nicht nur tatenkräftig, aktiv und fleißig sein, sondern auch sportlich und attraktiv. Übergewichtig bzw. nicht schlank zu sein, wird in den westlichen Industrieländern als unattraktiv empfunden und auf jede erdenkliche Art und Weise bekämpft (Diäten, Abnehmpräparate, Sport), während es im Vergleich dazu in Entwicklungsländern geradezu ein Statussymbol ist.

Vor allem Eigenschaften wie Leistungsfähigkeit, Durchhaltevermögen und Willensstärke werden übergewichtigen Menschen allein aufgrund ihrer Leibesfülle abgesprochen (siehe unten). Ebenso ein Stereotyp, wenn auch positiver Natur, ist das Bild des gemütlichen, genießerischen und lebenslustigen Dicken. Die deutliche Ablehnung von Übergewicht in unserer Gesellschaft lässt jedoch vermuten, dass übergewichtigen Menschen im Allgemeinen eher weniger positive Attribute zugeschrieben werden. Dass es sich bei den Klischees um die bequemen, maßlosen Dicken um stereotype Halbwahrheiten handelt, belegen neue Untersuchungen zur genetischen Veranlagung von Übergewicht, wie sie die Frankfurter-Rundschau vom 17.02.2005 zitiert. Es konnte gezeigt werden, dass extremes Übergewicht zu 70 Prozent auf die genetische Veranlagung zurückzuführen ist. Mittlerweile sind sechs Gen-Defekte für Adipositas bekannt, die zu unterschiedlichen Auswirkungen auf das Hungergefühl und zum anderen auf den Stoffwechsel haben.

Übergewicht ist ein Gesundheitsrisiko. Die Weltgesundheitsorganisation (WHO) hat Adipositas oder Fettsucht als Krankheit definiert. Nicht zuletzt leistet die offene Diskussion über die Belastung des Gesundheitssystems durch übergewichtige Personen einen Beitrag zur Meinungsbildung innerhalb der Gesellschaft. Eine Schlagzeile des Deutschen Ärzteblatt vom 14. Februar 2005, in dem eine amerikanische Studie aus dem *International Journal of Obesity* zitiert wird, lautet beispielsweise: „Übergewichtige belasten Krankenkassen doppelt“. Eine weithin akzeptierte Kenngröße für Über-, Normal- oder Untergewicht ist der Body-Mass-Index (BMI), der das Verhältnis aus Körpergewicht in *kg* zu quadrierter Körpergröße in *m* angibt (James, Ferro-Luzzi & Waterlow, 1988). Von Übergewicht wird ab einem BMI von mehr als 25 gesprochen, von starkem Übergewicht, Adipositas oder auch krankhaftem Übergewicht ab einem BMI von 30 und von massiver Adipositas, wenn der BMI größer als 40 ist.



Übergewicht spielt in den Industrieländern mehr und mehr eine Rolle. Das Deutsche Ärzteblatt (100, 2003) schreibt. „Die Anzahl übergewichtiger Personen hat sich in den letzten Jahren in allen OECD-Ländern erhöht. Mit einem Anteil von 11,5 Prozent befindet sich Deutschland hierbei im unteren Drittel.“ Laut Statistischem Bundesamt (2004) sind 13% der Bevölkerung stark übergewichtig, wobei die Häufigkeit von Übergewicht unter Männern etwas höher ist (14%), als unter Frauen (12%). Etwa ein Prozent der Bevölkerung ist mit einem BMI über 40 massiv übergewichtig.

### 3.2.1.2 *Homosexuelle Menschen*

Unter Homosexualität wird eine gleichgeschlechtliche sexuelle Orientierung verstanden. Verschiedene Nachschlagewerke definieren „Homosexualität“ wie folgt:

*„gleichgeschlechtliche, sexuelle Anziehung und oder Bestätigung, häufige Form der Abweichung von der sexuellen Norm“ (Brockhaus Enzyklopädie)*

*„sexueller Antrieb, der vorwiegend auf Artgenossen gerichtet ist, die demselben Geschlecht angehören“ (Meyers Enzyklopädisches Lexikon, 1971-1981)*

*„sich auf das eigene Geschlecht richtendes sexuelles Empfinden und Verhalten“ (Duden, 2005)*

*„Homosexualität bezeichnet eine sexuelle Orientierung, bei der romantische Liebe und sexuelles Begehren ausschließlich oder vorwiegend gegenüber Personen gleichen Geschlechts empfunden werden.“ (Wikipedia, freie Enzyklopädie)*

Der Begriff geht auf das griechische Wort *homos* für „gleich“ bzw. „gleichartig“ und die lateinische Bezeichnung *sexus* für „Geschlecht“ zurück. Gleichgeschlechtliche Liebe hat es zu allen Zeiten der Menschheit gegeben. Die begriffliche Kennzeichnung erfolgte erst relativ spät; 1869 prägte der Psychiater Karl Maria Kertbeny den Begriff Homosexualität. Die Unterscheidung zwischen hetero- und homosexuellen Personen kennzeichnet Gesellschaften mit einem *Bedürfnis* zur Unterscheidung aufgrund starker Geschlechtsrollentrennung und rassistischer Vorurteile (Boswell, 1980). Homosexualität wurde lange Zeit als Krankheit betrachtet. Erst 1993 strich die WHO Homosexualität aus ihrer Klassifikation.

Homosexualität wurde in der Antike nicht als moralisch verwerflich betrachtet, in der Gegenwart dagegen schon. Die gleichgeschlechtliche Liebe steht den christlichen Konventionen, von denen unsere abendländische Gesellschaft geprägt ist, deutlich entgegen, sie bedroht in den Augen vieler religiöser und politischer Verantwortungsträger die kirchliche und nach wie vor auch staatliche Exklusivinstitution der Ehe und Familie und nicht zuletzt eine sittlich-konservative Weltanschauung.

Der Umgang mit Homosexualität variiert von Kultur zu Kultur, während homosexuellen Menschen in einigen Ländern noch mit Verfolgung, Bestrafung, ja gar Todesstrafe rechnen müssen, werden sie in

anderen Ländern geduldet und in einigen, wie beispielsweise den Niederlanden, akzeptiert, was sich deutlich an den gesetzlichen Bestimmungen ablesen lässt. Seit 2001 gibt es im deutschen Recht die eingetragene Lebenspartnerschaft (BGB, Schönfelder), womit gleichgeschlechtlichen Partnerschaften jedoch nicht viel mehr als eine rechtliche Zurkenntnisnahme gewährt wird. Nach wie sind viele Rechte ausschließlich der Ehe vorbehalten, wie die finanziellen Steuer-, Renten- und Krankenversicherungsvergünstigungen, das Recht auf Asyl oder Staatsbürgerschaft für einen ausländischen Partner oder die Adoption von Kindern. Trotz der Einwände insbesondere der Unionsparteien CDU und CSU ist eine Erweiterung der Rechte homosexueller Paare abzusehen, so traten am 1. Januar 2005 das lange umkämpfte Lebenspartnerschaftsgesetzüberarbeitungsgesetz in Kraft. Dennoch sind die Hürden einer Gleichberechtigung nicht nur in gesetzlichen Fragen teilweise noch unüberbrückbar.

Nach wie vor ist Homosexualität mit einem sozialen Stigma behaftet. Insofern sind Prävalenzschätzungen nicht unbedingt aussagekräftig, zudem ist die Befundlage durch ausgesprochene Variabilität gekennzeichnet. Schmidt (1995) berichtet von nicht mehr als 6% homosexueller Männer und 3% homosexueller Frauen. Einer repräsentativen Umfrage von Steffens & Wagner (2004) zufolge, bei der nur 40.7% der befragten Personen überhaupt bereit waren (N=14 942), ihre sexuelle Orientierung offen zu legen, gaben 0.6% der Frauen und 1.3% der Männer an, homosexuell zu sein. Weitere 2.5% der Frauen und 2.8% der Männer bezeichneten sich als bisexuell orientiert.

### 3.2.2 Einstellungen zu den ausgewählten Stereotypen

#### 3.2.2.1 Übergewichtige Menschen

Übergewichtige Menschen als stereotypisierte und diskriminierte soziale Gruppe finden innerhalb der Vorurteilsforschung erst nach und nach mehr Beachtung, dabei handelt es sich bei Übergewicht um ein eindeutig stigmatisiertes Merkmal, zumindest in den westlichen Industrienationen. Crandall & Biernat (1990, 227) schreiben:

*The obese in this country are a strongly disliked group; being fat is associated with more negative characteristics than nearly any other stigma (Allon, 1982). Whether this prejudice stems from viewing the obese as aesthetically, displeasing, morally and emotionally impaired, or socially handicapped, anti-fat attitudes seem to be widely expressed."*

Die Autoren umschreiben die Einstellungen gegenüber übergewichtigen Menschen als breites soziales Phänomen gleichsam unter Männern und Frauen, in allen Altersgruppen und in verschiedenen sozialen und nationalen Gruppen. Arbeitgeber stellen ungern übergewichtige Bewerber ein, selbst Mediziner haben eine Reihe an Vorurteilen gegenüber Übergewichtigen und empfinden sie als eine der unan-

genehmsten Patientengruppe. Im Kollegium werden dicke Menschen häufiger Opfer von Mobbing, als schlanke (Rothblum, Brand, Miller & Oetjen, 1990). Hebl & Mannix (2003) beobachten in diesem Zusammenhang sogar Halo-Effekte: Männliche Bewerber wurden vergleichsweise negativer bewertet, wenn zusammen mit einer übergewichtigen Frau gesehen wurden.

Grover, Keel & Mitchell (2003, 125) stellen fest:

*“Within our society, fat individuals are judged to be less attractive, less intelligent, less competent, less disciplined, and more lazy compared with thin individuals.”*

Innerhalb der Übergewichtigen sind Frauen noch stärker stigmatisiert, als Männer (Millman, 1980) und sie leiden stärker darunter (Harris, Walters & Waschull, 1991; Wooley, Wooley & Dyrenforth, 1979). Insbesondere dicke Frauen haben einen sehr geringen Selbstwert (Crandall & Biernat, 1990; Grover, Keel & Mitchell, 2003). Es sind jedoch auch die Frauen im Allgemeinen, die negativere Einstellungen zu Übergewicht besitzen (Crandall & Biernat, 1990; Tiggemann & Rothblum, 1988) und sich größere Sorgen um ihr Gewicht machen (Harris & Foltz, 1999). Die hohe Prävalenz von Essstörungen z.T. bereits in den Schulklassen und immer noch vorzugsweise unter Mädchen und Frauen veranschaulicht, welche Folgen diese gesellschaftliche Schlankkeitsnorm und die Stigmatisierung von Personen, die von dieser Norm abweichen, haben können.

Als Kernproblem der Stigmatisierung übergewichtiger Menschen führen Crandall & Biernat (1990) sowie Crandall & Martinez (1996) an, dass das Körpergewicht als Merkmal, welches unter eigener Kontrolle ist, angesehen wird und übergewichtige Menschen also selbst für ihre Körperfülle verantwortlich gemacht werden (vgl. DeJong, 1980). Dicken Frauen wird stereotyp unterstellt, zuviel zu essen; dicke Männer würden dagegen zu wenig Sport treiben (Weiner, 1986). Dabei unterscheiden sich dicke und dünne Menschen nicht in der Auffassung von der willentlichen Steuerbarkeit der Körperfettmasse (Mackenzie, 1984). Crandall et al. (2001) weisen in sechs verschiedenen Nationen nach, dass neben einer kulturell negativ geprägten Grundhaltung zu Fettleibigkeit die Attribution der Verantwortung auf die Übergewichtigen selbst für die Vorhersage von Vorurteilen gegenüber dicken Menschen bedeutsam ist. Wie oben bereits berichtet handelt es sich dabei jedoch um ein Vorurteil, denn Übergewicht ist einerseits genetisch bedingt, wie eine Adoptionsstudie von Stunkard et al. (1986) zeigen konnte und andererseits auf physiologische Unterschiede zurückzuführen. Den generell angenommen Zusammenhang von Übergewicht und Gefräßig-Seins konnte ein Review-Untersuchung von Garrow (1974) eindrucksvoll widerlegen, der in nahezu allen Studien keine Unterschiede in der Kalorienzufuhr zwischen normal- und übergewichtigen Probanden feststellte. Die übergewichtigen Untersuchungsteilnehmer nahmen zum Teil sogar eine geringere Kalorienmenge zu sich.

Negative Einstellungen zu Übergewicht basieren nach Crandall & Biernat (1990) vor allem in ideologischen Überzeugungen. Die Autoren finden sowohl für männliche, als auch für weibliche Probanden

signifikant positive Zusammenhänge zwischen Vorurteilen gegenüber dicken Menschen und politischem Konservatismus (jeweils .10), symbolischem Rassismus (Männer: .20, Frauen: .24) und Right-Wing Autoritarismus (Männer: .68, Frauen: .36).

Crandall (1994) findet keinen Ingroup-Bias unter übergewichtigen Menschen. Niemand, auch eine tatsächlich übergewichtige Person, möchte zur Gruppe der „Dicken“ gehören. Dass das Merkmal körperlicher Fettanteil relativ normal verteilt ist, somit immer noch dickere Menschen vorstellbar sind und der BMI, das Klassifizierungsmaß für die Körperfettmasse, einer Person nicht an der Stirn ablesen werden kann, eröffnet den notwendige Interpretationsspielraum. Dicke Personen tragen ebenso Vorurteile gegenüber Übergewicht, wie schlanke Personen. Diesen Befund stützen auch die Ergebnisse von Grover, Keel & Mitchell (2003), die sowohl in expliziten, als auch in impliziten Einstellungen (IAT, verbale Stimuli) zu Übergewicht keine Unterschiede zwischen normal- und übergewichtigen Personen finden. In beiden Messungen finden die Autoren negative Einstellungen zu Übergewicht, es besteht jedoch keinerlei Korrespondenz zwischen den explizite und den impliziten Einstellungen. Chambliss, Finley & Blair (2003) finden eine negative Korrelation zwischen einem Paper-Pencil-IAT (dicke vs. dünne Menschen, faul vs. motiviert, verbale Stimuli) und der Differenz der entsprechenden semantischen Differentiale (faul vs. motiviert) für dicke und dünne Menschen von -.20. Zwischen anderen impliziten und expliziten Einstellungsmaßen (gut vs. schlecht, statt faul vs. motiviert) bestand kein Zusammenhang.

Teachman et al. (2003) finden eine starke Ablehnung übergewichtiger Menschen im IAT, jedoch keine negativen Einstellungen in der expliziten Einstellungsmessung. Manipulationen zur Attribution der Verantwortung für Übergewicht bewirkten Veränderungen in den impliziten Einstellungen in erwartungsgemäßer Richtung. Informationen, dass übermäßiges Essen die hauptsächliche Ursache für Übergewicht ist, steigerte den impliziten Bias, während Hinweise, dass genetische Dispositionen hauptverantwortlich für Übergewicht sind, die implizite Ablehnung dicker Menschen abschwächte.

Bessenoff & Sherman (2000) finden keinen Zusammenhang zwischen einer Primingmessung, in der Bilder von normal- und übergewichtigen Personen die Primes darstellten und Adjektive bzw. Substantive, die nach lexikalischer Kategorie zu klassifizierenden Target-Stimuli, und expliziten Einstellungen zu Übergewicht. Dagegen bestehen signifikante Zusammenhänge zwischen der automatischen Bewertung normal- und übergewichtiger Personen im Priming und einer Verhaltensmessung ( $r=.19$ ). Je stärker die automatisch negative Aktivierung auf übergewichtige im Vergleich zu normalgewichtigen Stimuli, desto größer war die selbst gewählte Distanz des eigenen Stuhls zu einem übergewichtigen Konföderierten.

### 3.2.2.2 *Homosexuelle Menschen*

Im Jahre 1969 tauchte zum ersten Mal der Begriff „Homophobie“, eine negative Haltung homosexuellen Menschen gegenüber kennzeichnend, in den Printmedien auf (Weinberg, 1972, zitiert nach Herek, 2000). Die prototypischen Homosexuellen sind offensichtlich schwule Männer; in einem Interview nahmen 75% der Probanden an, alle Homosexuellen seien männlich (Haddock, Zanna & Esses, 1993). Herek (1994) dokumentiert einer repräsentativen Umfrage in den USA gemäß, dass mehr als die Hälfte heterosexueller Befragter schwule Männer oder lesbische Frauen Ekel erregend finden. Zwar zeigen Gefühlsthermometermessungen von 1984 bis 1996 eine konstante Verbesserung der Einstellungen zu homosexuellen Personen an (Yang, 1997), dennoch verzeichnet Herek (2000) nach wie vor für homosexuelle Personen einen Rang unter den am schlechtesten bewerteten Gesellschaftsgruppen in Amerika. Zwar finden Hebl, Foster, Mannix & Dovidio (2002) keine formale Diskriminierung von homosexuellen Bewerbern (Konföderierte) durch die Arbeitgeber, jedoch wird ein deutlich negativeres Sozialverhalten seitens der Arbeitgeber beobachtet.

Allgemein besteht ein deutlicher Unterschied in der Bewertung homosexueller Männer und Frauen. In der Regel werden schwule Männer negativer bewertet, als lesbische Frauen (Herek & Capitanio, 1996; Kite & Whitley, 1998; LaMar & Kite, 1998; Steffens & Wagner, 2004). Außerdem haben heterosexuelle Männer im Vergleich zu Frauen sehr viel negativere Einstellungen gegenüber Homosexualität (Hayes, 1995; Kite & Whitley, 1998). Damit wäre bereits das erste stabile Korrelat von Vorurteilen gegenüber Homosexualität genannt: Das Geschlecht. In verschiedenen Untersuchungen konnten Prädiktoren für die Einstellungen gegenüber Homosexuellen etabliert werden, die sowohl demografischer, als auch eher ideologischer Natur sind. Dass Männer Homosexualität negativer gegenüber stehen wird auf die unterschiedlichen Geschlechtsrollenüberzeugungen von Männern und Frauen zurückgeführt.

Andere relevante demografische Variablen sind u.a. das Alter, der Bildungsstand, der sozioökonomische Status (SES), wobei ältere, weniger gebildete Personen und Personen mit geringerem SES stärkere Vorurteile haben (Hayes, 1995; Herek, 1994; Kämpfe, 2002; Morrison, Parriag & Morrison, 1999). Allgemeinere soziale Einstellungsdispositionen stehen ebenso mit den Einstellungen zu Homosexualität in Beziehung. Sherrod & Nardi (1998) zeigen die positive Beziehung zwischen Konservatismus und der Ablehnung homosexueller Personen. Altemeyer (1998) zeigt in diesem Rahmen zudem die Bedeutung von Autoritarismus und sozialer Dominanzorientierung auf (siehe oben; vgl. Kämpfe (2002)). Ein ebenfalls bedeutsamer Aspekt ist die Bekanntschaft zu einem oder mehreren homosexuellen Personen. Probanden, die angeben, eine homosexuelle Person in ihrem Freundeskreis oder ihrer Familie zu haben, weisen deutlich positivere Einstellungen gegenüber Homosexualität im Allgemeinen auf (Herek & Capitanio, 1996; Kämpfe, 2002). Steffens & Wagner (2004) weisen das Alter, die persönlichen Kontakterfahrungen, die gleich- und gegengeschlechtliche sexuelle Anziehung,

die Parteienpräferenz sowie das Geschlecht als beste Prädiktoren von Einstellungen zu Homosexualität aus.

Jellison, McConnell & Gabriel (2004) setzen einen IAT mit bildhaften Stimuli romantischer Paare zur Messung automatischer Assoziationen zu Homosexualität relativ zu Heterosexualität ein. Sie berichten Korrelationen von .31 bis .51 zwischen dem IAT-Effekt und verschiedenen expliziten Einstellungsmessungen (u.a. Gefühlsthermometer) zur sexuellen Orientierung in einer männlichen Stichprobe. Sie finden außerdem einen Zusammenhang von .29 zwischen der Motivation, keine Vorurteile gegenüber Homosexuellen zu haben und expliziten Einstellungen, nicht jedoch impliziten Einstellungen.

Banse, Seise und Zerbes (2001) testen einen IAT Homosexualität, ebenfalls mit bildhaften Stimuli, im deutschsprachigen Raum mit einer Stichprobe heterosexueller und homosexueller Männern und Frauen. Sie replizieren zunächst den bereits berichteten Befund positiverer Einstellungen zu Homosexualität unter Frauen im Vergleich zu Männern, und zwar in den expliziten und impliziten Messungen. Zudem haben homosexuelle Probanden deutlich positivere Einstellungen zu Homosexualität, was sich ebenfalls in allen Maßen niederschlägt. Die Autoren dokumentieren einen Zusammenhang von  $r=.52$  zwischen dem IAT und den aggregierten Skalen kognitive und affektive Einstellungen zu Homosexualität über alle Probanden hinweg. Die statistische Kontrolle von sexueller Orientierung und Geschlecht resultiert in einem Absinken der Korrelation auf .29.

Für die Messung von Einstellungen zu Homosexualität ist zu beachten, dass männliche und weibliche Homosexualität unterschiedliche Einstellungen evoziert (Herek & Capitano, 1996; LaMar & Kite, 1998; Kite & Whitley, 1998; Steffens & Wagner, 2004). Während die Gruppe schwuler Männer einen etablierten sozialen Stereotyp darstellt, sind die Stereotypeninhalte für die Gruppe lesbischer Frauen diverser und weniger sozial geteilt (Morrison, McLeod, Morrison, Anderson, & O'Connor, 1997; Steffens, in press; Steffens & Wagner, 2004). Dies hängt möglicherweise mit den gesellschaftlich strengeren Normen hinsichtlich Nähe und Intimität zwischen Männern, als zwischen Frauen zusammen.

### **3.3 Der empirische Teil vorliegender Arbeit**

#### **3.3.1 Aufbau**

Vorliegende Arbeit zur Konstruktvalidität von Einstellungen umfasst drei empirische Untersuchungen. Die ersten beiden Studien sind der Erprobung bzw. Entwicklung ausgewählter Methoden zur Erfassung expliziter und impliziter Einstellungen gewidmet. In Studie I werden der Implizite

Assoziationstest und der Extrinsic Affective Simon Task zur Messung impliziter Einstellungen zu Homosexualität getestet. Studie II umfasst die Entwicklung einer offenen Assoziativmessung in Anlehnung an den Thematischen Apperzeptionstest, mit der sowohl implizite, als auch explizite Einstellungs- und Motivationsaspekte erfasst werden können. Studie III fokussiert schließlich die globalen Fragestellungen der Konstruktvalidität expliziter und impliziter Einstellungen vorliegender Arbeit.

### 3.3.2 Zeichenlegende

Die Dokumentation der empirischen Befunde von Studie I bis III erfolgt in einem einheitlichen statistischen Beschriftungssystem, das in folgender Legende beschrieben wird:

<b>Beschriftung</b>	<b>Bedeutung</b>
$\alpha$	Cronbach's Alpha
$\Delta$	(Mittelwerts-) Differenzen
$\chi^2$	Prüfgröße des Chi <sup>2</sup> -Test
<i>Beta</i>	Standardisiertes Regressionsgewicht
<i>df</i>	Freiheitsgrade
<i>F</i>	Prüfgröße des F-Test
<i>M</i>	Mittelwert
<i>Max</i>	Maximum
<i>Min</i>	Minimum
<i>NFI</i>	Normed Fit Index
<i>p</i>	Wahrscheinlichkeitsniveau (p-Wert)
	<b><i>p</i></b> <.01
	<i>p</i> p<.05
	<i>p</i> nicht signifikant
<i>r</i>	Produkt-Moment-Korrelation
<i>R<sup>2</sup></i>	Varianzaufklärung
<i>RMSEA</i>	Root Mean Square of Error Approximation
<i>SD</i>	Standardabweichung
<i>SE</i>	Standardfehler
<i>T</i>	Prüfgröße des T-Test

**Tabelle 3-1,** Zeichenlegende der empirischen Datendokumentation

**- Empirischer Teil -**



## **4 Studie I : Auswirkungen einer fixierten Stimulusabfolge auf die Reliabilität und Validität von IAT und EAST**

### **4.1 Theoretische Grundlagen und Anliegen von Studie I**

Der IAT-Ablauf ist weitestgehend standardisiert. Dennoch können verschiedene Variationsmöglichkeiten angeführt werden. An dieser Stelle werden die Hauptvariationen der Blockreihenfolge, der Stimulusabfolge sowie der Antworttastenbelegung betrachtet.

Im klassischen IAT erfolgt der kompatible- vor dem inkompatiblen Block. Die ist auch in der Mehrzahl an Untersuchungen der Fall. Kritisch daran ist das Zusammenfallen von Effekten aufgrund der eigentlich interessierenden Einstellungsinkompatibilität mit den Task-Switch-Kosten aufgrund des notwendigen Umlernens der Tastenbelegung sowie Ermüdungserscheinungen. Dies könnte die Reaktionszeiten und Fehlerraten im inkompatiblen Block künstlich erhöhen, was dann fälschlicherweise als Zeichen für eine geringere kognitive Assoziationsstärke zwischen den einstellungsinkongruent kombinierten Konzepten (z.B. „Homosexuell“ und „angenehm“ vs. „Heterosexuell“ und „unangenehm“) im Vergleich zu den kongruent kombinierten Konzepten im kompatiblen Block interpretiert werden würde. Den Untersuchungen zur Auswirkung der Blockreihenfolge gemäß lässt sich entsprechend eine Bedeutung für die Höhe des IAT-Effekts festhalten. Greenwald, McGhee & Schwartz (1998) finden etwa doppelt so hohe IAT-Effekte, wenn die kompatible Aufgabe der inkompatiblen voran geht. Schmukle & Egloff (DGPs, 2004) dagegen variieren die Reihenfolge der Blöcke ohne bedeutsame Auswirkungen auf die Ergebnisse festzustellen. Von einem rein betragsmäßigen Unterschied im IAT-Effekt sollten keinerlei Veränderungen im Korrelationsmuster der Maße resultieren. Prinzipiell sind jedoch höhere Effektstärken wünschenswert, da sie die natürliche Variationsbreite der Maße besser widerspiegeln. Einschränkungen in der Merkmalsvarianz führen zur Unterschätzung der Populationskorrelation (vgl. Bortz, 2004).

Eine weitere Variationsmöglichkeit im Aufbau des IAT und prinzipiell aller impliziten Maße betrifft die Anzahl an Messblöcken. Der klassische IAT besteht aus insgesamt fünf Blöcken, drei Übungsblöcken und jeweils einem gemischten kompatiblen und inkompatiblen Messblock. Greenwald und Kollegen setzen später einen IAT mit jeweils zwei gemischten kompatiblen und inkompatiblen Messblöcken ein (z.B. Greenwald, Nosek & Banaji, 2003), woraus ein IAT mit sieben Blöcken resultiert. Die zusätzliche Bearbeitung von zwei Messblöcken dient einerseits zur Übung der kompatiblen und inkompatiblen gemischten Aufgaben und liefert andererseits einen weiteren IAT-Effekt.

Für die Tastenbelegung existieren ebenfalls mehrere theoretische Variationsmöglichkeiten. So kann zwischen den Versuchsteilnehmern variiert werden, mit welcher Taste bspw. die angenehmen und mit welcher die unangenehmen Stimuli zu beantworten sind. Eine systematische Überprüfung der Ant-

worttastenbelegung liegt nach Wissen der Autorin nicht vor, da von der Zuordnung der Stimuluskategorien zu den Antworttasten offensichtlich kein bedeutsamer Einfluss auf die resultierenden Effekte zu erwarten ist.

Eine bereits kurz erwähnte Ablaufvariation, die für die Reliabilitätsproblematik von großer Bedeutung sein kann, betrifft die Darbietung der Stimuli in randomisierter vs. fixierter Reihenfolge (vgl. Abschnitt 1.6.3). Zur Interpretation von Haupteffekten in sozialpsychologischen Experimentalstudien neben Zwischensubjektvariationen der Blockreihenfolge und Tastenbelegung eingeführt, die aufgrund der daraus resultierenden ungleichen Bedingungen für eine Untersuchung interindividueller Differenzen nicht zweckmäßig sind (vgl. Teige, Schnabel, Banse & Asendorpf, 2004). Die durch Randomisierung der Trialabfolge und durch Variation des Experimentalaufbaus sowie der Tastenbelegung u.ä. erzielte Verallgemeinerbarkeit der Effekte für eine Stichprobe, resultiert in einer mangelnden Vergleichbarkeit individueller Effekte. Die prozeduralen Variationen führen zu einer allgemeinen Erhöhung der Varianz der Messwerte. Verschiedene Autoren (z.B. Banse, 2001; De Houwer, 2003a; Teige et al., 2004) raten zugunsten der Reliabilität sowie der Untersuchung differentieller Fragestellungen zu einer Konstanthaltung der experimentellen Bedingungen. Der Frage, ob sich eine feste Stimulussequenz allgemein begünstigend auf die Verfahrensreliabilitäten von IAT und EAST und in Folge dessen auch auf ihr Zusammenhangsmuster auswirkt, wird in Studie II nachgegangen.

Trotz der weiten Verbreitung reaktionszeitbasierter Messverfahren wie dem IAT besteht nach wie vor entscheidender Klärungsbedarf insbesondere bezüglich der Gütekriterien der Maße. Studie I ist angelegt, die Auswirkungen einer fixierten Stimulusabfolge für alle Probanden auf die Reliabilität von IAT und EAST für die Zusammenhänge zwischen den Maßen beider Verfahren sowie ihre Beziehungen zu expliziten Maßen zu untersuchen. Besonderer Fokus liegt dabei auf der Entwicklung eines reliablen EAST, da dieses Verfahren im Gegensatz zum IAT *nicht* als etabliert bezeichnet werden kann (vgl. Teige et al., 2004). Zu diesem Zweck werden in einer experimentellen Studie Einstellungen gegenüber Homosexualität mittels IAT, EAST und expliziter Selbstberichtsmessung erfasst. Da für die implizite Einstellungsmessung zu Übergewicht, dem zweiten ausgewählten Einstellungsobjekt in vorliegender Arbeit, bereits elaboriertes Bildmaterial sowie ein entsprechender EAST vorliegen (Degner, in prep.), wird an dieser Stelle die notwendige Entwicklung eines Homosexualitäts-EAST fokussiert. Dabei wird eine Einschränkung auf männliche Homosexualität vorgenommen. Dies rechtfertigt die Betrachtung der bisherigen Befundlage zu Einstellungen gegenüber Homosexualität, deren zufolge für die überwältigende Mehrheit der Befragten der Prototyp für Homosexualität *männliche Homosexualität* darstellt (vgl. Steffens & Wagner, 2004). Darüber hinaus werden mit männlicher und weiblicher Homosexualität prinzipiell sehr unterschiedliche Einstellungen verbunden (vgl. Abschnitt 3.2.2.2). Eine gemeinsame Untersuchung männlicher und weiblicher Homosexualität wäre nur bei getrennter Betrachtung und

entsprechend doppelter Erhebung der Einstellungsmaße zu rechtfertigen. Aufgrund der mangelnden Konsistenz des Stereotypen „weibliche Homosexualität“ (vgl. Herek, 1991; Kite, 1984), wird auf die Betrachtung dieses Einstellungsobjektes verzichtet.

## 4.2 Methode

### 4.2.1 Stichprobe

Die Untersuchung wurde im März 2004 an der Friedrich-Schiller-Universität Jena durchgeführt. An dem Experiment nahmen 51 studentische Versuchspersonen teil, die mit einer Versuchspersonenstunde entlohnt wurden. Ein Proband musste aufgrund eines partiellen Datenverlustes von den Analysen ausgeschlossen werden. Von den verbleibenden 50 Probanden sind  $N=43$  weiblich (86%). Das Alter der Untersuchungsteilnehmer rangiert von 19-30 Jahre und beträgt im Mittel 21.6 Jahre ( $SD=2.62$ , vgl. Abb. 4-1). Bis auf vier Probanden, die angaben, einen höheren Bildungsabschluss (Diplom, Magister) zu besitzen, hatten alle Personen die Allgemeine Hochschulreife. 22 Probanden (44%) sind in Thüringen aufgewachsen.

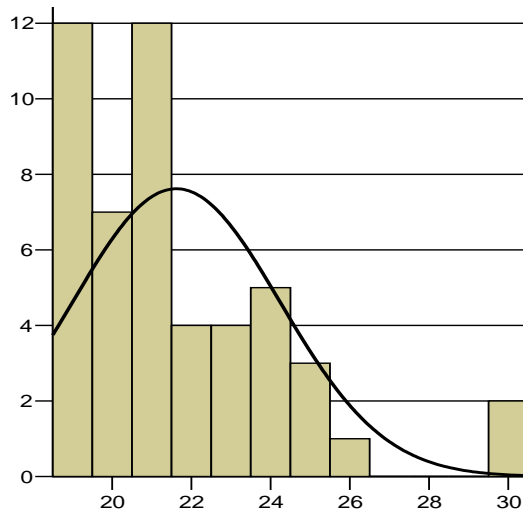


Abbildung 4-1; Altersverteilung

Zwei Teilnehmer gaben an, bisexuell orientiert zu sein, die überwiegend Zahl bezeichnete sich als heterosexuell orientiert. 12 Personen haben keine homosexuelle Person in ihrem Freundes- und Bekanntenkreis, 35 Personen haben einen bis vier homosexuelle Bekannte und drei Personen gaben an, 5 bis 9 Homosexuelle Freunde/innen zu haben.

Etwas mehr als die Hälfte der Probanden ( $N=29$ ) hatte zuvor noch an keinem Laborexperiment teilgenommen.

### 4.2.2 Material

#### 4.2.2.1 Implizite Maße

Das evaluative Stimulusmaterial für die impliziten Messungen stellen 6 männliche und 6 weibliche Portraitfotografien dar, wovon jeweils eine Hälfte einen positiven und die andere einen negativen Ge-

sichtsausdruck hat. Diese Bilder wurden aus der Datenbank für Emotionsausdrücke *The Karolinska Directed Emotional Faces* (Lundqvist, Flykt, & Öhman, 1998) ausgewählt. Für die Klassifizierungsaufgabe zum Zielkonzept „Homosexuell vs. Heterosexuell“ werden 12 Paarfotos verwendet; jeweils sechs Stimuli bilden ein homosexuelles- bzw. ein heterosexuelle Pärchen ab. Die Bilder stellen jeweils romantische Paare dar und entstammen einer Grundgesamtheit von jeweils 46 heterosexuellen und homosexuellen Paarbildern, die einem Vortest zufolge ausgewählt wurden.

Im Vortest der Stimuli für das Einstellungsobjekt Homosexualität und das Vergleichskonzept Heterosexualität wurden N=30 Probanden dazu aufgefordert, die insgesamt 92 randomisiert dargebotenen Paarbilder schnellstmöglich mittels Tastendruck einer von zwei Kategorien zuzuordnen. Die Kategorien waren mit „gleichgeschlechtlich“ und „gemischtgeschlechtlich“ bezeichnet. Die Kategorisierungsaufgabe lieferte die ersten beiden Kriterien für die Stimulusauswahl: die Reaktionszeit für die Zuordnung und die Korrektheit. Im Anschluss an die Kategorisierungsaufgabe wurden die Bilder erneut dargeboten und nach Prototypikalität für homosexuelle bzw. heterosexuelle Paare, nach Sympathie und nach dem romantischen Eindruck, den sie erzeugten, jeweils auf einer 7-Punkte-Skala eingeschätzt. Zur Auswahl geeigneter Stimuli wurden zunächst Rangreihen der Geschwindigkeit-mal-Korrektheitsergebnisse für die Einordnung der Paarbilder in die Kategorie Homosexuell bzw. Heterosexuelle gebildet. In einem zweiten Schritt wurden die gleich- und gemischtgeschlechtlichen Paarbilder hinsichtlich der Prototypikalitäts-, Sympathie- und Romantik-Ratings gematcht.

Für den EAST Homosexualität wurden alle 12 ausgewählten Stimuli des Zielkonzepts in einem Grafikbearbeitungsprogramm (ArcSoft Photo Studio 5) mit jeweils einer bläulichen und einer grünlichen Farbschicht überzogen. Die Farben werden jeweils nach den Originalangaben von De Houwer (2003) erstellt; die bläuliche Farbschicht setzt sich aus folgenden Farbwerten zusammen: Magenta: 0, Grün: 125, Blau: 150; die grünliche Farbschicht dagegen wie folgt: Magenta: 0, Grün: 150, Blau: 125. Das resultierende Zielkonzept-Stimulusmaterial für den EAST stellen demzufolge nach 24 Paarbilder dar, wobei jedes Paarbild einmal bläulich und einmal grünlich eingefärbt ist. Die beiden Farben sind relativ ähnlich gehalten, um sicher zu stellen, dass die Probanden genau hinsehen müssen, um die Zielkonzeptstimuli aufgrund der Farbe zu beantworten. Wären die beiden Oberflächenmerkmale auch ohne Aufmerksamkeitsaufwand eindeutig unterscheidbar, wäre die notwendige inhaltliche Verarbeitung der Paarbilder gefährdet. Eine Aufführung aller verwendeten Stimuli befindet sich im Anhang.

Der Bildschirmhintergrund für beide implizite Messungen ist schwarz (Farbwerte: 0, 0, 0). Die Instruktionen sowie die Markierung der Tastenbelegung während der Blöcke wird in Weiß dargeboten (Farbwerte: 255, 255, 255). Der Aufbau der einzelnen Trials sieht ein Intertrialintervall von 1200 ms vor. Entsprechend der Monitor-Refresh-Rate erfolgt für jeweils 259 ms ein Fixationskreuz in der Mitte des Bildschirms, das mit dem Stimulus ersetzt wird, der wiederum nach 541 ms schwarz überblendet wird.

#### 4.2.2.1.1 Aufbau des IAT

Der eingesetzte IAT umfasst 5 Blöcke. Als Antworttasten fungieren das [A] und die [5] im Nummernblock. Über das gesamte Experiment hinweg sind die negativen Stimuli des Valenzkonzepts mit links ([A]) und die positiven mit rechts ([5]) zu beantworten. Der erste Block besteht aus 20 evaluativen Trials, die gemäß des Gesichtsausdrucks in eine der beiden Antwortkategorien einzuordnen sind („negativ“ = [A] vs. „positiv“ = [5]). Im zweiten Block werden 20 Klassifizierungstrials präsentiert. Die Probanden werden instruiert, gleichgeschlechtliche romantischen Paare mit rechts zu beantworten („homosexuell“ = [A]) und gemischtgeschlechtliche romantische Paare mit links („heterosexuell“ = [5]). Der dritte Block stellt den ersten gemischten Messblock dar. Den Probanden werden in festgelegter Abfolge insgesamt 96 Trials aus beiden Konzepten (46 evaluative und 46 Zielkonzeptstimuli) präsentiert. Aus der Tastenbelegung „positiv oder heterosexuell“ = [5] bzw. „negativ oder homosexuell“ = [A] resultiert eine *vorurteilskompatible* Stimuluskombination. Im anschließend vierten Block erfolgt das Umlernen der Tastenbelegung für das Zielkonzept mit 20 Trials. Nunmehr sollen Bilder gemischtgeschlechtlicher romantischer Paare mit links beantwortet werden („heterosexuell“ = [A]), während Bilder gleichgeschlechtlicher romantischer Paare mit rechts zu beantworten sind („homosexuell“ = [5]). Der fünfte Block stellt den zweiten kritischen Messblock, nun mit *vorurteilsinkompatibler* Tastenbelegung, dar. Wie im dritten Block werden 96 gemischte Stimuli dargeboten, wobei jetzt negative Gesichtsausdrücke und heterosexuelle Paarbilder mit rechts (Taste [A]) zu beantworten sind und positive Gesichtsausdrücke und homosexuelle Paarbilder mit links (Taste [5]).

Für die fixierten Stimulusabfolgen der beiden kritischen Messblöcke (3. und 5. Block) wurden jeweils vier unterschiedlichen Sequenzen (à 24 Trials) generiert, in denen jeder Stimulus jeweils einmal präsentiert wird. Dies erfolgt einerseits um eine Ausgewogenheit der Stimuli innerhalb eines Block zu erzielen und andererseits, um die Reliabilität mittels Split-Half-Methode bzw. als interne Konsistenz ermitteln zu können.

Mit Ausnahme der beiden Messblöcke erhalten die Probanden eine Fehlerrückmeldung, wenn ein Stimulus mit einer falschen Taste beantwortet wird. Bei einer inkorrekten Antwort erscheint das Wort „Fehler“ in roter Schrift (Farbwerte: 255, 0, 0) in der Mitte des Bildschirms. Für die Bearbeitung eines IAT-Experiments benötigten die Probanden im Durchschnitt 12 Minuten.

#### 4.2.2.1.2 Aufbau des EAST

Der hier eingesetzte EAST besteht aus 7 Blöcken, wovon die ersten drei Blöcke der Übung dienen, und aus den letzten vier Blöcken die Messdaten resultieren. Im ersten Block, in dem 20 evaluative Stimuli präsentiert werden, lernen die Probanden die valenzbezogene Bedeutung der Antworttasten ([A]= „negativ“ vs. [5]= „positiv“). Im zweiten Block wird jeder Klassifizierungsstimulus einmal in

grüner und einmal in blauer Farbe dargeboten. Dieser 2. Übungsblock umfasst entsprechend 24 Trials, die gemäß ihrer Färbung mit der „Positiv-“, bzw. „Negativ-Taste“ beantwortet werden sollen. Im dritten Block werden 36 gemischte Trials zur kombinierten Übung beider Aufgaben dargeboten, die Versuchspersonen sollen die evaluativen Stimuli auf der Basis ihrer Valenz (Gesichtsausdruck) und die farbigen Zielkonzeptstimuli aufgrund ihrer Färbung mit der „Positiv-“, bzw. „Negativ-Taste“ beantworten. Im Anschluss an die gemischte Übung folgen vier Messblöcke á 53 gemischte Trials, die mit jeweils 5 evaluativen Trials beginnen, um die extrinsisch affektive Tastenbelegung zu festigen. Auch hier wurden im Vorfeld Stimulusabfolgen von  $N=48$  Stimuli generiert (abzüglich der einleitenden evaluativen Trials), in denen pro Block jeder evaluative Stimulus zweimal und jeder Zielkonzeptstimulus einmal in jeder Farbe präsentiert wird. Analog zum IAT bekommen die Probanden nur in den Übungsblöcken (Block 1 bis 3) Fehlerrückmeldungen im Falle nicht korrekter Reaktionen („Fehler“).

Da der EAST ein junges Verfahren ist, zu dem bislang nur wenige Erkenntnisse vorliegen, wurde die notwendige Valenzneutralität der beiden Farben, in denen die Zielkonzeptstimuli dargeboten werden, in einer Zwischensubjektvariation der Farbbedeutung geprüft. Für eine Hälfte der Probanden stellt Blau die „positive“ und Grün die „negative“ Farbe dar, für die andere Hälfte verhält es sich entsprechend umgekehrt. Das bedeutet, für die Hälfte der Probanden, sollten die grünlichen Zielkonzeptstimuli mit den negativen Gesichtsausdrücken (Taste [A]) und die bläulichen Zielkonzeptstimuli mit den positiven Gesichtsausdrücken (Taste [5]) auf einer Taste beantwortet werden. Für die andere Hälfte der Stichprobe waren bläulichen Zielkonzeptstimuli mit den negativen Gesichtsausdrücken auf der einen Taste ([A]) und grünliche Zielkonzeptstimuli mit den positiven Gesichtsausdrücken auf der anderen Taste ([5]).

Diese Zwischensubjektvariation kann ihrerseits die uneingeschränkte Vergleichbarkeit aller Probanden verringern, da die resultierenden Teilstichproben jeweils nur  $N=25$  Probanden umfassen. Bei einem Vergleich der Effekte dieser beiden Gruppen muss also die begrenzte Repräsentativität kleiner Stichproben bedacht werden, in denen Unterschiede in Merkmalsausprägungen mit höherer Wahrscheinlichkeit auf unausgewogene Stichprobencharakteristika zurückzuführen sein können, als bei größeren Stichproben. Die Gegebenheit der Normalverteilung der Maße ist jedoch auch bei der vorliegenden Gesamtstichprobe von  $N=50$  noch nicht anzunehmen. Aus diesem Grund und da die extrinsisch zugewiesene Wertigkeit der Farbe keine allgemeingültige inhaltliche Bedeutung haben sollte, wird die Farb-Valenz-Variation für den EAST getestet und die sich mit der zufälligen Stichprobenaufteilung eröffnende Möglichkeit zur Kreuzvalidierung der Effekte genutzt, auch wenn die Ergebnisse aufgrund der Stichprobengröße vorsichtig zu interpretieren sind. Die Dauer eines EAST beträgt ca. 15 Minuten.

#### 4.2.2.2 Explizite Maße

Der explizite Selbstbericht zu kognitiven und affektiven Einstellungen gegenüber Homosexualität enthielt zusätzlich einige ausgewählte Korrelate expliziter sozialer Einstellungen sowie die Motivation zur Vorurteilsfreiheit. Tabelle 1 gibt Aufschluss über den Aufbau des Fragebogens.

Die Beantwortung der Items erfolgt auf einer fünfstufigen Likertskala mit den Ankern 1: „lehne vollkommen ab“ 3: „weder, noch“ und 5: „stimme vollkommen zu“. Alle Items wurden in Richtung der entsprechenden Konstruktsbezeichnung kodiert, so dass hohe Werte eine hohe Ausprägung im Sinne des gemessenen Konstrukts anzeigen.

Zusätzlich wurden einige Kontroll-Items zu den beiden vorausgegangenen Experimenten erhoben sowie Geschlecht (1=weiblich, 2=männlich), Alter, höchster Bildungsabschluss und Geburts-Bundesland. Die Versuchsteilnehmer wurden gebeten, ihre sexuelle Orientierung mit 1= „heterosexuell“, 2= „bisexuell“ oder 3= „homosexuell“ anzugeben. Um die Bedeutung der Kontakterfahrungen der Probanden einschätzen zu können, wurde die Anzahl homo- oder bisexueller Freunde mit vier Antwortmöglichkeiten von 1= „keine“, 2= „etwa 1 bis 4“, 3= „etwa 5 bis 9“ bis 4= „mehr als 10“ erfragt. Zusätzlich wurde die Semesterzahl der Probanden erfragt.

Skala	Abk.	N Items
<i>Right-Wing-Authoritarianism</i> (Altemeyer, 1996, Kämpfe 2002 <sup>2</sup> )	RWA	12
<i>Social Dominance Orientation</i> (Sidanius & Pratto, 1999, Kämpfe 2002 <sup>1</sup> )	SDO	16
einzelne Items <i>Ambiguitätsintoleranz</i> (z.B. McDonald, 1970; Reis, 1997; Neuberg & Newsom, 1993)	AIT	8
<i>Kognitive Einstellungen zur Homosexualität</i> (Seise, Banse & Neyer, 2002) und ausgewählte Homophobie-Items (Kämpfe, 2002 <sup>2</sup> )	HOKO	17
<i>Affektive Einstellungen zur Homosexualität</i> (Seise, Banse & Neyer, 2002)	HOAF	9
<i>Modern Racism</i> (McConahay, 1986; überarbeitete und übersetzte Fassung)	MRS	7
<i>Internal Motivation to Respond Without Prejudice Scales</i> (Plant & Devine, 1998, übersetzte und psychometrisch getestete Fassung)	IMS	4
<i>External Motivation to Respond Without Prejudice Scales</i> (Plant & Devine, 1998, übersetzte und psychometrisch getestete Fassung)	EMS	5
<i>Motivation zu vorurteilsfreiem Verhalten</i> (MVV-16, Banse & Gawronski, 2003)	MVV	16
<i>Sexuelle Orientierung</i>	Sex-O	1
<i>Anzahl homo- oder bisexueller Freunde und Bekannte</i>	FreundH	1

**Tabelle 4-1**, Skalen und Items der expliziten Fragebogenmessung

<sup>2</sup> Übersetzung der Original-Items in Zusammenarbeit mit Silke Zachariae, 2002

### 4.2.3 Durchführung

Die Untersuchung fand in einem Computerlabor statt, in dem maximal fünf Personen gleichzeitig getestet werden können. Der Versuchsaufbau war für alle Probanden gleichsam festgelegt. nach einer kurzen Einführung durch den Versuchsleiter wurde zunächst der IAT Homosexualität durchgeführt. Im Anschluss daran erfolgt die Bearbeitung des EAST Homosexualität und schließlich wurde der Selbstbericht ausgeteilt. Die Durchführung der computergestützten Experimente, inklusive Instruktionen, und die Datenerfassung wurden mit der Software Inquisit 2.0 gesteuert. Im Durchschnitt erfordert eine Erhebungssitzung etwa 50 Minuten.

Nach Abschluss aller Messungen wurden die Versuchsteilnehmer kurz über den Hintergrund der Untersuchung aufgeklärt und nach dem Ausfüllen des Formulars für die Versuchspersonenstunden verabschiedet.

## 4.3 Datenbehandlung

Greenwald et al (2004) schlagen einen verbesserten Scoring-Algorithmus für IAT-Daten vor, der an den Daten des internetbasierten Einsatzes des IAT im Rahmen des „Project Implizit“ entwickelt wurde. Aus zwei Gründen erfolgt die Datenbehandlung jedoch nicht gemäß diesem Algorithmus, sondern in Anlehnung an De Houwer (2003): Das Prozedere der IAT-Erhebung, auf welches der neue Algorithmus von Greenwald und Kollegen abzielt, ist ein deutlich anderes, als in der vorliegenden Untersuchung. Experimentaldaten, die unter unkontrollierbaren und somit schwer vergleichbaren Bedingungen erhoben werden, sind prinzipiell fragwürdig. Der Anspruch, implizite, also unintendierte und unbeeinflusste Kognitionen zu erfassen, geht in den Augen der Autorin u.a. nicht mit Reaktionszeiten von maximal 10 000 ms einher, was als obere Schranke individueller Reaktionszeiten für die internetbasierten Daten festgelegt wurde. In den Daten vorliegender Laborerhebung wird bspw. keine Reaktionszeit von mehr als 9 658 ms (IAT) beobachtet.

Aus Gründen der Vergleichbarkeit der beiden impliziten Einstellungsmaße soll mit den resultierenden Daten beider Verfahren zudem gleichermaßen verfahren werden (vgl. Teige et al., 2004). In Anlehnung Greenwald, McGhee & Schwartz (1998) und De Houwer (2003) wird folgende Behandlung der Rohdaten vorgenommen:

- Eliminierung der Reaktionszeiten von Fehlertrials
- Logarithmierung der Reaktionszeiten
- Ausschluss von Personen mit einer Fehlerquote von mehr als 20%
- Ausschluss von Personen mit einer Fehlerquote, die mehr als 3 Standardabweichungen vom Gruppenmittelwert abweicht



Die logarithmische Transformation der Reaktionszeiten dient zur Stabilisierung der Varianzen (vgl. Greenwald, McGhee & Schwartz, 1998). Da Rohlatenzen in aller Regel schief verteilt sind, liegen für viele statistische Analyseverfahren Voraussetzungsverletzungen vor, denen die Logarithmierungstransformation wirksam begegnen kann.

Für die transformierten Daten wurden folgende drei Möglichkeiten festzulegender Reaktionszeitschranken getestet:

- a) Eliminierung von Trials, deren Reaktionszeiten mindestens 2 Standardabweichungen vom Mittelwert abweichen
- b) Eliminierung von Trials, deren Reaktionszeiten mindestens 3 Standardabweichungen vom Mittelwert abweichen
- c) Rekodierung extremer Reaktionszeiten auf die Schranken 300 und 3000 ms

Entscheidungskriterium hierbei waren sowohl die internen Zusammenhänge, als auch die Korrelationen der resultierenden Maße mit der Fragebogenmessung. Aufgrund geringfügiger Differenzen zwischen den verschiedenen Reaktionszeitschranken hinsichtlich dieser Kriterien, wurde verbunden den Erfahrungen anderer Experimentalpsychologen schließlich Version b), also der Ausschluss von Trials, deren Reaktionszeiten mindestens 3 Standardabweichungen vom Mittelwert abweichen, festgelegt. Sofern nicht anders ausgewiesen, basieren alle im Folgenden berichteten Ergebnisse auf dieser Transformation. Zum Zwecke der Anschaulichkeit werden vereinzelt die Rohlatenzen dokumentiert.

## **4.4 Ergebnisse**

### **4.4.1 Explizite Einstellungen**

Die Reliabilitäten der Fragebogenskalen sind mit Ausnahme von Modern Racism (MRS,  $\alpha=.60$ ) und Ambiguitätsintoleranz (AIT,  $\alpha=.67$ ) zufrieden stellend bis sehr gut (vgl. Tab. 4-2). Während die kognitiven Einstellungen zu Homosexualität mit 1.94 weit unter dem theoretischen Skalenmittel von 3 liegen, wird für die affektiven Einstellungen zu Homosexualität ein Stichprobenmittelwert relativ nahe dem theoretischen Mittel der Skala erhalten. Verhältnismäßig hoch sind die Mittelwerte für Internale Motivation zur Vorurteilsfreiheit (IMS) und die Motivation zu vorurteilsfreiem Verhalten (MVV). Insgesamt zeichnet sich also die zu erwartende liberale Haltung von Stichproben ab (vgl. Six, 1996).

Skala	Min	Max	M (SD)	Cronbach's Alpha
<b>HOKO</b>	1.22	3.83	1.94 (.61)	.93
<b>HOAF</b>	1.56	4.56	2.83 (.55)	.86
<b>RWA</b>	1.25	3.75	2.14 (.55)	.77
<b>SDO</b>	1.50	3.50	2.34 (.49)	.78
<b>AIT</b>	1.50	4.13	2.83 (.58)	.67
<b>MRS</b>	1.14	3.43	2.12 (.47)	.60
<b>IMS</b>	2.33	5.00	4.34 (.62)	.79
<b>EMS</b>	1.00	4.50	2.45 (.83)	.73
<b>MVV</b>	2.82	4.53	3.66 (.41)	.74

**Tabelle 4-2.** Deskriptive Statistiken für Fragebogenskalen; HOKO=kognitive Einstellungen zu Homosexualität, HOAF=affektive Einstellungen zu Homosexualität, RWA=Right-Wing Authoritarianism, SDO=Soziale Dominanzorientierung, AIT=Ambiguitätsintoleranz, MRS=Modern Racism, IMS/EMS=Internal/External to Respond without Prejudice, MVV=Motivation zu vorurteilsfreiem Verhalten

Das Korrelationsmuster der expliziten Maße (vg. Tab. 4-3) zeigt erwartungskonform hohe Korrelationen zwischen kognitiven und affektiven Einstellungen zu Homosexualität ( $r=.68$ ,  $p<.001$ ). Ein vergleichbar hoher Zusammenhang von  $.61$  ( $p<.001$ ) besteht zwischen RWA und HOKO, wogegen die affektiven Einstellungen nur zu  $.31$  ( $p<.05$ ) mit RWA korrelieren. Autoritäre Probanden weisen zudem höhere Maße an Ambiguitätsintoleranz ( $r=.36$ ,  $p<.01$ ) und Modern Racism ( $r=.35$ ,  $p<.05$ ) und niedrigere IMS-Ausprägungen auf ( $r=-.42$ ,  $p<.01$ ). Soziale Dominanzorientierung steht nicht in Beziehung zur Vorurteilmessung, dagegen bestehen signifikant negative Zusammenhänge zwischen SDO und IMS ( $r=-.40$ ,  $p<.01$ ) sowie MVV ( $r=-.51$ ,  $p<.001$ ). Je dominanzorientierter eine Person also ist, desto geringer fällt ihre (internale) Motivation zur Vorurteilsfreiheit aus. Zudem besteht eine positive Korrelation zwischen SDO und MRS ( $r=.35$ ,  $p<.05$ ).

Ein signifikant negativer Zusammenhang wird zwischen HOKO und IMS erhalten ( $r=-.46$ ,  $p<.001$ ). Wie zu erwarten haben Personen, die internal zur Vorurteilsfreiheit motiviert sind, positivere explizite Einstellungen zu Homosexualität. Darüber hinaus besteht keine weitere Beziehung zwischen den selbstberichteten Einstellungen und der Motivation zur Vorurteilsfreiheit Motivationsmaß. Ein positiver Zusammenhang besteht zwischen AIT und EMS ( $r=.44$ ,  $p<.01$ ), der anzeigt, dass mit steigender Ambiguitätsintoleranz die externale Motivation zur Vorurteilsfreiheit zunimmt. Modern Racism korreliert negativ mit IMS ( $r=-.39$ ,  $p<.01$ ) und MVV ( $r=-.42$ ,  $p<.01$ ). Wie Devine (1989) und Devine et al. (2002) berichten, sind IMS und EMS unkorreliert.

Das Geschlecht steht in negativer Beziehung zu AIT ( $r=-.31$ ,  $p<.05$ ); demzufolge nach sind Frauen ambiguitätsintoleranter als Männer. Die Anzahl homo- oder bisexueller Freunde steht gleichsam zu HOKO und HOAF in negativem Zusammenhang, signifikant ist die Korrelation jedoch nur für

HOKO ( $r=-.31, p<.05$ ). Zudem geben ambiguitätsintolerantere Probanden an, weniger homo- oder bisexuelle Freunde zu haben ( $r=-.30, p<.05$ ). Der positive Zusammenhang zwischen Alter und Geschlecht ( $r=.30, p<.05$ ) deutet an, dass die männlichen Probanden signifikant älter sind, als die weiblichen.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. <b>HOKO</b>												
2. <b>HOAF</b>	<b>.68</b>											
3. <b>RWA</b>	<b>.61</b>	.34										
4. <b>SDO</b>	.11	-.11	.21									
5. <b>AIT</b>	.23	.26	<b>.36</b>	-.10								
6. <b>MRS</b>	.28	.16	.35	.35	.08							
7. <b>IMS</b>	<b>-.46</b>	-.26	<b>-.42</b>	<b>-.40</b>	-.20	<b>-.39</b>						
8. <b>EMS</b>	.06	-.15	.27	.02	<b>.44</b>	.14	-.02					
9. <b>MVV</b>	-.16	-.12	-.09	<b>-.51</b>	.11	<b>-.42</b>	<b>.50</b>	.22				
10. <b>Geschlecht</b>	-.01	.17	-.14	-.11	-.31	-.14	.24	-.11	-.13			
11. <b>Alter</b>	.26	.28	.03	-.02	-.11	-.19	-.01	-.04	.09	.30		
12. <b>Sex-O</b>	-.19	-.19	-.18	-.25	-.05	-.21	.22	-.02	.20	-.08	-.13	
13. <b>FreundH</b>	-.31	-.28	-.09	.25	-.30	.04	-.02	-.13	-.05	.03	.22	.07

**Tabelle 4-3**, Korrelationen der Fragebogenmaße; Sex-O=Sexuelle Orientierung; FreundH= Anzahl homo- o. bisexueller Freunde und Bekannte; weitere Abkürzungen siehe Tab. 4-2

#### 4.4.2 Implizite Einstellungen

Zunächst werden die impliziten Messungen anhand der Kontroll-Items zur subjektiven Aufgabenschwierigkeit, zur verfolgten Strategie und zur Verständlichkeit der Instruktionen in IAT und EAST miteinander verglichen. Als objektive Kriterien zur Evaluation der beiden Messungen werden die individuellen Mittelwerte für Reaktionszeiten und Fehlerraten herangezogen. Im Anschluss daran erfolgt die Ergebnisdokumentation für den IAT, den EAST. Berichtet werden Haupteffekte, Reliabilitäten und Validitäten für Reaktionszeit- und Fehlereffekte. Da die Auswertung des EAST vergleichsweise komplizierter ist, als die des IAT und zudem der stärkere Fokus der Studie auf der Entwicklung eines reliablen EAST liegt, nimmt die Datenanalyse zu diesem Verfahren ungleichmäßig mehr Raum in der Dokumentation ein.

Im Anschluss an die Einzelauswertung der impliziten Maße erfolgt eine Gesamtbetrachtung von IAT und EAST gemeinsam mit den expliziten Fragebogenwerten.

#### 4.4.2.1 Evaluation der impliziten Maße

Eine Analyse der Kontrollfragen zu den Experimenten erbringt signifikante Unterschiede zwischen beiden Experimenten im subjektiven Schwierigkeitsgrad. Der EAST wird im Vergleich zum IAT als deutlich schwieriger erlebt. Sowohl in der Strategie, die die Probanden verfolgen, als auch bezüglich der Instruktionsverständlichkeit sind IAT und EAST jedoch miteinander vergleichbar (Tab. 4-4).

Subjektive Kontrollvariablen		M (SD)	$\Delta$	T	p
<b>Schwere der Aufgabe:</b> „sehr leicht“ (1) bis „sehr schwer“ (7)	IAT	2.64 (1.06)			
	EAST	3.76 (1.30)	-1.12	-6.08	.00
<b>Verfolgte Strategie:</b> „korrekte-„ (1) bis „schnelle Antworten“ (7)	IAT	3.78 (1.63)			
	EAST	3.46 (1.54)	.32	1.61	.11
<b>Verständlichkeit der Instruktion</b> „sehr schwer-„ (1) bis „sehr gut verständlich“ (7)	IAT	5.56 (2.04)			
	EAST	5.40 (1.99)	.16	1.31	.20

**Tabelle 4-4**, Gegenüberstellung von IAT und EAST anhand der via Fragebogen erhobenen Kontrollvariablen,  $df=49$

Die mittels Fragebogen erhobene Einschätzung der Probanden zur Schwierigkeit der Experimente spiegelt sich auch in der Gesamtbetrachtung der Experimentaldaten wider (Tab. 4-5). Die Kategorisierungstrials im IAT wurden signifikant schneller und korrekter beantwortet als im EAST. Diese deutlich geringere Fehlerrate im IAT betrifft insbesondere die kompatiblen Trials, für die im EAST etwa die dreifache Fehlerquote zu verzeichnen ist ( $M_{IAT}=4.79$ ,  $M_{EAST}=12.58$ ,  $T=6.95$ ,  $df=49$ ,  $p<.00$ ).

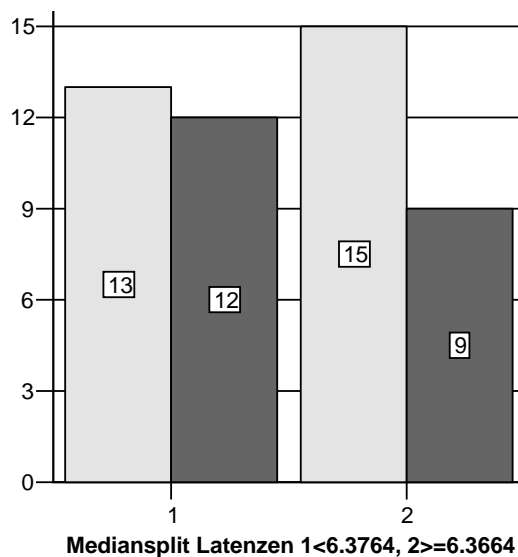
objektive Kontrollvariablen		M (SD)	$\Delta$	T	p
Individueller Latenzmittelwert	IAT	650.66 (147.09)			
	EAST	608.65 (105.68)	-42.00	-3.33	.00
Individuelle Fehlerraten	IAT	8.21 (4.67)			
	EAST	13.21 (8.09)	-5.00	-4.43	.00

**Tabelle 4-5**, Reaktionszeiten (Rohwerte) und Fehlerraten für IAT und EAST,  $df=48$

Möglicherweise führt die schwierigere Aufgabenstellung im EAST dazu, dass entweder auf schnelles oder korrektes Antworten geachtet werden kann, und entsprechend einer der beiden EAST-Scores an Aussagekraft einbüßt. Zur Überprüfung der Verteilung von Reaktionsgeschwindigkeiten und Fehlerraten

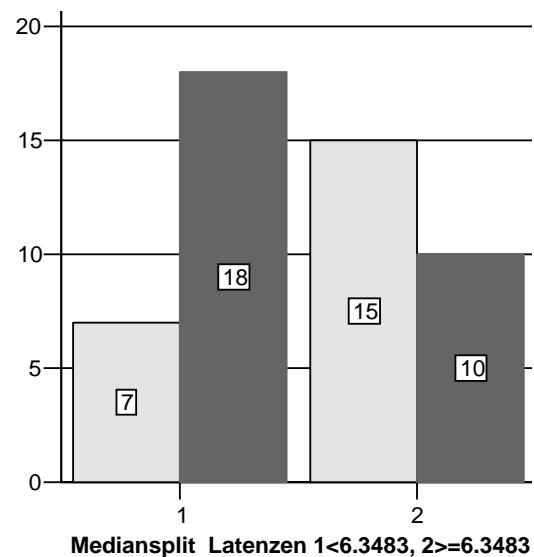
ten wird die Stichprobe sowohl für die IAT-, als auch die EAST-Daten mittels Mediansplit zum einen in „schnelle“ (1) und „langsame“ (2) Probanden geteilt (Reaktionszeiten:  $\text{Median}_{\text{IAT}}=6.3764$ ,  $\text{Median}_{\text{EAST}}=6.3483$ ) und zum anderen in „korrekte“ (1) und „nicht korrekte“ (2) Probanden (Fehlerraten:  $\text{Median}_{\text{IAT}}=7.29\%$ ;  $\text{Median}_{\text{EAST}}=11.45\%$ ).

Eine Analyse der Häufigkeiten zeigt einen deutlichen Unterschied in der Zellbesetzung zwischen IAT und EAST. Latenzen und Fehlerraten im EAST stehen in signifikant negativer Beziehung (Spearman-Korrelation=  $-0.32$ ,  $p=0.02$ ;  $\chi^2=5.195$ ,  $df=1$ ,  $p=0.02$ ). Schnelligkeit geht mit einer höheren Fehlerrate einher, während die langsamen Probanden bedeutend weniger Fehler machen (vgl. Abb. 4-2). Im Vergleich dazu ist die Zellbesetzung für die IAT-Daten nicht signifikant unterschiedlich (Spearman-Korrelation=  $-0.11$ ,  $p=0.47$ ;  $\chi^2=0.551$ ,  $df=1$ ,  $p=0.46$ , vgl. Abb. 4-3)



**Abbildung 4-2, IAT**

hellgrau=Fehlerrate < 7.29% (1)  
dunkelgrau=Fehlerrate >= 7.29% (2)



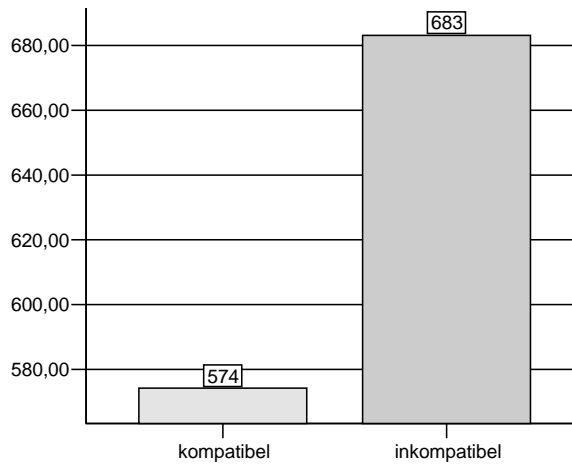
**Abbildung 4-3, EAST**

hellgrau=Fehlerrate < 11.45% (1)  
dunkelgrau=Fehlerrate >= 11.45% (2)

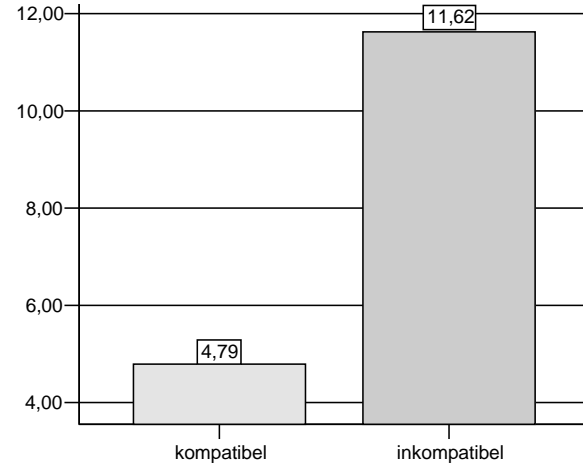
#### 4.4.2.2 Ergebnisse IAT

##### 4.4.2.2.1 Haupteffekte

Die Differenz zwischen dem inkompatiblen und dem kompatiblen Block resultiert in signifikanten IAT-Haupteffekten, und zwar sowohl für die Reaktionszeiten (Abb. 4-4), als auch für die Fehlerraten (Abb. 4-5) (vgl. Tab. 4-6).



**Abbildung 4-4**, IAT, Reaktionszeiten (Rohwerte) für Block 3 (kompatibel) und Block 5 (inkompatibel)



**Abbildung 4-5**, IAT, Fehlerraten für Block 3 (kompatibel) und Block 5 (inkompatibel)

Block	logarithmierte Latenzen in ms				Fehlerraten in %			
	M (SD)	T	p		M (SD)	T	p	
3	6.32 (.15)				4.79 (4.52)			
5	6.46 (.23)				11.63 (7.65)			
IAT	.14 (.18)	5.36	.000		6.83 (8.41)	5.75	.000	

**Tabelle 4-6**, IAT-Haupteffekte für Latenzen und Fehlerraten, Test des IAT-Effekts gegen Null,  $df=49$

#### 4.4.2.2.2 *Reliabilität*

Zur Berechnung der Reliabilität werden entsprechend der vier vollständigen Stimulussequenzen (siehe oben) pro Block jeweils vier IAT-Scores auf der Basis der Reaktionszeiten sowie der Fehlerraten berechnet. In Tabelle 4-7 sind sowohl die Intrakorrelationen (Interne Konsistenz) verzeichnet, als auch die Zusammenhänge zwischen Latenz- und Fehler-IAT-Effekten. Die interne Konsistenz (Cronbach's Alpha) der vier Unter-IATs beträgt .87 für die Reaktionszeit- und .53 für die fehlerbasierten Effekte. Werden IAT-Effekte auf der Grundlage der jeweils ersten und zweiten Hälfte der Experimentalblöcke berechnet, korrelieren die Reaktionszeiteffekte zu .77 ( $p<.001$ ), woraus eine Split-Half-Reliabilität von .87 resultiert. Die Korrelation der beiden Hälften für die Fehlereffekte beträgt .43, die Split-Half-Reliabilität entsprechend .60. Reaktionszeit- und Fehlereffekt korrelieren zu .44 ( $p<.01$ ).

Intrakorrelationen der Latenz-IATs (oben) und der Fehler-IATs (unten)					Interkorrelationen von Latenz- und Fehler-IATs				
Seq.					Fehler	Latenzen			
	1	2	3	4		1	2	3	4
1		<b>.56</b>	<b>.60</b>	<b>.38</b>	1.	.29	.35	<b>.40</b>	<b>.49</b>
2	.28		<b>.80</b>	<b>.80</b>	2.	.06	.29	.32	<b>.46</b>
3	.31	.38		<b>.72</b>	3.	-.10	.17	.06	<b>.38</b>
4	.23	.11	.02		4.	.24	.00	.19	.12

**Tabelle 4-7**, Reliabilitätsberechnung für 4 IAT-Sequenzen (Seq.), Korrelationen der latenzbasierten IATs: obere Matrix (dunkelgrau schattiert), Korrelationen der fehlerbasierten IATs: untere Matrix (hellgrau schattiert), fett schwarz gedruckte Werte:  $p < .01$ , schwarz gedruckte Werte:  $p < .05$ , grau gedruckte Werte: nicht signifikant

#### 4.4.2.2.3 Validität

Einer Betrachtung der Zusammenhänge zwischen den impliziten Assoziationen via IAT und den expliziten Einstellungen zu Homosexualität gemäß Fragebogen zufolge sind substantielle Korrelationen insbesondere für den reaktionszeitbasierten IAT-Effekt zu verzeichnen (Tab. 4-8). Hier erreichen 5 von 9 Zusammenhängen statistische Bedeutsamkeit. Der Reaktionszeit-IAT steht in positivem Zusammenhang zu kognitiven und affektiven Einstellungen zu Homosexualität (HOKO:  $r = .44$ , HOAF:  $r = .40$ ,  $p < .01$ ) und des Weiteren zu RWA ( $r = .40$ ,  $p < .01$ ), SDO ( $r = .28$ ,  $p < .05$ ) und Modern Racism ( $r = .32$ ,  $p < .05$ ).

Die Korrelationen zwischen dem fehlerbasierte IAT und den beiden expliziten Einstellungsskalen sind zwar nicht signifikant, aber dennoch positiv (HOKO:  $r = .24$ ,  $p < .10$ ; HOAF:  $r = .24$ ,  $p < .11$ ). Ähnliches gilt für die Beziehung zwischen Fehler-IAT und Autoritarismus ( $r = .25$ ,  $p < .09$ ).

	HOKO	HOAF	RWA	SDO	AIT	MRS	IMS	EMS	MVV
Reaktionszeit-IAT	<b>.44</b>	<b>.40</b>	<b>.40</b>	.28	-.05	.32	-.19	-.09	-.23
Fehler-IAT	.24	.24	.25	.00	.08	.14	-.10	-.19	-.16

**Tabelle 4-8**, Zusammenhänge zwischen IAT und expliziten Vorurteilen und weiteren Fragebogenmaßen, Abkürzungen siehe Tab. 4-2

Zwischen beiden IAT-Effekten und dem Geschlecht, dem Alter, der sexuellen Orientierung und der Anzahl homo- oder bisexueller Personen im Freundes- und Bekanntenkreis der Probanden bestehen keine statistisch bedeutsamen Zusammenhänge (vgl. Tab. 4-9). Tendenziell werden höhere Fehlereffekte für Männer im Vergleich zu Frauen und ältere Probanden festgestellt.

	<b>Geschlecht</b>	<b>Alter</b>	<b>Sex-O</b>	<b>FreundH</b>
Reaktionszeit-IAT	.02	-.06	-.08	.02
Fehler-IAT	.15	.14	.11	.06

**Tabelle 4-9**, Zusammenhänge zwischen IAT und expliziten Vorurteilen und demografischen sowie Kontaktvariablen, Sex-O=Sexuelle Orientierung; FreundH= Anzahl homo- oder bisexueller Freunde und Bekannte

#### 4.4.2.3 *Ergebnisse EAST*

##### 4.4.2.3.1 *Haupteffekte*

Der Gesamtwert des EAST liefert einen signifikanten Haupteffekt für die Reaktionszeiten. Die über alle Blöcke hinweg gemittelte Differenz der Latenzen von inkompatiblen und kompatiblen Stimuli ist positiv, das heißt im Sinne des Vorurteils gerichtet, und unterscheidet sich signifikant von Null (Tab. 4-10). Dies ist hauptsächlich auf den ersten und dritten Messblock zurückzuführen, der EAST-Effekt des vierten Blocks ist zwar positiv, jedoch nicht statistisch bedeutsam. Im zweiten Block resultiert die Reaktionszeitdifferenz für inkompatible und kompatible Stimuli in einem vorurteilskonträren negativen EAST-Effekt, der allerdings nicht signifikant ist. Bei den einzelnen Blöcken sowie dem Gesamtwert handelt sich um sehr geringe Effektgrößen (Block 1:  $d=.25$ , Block 3:  $d=.24$ , Gesamt:  $d=.16^2$ ). Innerhalb der Fehlerdaten werden weder in den einzelnen Blöcken, noch für den Gesamtwert signifikante EAST-Effekte beobachtet. Wie bereits für die Reaktionszeiten zeigen auch die Fehlerraten einen vorurteilskonträren negativen Effekt im zweiten Block an.

<b>Block</b>	<b>Rohlatenzen in ms</b>	<b>logarithmierte Latenzen in ms</b>			<b>Fehlerraten in %</b>		
	<b>M (SD)</b>	<b>M (SD)</b>	<b>T</b>	<b>p</b>	<b>M (SD)</b>	<b>T</b>	<b>p</b>
1	17.37 (68.13)	.03 (.08)	2.87	.006	2.00 (14.92)	.95	.348
2	-4.84 (67.79)	-.01 (.11)	-.50	.623	-1.17 (15.70)	-.53	.602
3	25.94 (74.87)	.03 (.10)	2.07	.043	2.67 (15.47)	1.22	.229
4	13.14 (68.38)	.02 (.09)	1.45	.153	1.50 (14.45)	.73	.467
Gesamt	12.90 (36.90)	.02 (.05)	2.46	.017	1.25 (8.97)	.99	.329

**Tabelle 4-10**, EAST-Effekte, Differenz inkompatible Trials-kompatible Trials

<sup>2</sup> Berechnet nach Bortz (2004)



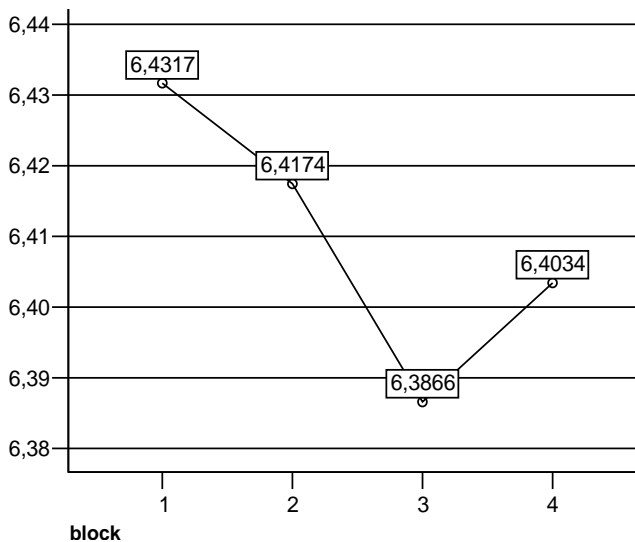
Zur simultanen Analyse der Effekte von extrinsischer Valenz und Stimulusart auf Reaktionszeiten und Fehlerraten werden in einem nächsten Schritt je Messblock Latenzmittelwerte und Fehlerraten für die verschiedenen Kombinationen von Stimulusart (heterosexuell vs. homosexuell) und extrinsischer Valenz (Blau=positiv vs. Grün=negativ) berechnet, also für Trials, die

- a) ein heterosexuelles Pärchen in der positiven Farbe (positiv x positiv),
- b) ein heterosexuelles Pärchen in der negativen Farbe (positiv x negativ),
- c) ein homosexuelles Pärchen in der positiven Farbe (negativ x positiv) und
- d) ein homosexuelles Pärchen in der negativen Farbe (negativ x negativ) präsentieren.

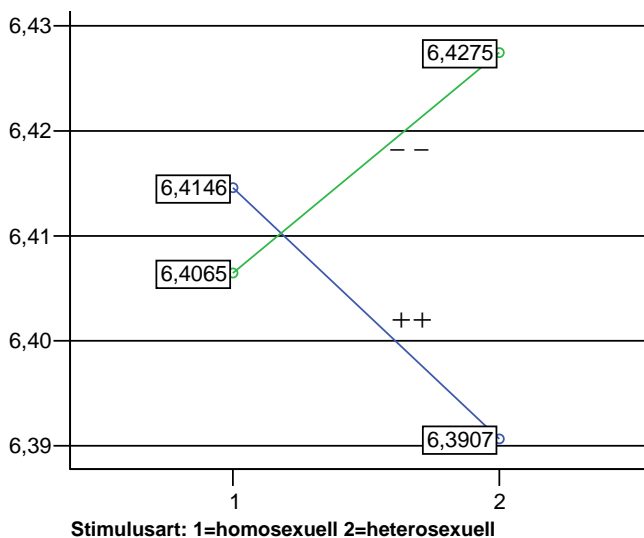
Das resultierende 4 (Block) x 2 (Stimulusart: heterosexuell x homosexuell) x 2 (extrinsische Valenz: positiv x negativ)-Design wird für Latenzen und Fehlerraten jeweils in einer ANOVA mit Messwiederholung analysiert. Tabelle 4-11 enthält die Mittelwerte und Standardabweichungen der Reaktionszeiten und Fehlerraten.

Block	Stimulus - Response	Reaktionszeiten		Fehlerraten		
		M	(SD)	M	(SD)	
1	Homosexuell x	„positive“ Farbe	6.43	(.22)	16.33	(15.61)
		„negative“ Farbe	6.39	(.20)	9.33	(13.54)
	Heterosexuell x	„positive“ Farbe	6.45	(.24)	14.67	(18.02)
		„negative“ Farbe	6.46	(.21)	11.67	(15.15)
2	Homosexuell x	„positive“ Farbe	6.41	(.23)	13.67	(16.73)
		„negative“ Farbe	6.44	(.23)	15.67	(18.57)
	Heterosexuell x	„positive“ Farbe	6.40	(.20)	12.67	(18.31)
		„negative“ Farbe	6.43	(.21)	7.67	(13.13)
3	Homosexuell x	„positive“ Farbe	6.38	(.20)	10.67	(14.97)
		„negative“ Farbe	6.38	(.23)	17.33	(18.12)
	Heterosexuell x	„positive“ Farbe	6.37	(.19)	15.33	(14.22)
		„negative“ Farbe	6.42	(.20)	14.00	(17.29)
4	Homosexuell x	„positive“ Farbe	6.44	(.20)	14.00	(18.86)
		„negative“ Farbe	6.43	(.22)	15.33	(19.59)
	Heterosexuell x	„positive“ Farbe	6.36	(.21)	10.67	(16.41)
		„negative“ Farbe	6.40	(.18)	12.33	(17.43)

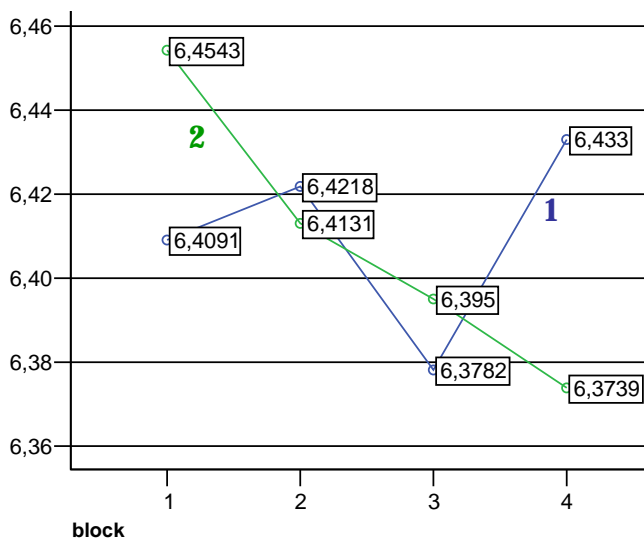
**Tabelle 4-11**, Deskriptive Statistik für ANOVA mit Messwiederholung; Stimulus-Response= Stimulusart x extrinsische Valenz



**Abbildung 4-6,** Reaktionszeiten, Haupteffekt des Blocks



**Abbildung 4-7,** Reaktionszeiten, Interaktion *Stimulusart* und *extrinsische Valenz*. ++: positiv, --: negativ

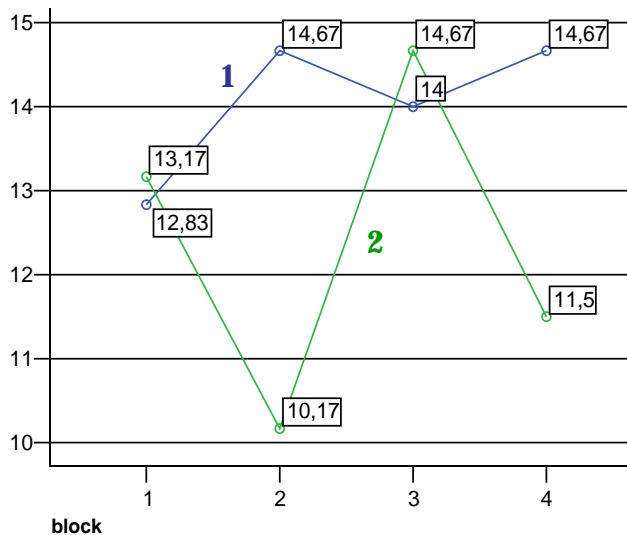


**Abbildung 4-8,** Reaktionszeiten, Interaktion *Block* und *Stimulusart*. 1=homosexuell, 2=heterosexuell

Die Analyse der Reaktionszeiten erbringt einen Haupteffekt des Blocks:  $F(3, 50) = 3.18, p < .05$  (Abb. 4-6) resultierend aus abnehmenden Reaktionszeiten vom ersten bis zum dritten Block und einem erneuten Reaktionszeitanstieg im letzten Block. Einen signifikanten Haupteffekt erbringt zudem die extrinsische Valenz:  $F(1, 51) = 4.17, p < .05$ , aufgrund kürzerer Reaktionszeiten bei extrinsisch positiven („positive Farbe“) im Vergleich zu extrinsisch negativen Stimuli („negative Farbe“).

Die kritische Interaktion zwischen Stimulusart und extrinsischer Valenz ist signifikant ( $F(1, 50) = 5.93, p = .004$ ; Abb. 4-7) und insbesondere auf Reaktionszeitdifferenzen für die heterosexuellen Stimuli zurückzuführen. Während die extrinsische Valenz für die homosexuellen Stimuli keinen deutlichen Effekt in den Reaktionszeiten hat, werden für die Darbietung von heterosexuellen Stimuli in der extrinsisch positiven Farbe deutlich kürzere Latenzen erhalten, als bei Darbietung in der negativen Farbe. Das bedeutet einen Reaktionszeitvorteil der einstellungskonformen Stimulus-Valenz-Kombination, jedoch nur für die heterosexuellen Stimuli.

Zwischen Block und Stimulusart besteht ebenfalls eine signifikante Interaktion ( $F(3, 50) = 9.99, p = .000$ ; Abb. 4-8): Nehmen die Reaktionszeiten für heterosexuelle Paarbilder über die Blöcke hinweg kontinuierlich ab, schwankt die Geschwindigkeit, mit der die Probanden die homosexuellen Paarbilder beantworten dagegen stark, wobei im letzten EAST-Block die höchsten Reaktionszeiten für homosexuelle Stimuli beobachtet werden.



**Abbildung 4-9**, Fehler, Interaktion *Block* und *Stimulusart*. 1=homosexuell, 2=heterosexuell

Die Analyse der Fehlerraten erbringt keinerlei signifikante Effekte auf dem 5%-Niveau. Marginal signifikant ist allein die Interaktion zwischen *Block* und *Stimulusart* ( $F(3, 50)=2.68$   $p=.058$ ; Abb. 4-9).

#### 4.4.2.3.2 *Reliabilität*

Für die Reliabilitätsberechnung der der EAST-Effekte werden zum Einen interne Konsistenzen (Cronbach's Alpha) der jeweils 4 Reaktionszeit- und Fehlereffekte betrachtet sowie die Korrelationen der reaktionszeit- und fehlerbasierten EAST-Differenzen inkongruenter und kongruenter Trials getrennt für homosexuelle und heterosexuelle Paarbilder zwischen den Messblöcken betrachtet.

Die interne Konsistenz der reaktionszeitbasierten EAST-Effekte gemäß den vier Messblöcken beträgt .20, für die Fehlereffekte wird ein höheres, jedoch immer noch unbefriedigendes Cronbach's Alpha von .38 erhalten. Von den insgesamt 28 Korrelationen zwischen den Reaktionszeitdifferenzen sind 9 statistisch bedeutsam, was einen überzufälligen Anteil signifikanter Ergebnisse von 32.14% bedeutet. Die reaktionszeitbasierten EAST-Effekte für homosexuelle Stimuli korrelieren signifikant positiv zwischen dem dritten und dem vierten Block ( $r=.34$ ,  $p<.05$ ), und negativ zwischen erstem und vierten Block ( $r=-.36$ ,  $p<.05$ ). Dieser negative Zusammenhang ist hinsichtlich der Reliabilität des Verfahrens ein unbefriedigendes Ergebnis. Für die heterosexuellen Stimuluseffekte bestehen positive Zusammenhänge zwischen erstem und vierten ( $r=.33$ ,  $p<.05$ ), sowie zwischen drittem und vierten Block ( $r=.40$ ,  $p<.01$ ).

Von den fünf signifikanten Korrelationen zwischen den EAST-Effekten für homo- und heterosexuelle Stimuli sind zwei Zusammenhänge positiver und drei Zusammenhänge negativer Natur. Dieser Befund ist ebenso wie die negativen Zusammenhänge innerhalb einer Stimulusart unbefriedigend, da die stimulusabhängigen EAST-Effekte die vorurteils*kongruenten* Differenzen inkompatibler und kompatibler Stimulus-Valenz-Kombinationen abbilden und insofern gleichgerichtet sein sollten (Tab. 4-12).

Reaktionszeiten		Homosexuelle Stimuli				Heterosexuelle Stimuli		
		Block	1	2	3	4	1	2
Homo- sexuell	1							
	2		.15					
	3		-.06	-.06				
	4		-.36	-.17	.34			
Hetero- sexuell	1		<b>-.56</b>	-.08	-.10	<b>-.50</b>		
	2		-.11	-.34	.06	-.08	.23	
	3		.01	.26	.27	.18	.21	.03
	4		<b>.37</b>	.20	.33	.00	.33	-.13

**Tabelle 4-12**, Reliabilitätsberechnung für die reaktionszeitbasierten EAST-Scores nach De Houwer (2003)

Die analoge Betrachtung der stimulusabhängigen Fehlerraten, zeigt lediglich 4 statistisch bedeutsame Zusammenhänge auf (Anteil signifikanter Ergebnisse: 14.29%). Ähnlich zu den Reaktionszeitbasierten EAST-Werten bestehen für die homosexuellen Stimuluseffekte positive Beziehungen zwischen erstem und zweitem ( $r=.32$ ,  $p<.05$ ), sowie dritten und vierten Block, wobei letztgenannte Korrelation von .26 nicht signifikant ist ( $p>.05$ ). Erster und vierter Block sind wiederum negativ korreliert ( $r=-.40$ ,  $p<.01$ ). Für die heterosexuellen Stimuluseffekte wird eine signifikant positive Beziehung zwischen zweitem und vierten Block erhalten ( $r=.36$ ,  $p<.01$ ).

Die Korrelationen von homo- und heterosexuellen Stimuluseffekten sind mit einer Ausnahme nicht signifikant. Hier erreicht lediglich die Korrelation zwischen dem EAST-Effekt für heterosexuelle Stimuli im ersten Block und dem Effekt für homosexuelle Stimuli im dritten Block statistische Bedeutung ( $r=.31$ ,  $p<.05$ , vgl. Tab. 4-13).

Fehlerraten		Homosexuelle Stimuli				Heterosexuelle Stimuli		
		Block	1	2	3	4	1	2
Homo- sexuell	1							
	2		.32					
	3		-.07	-.25				
	4		<b>-.40</b>	-.28	.26			
Hetero- sexuell	1		-.13	-.02	.31	-.02		
	2		-.06	.24	.12	.10	.23	
	3		-.13	.23	.12	.23	.00	-.03
	4		.13	-.01	.08	.20	.10	<b>.36</b>

**Tabelle 4-13**, Reliabilitätsberechnung für die fehlerbasierten EAST-Scores nach De Houwer (2003)

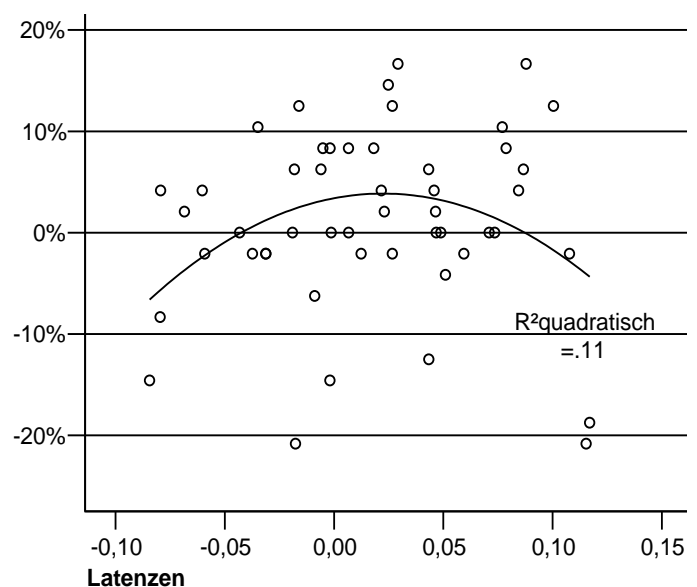
Von den 16 möglichen Korrelationen zwischen den jeweils vier Reaktionszeit- und Fehlereffekten für homosexuelle Stimuli erreichen drei Werte das Signifikanzniveau, wobei zwei Korrelationen davon ein negatives Vorzeichen aufweisen (Reaktionszeiten Block 4, Fehler Block 1:  $r=-.31$ ,  $p<.05$ ; Reaktionszeiten Block 1, Fehler Block 3:  $r=-.38$ ,  $p<.01$ ). Der erwartungskonforme, da positive Zusammenhang besteht zwischen Reaktionszeiten und Fehlereffekten des vierten Blocks ( $r=.51$ ,  $p<.001$ ). Von den 16 analog berechneten Korrelationen zwischen den Reaktionszeit- und Fehlereffekten für heterosexuelle Stimuli ist nur ein Zusammenhang statistisch bedeutsam und zudem positiv gerichtet (Reaktionszeiten Block 4, Fehler Block 2:  $r=.32$ ,  $p<.05$ ).

Reaktionszeiten	Fehler				Gesamt
	1	2	3	4	
1	-.03	-.11	.11	.17	.05
2	.13	.04	.23	.28	.28
3	-.37	.01	-.20	-.09	-.27
4	-.02	.20	-.29	.28	.07
Gesamt	-.13	.08	-.08	.29	.06

**Tabelle 4-14**, Korrelationen zwischen blockspezifischen Reaktionszeit- und Fehlereffekten des EAST

Über homo- und heterosexuelle Stimuli hinweg stehen reaktionszeit- und fehlerbasierte EAST-Effekte der Datenauswertung zufolge in keinem systematischen Zusammenhang, was die Analyse der Korrelationen zwischen den Reaktionszeit- und Fehlereffekten der einzelnen Blöcke erkennen lässt (Tab. 4-14). Die Korrelation der über alle Blöcke gemittelten EAST-Effekte beträgt .06 ( $p>.05$ ).

Die zum Teil beobachteten negativen Zusammenhänge zwischen Reaktionszeit- und Fehler-EAST-Effekten sind auf einen nichtlinearen Zusammenhang zwischen beiden Datentypen zurückzuführen. Bei der Vorhersage des Fehlereffekts durch den Reaktionszeiteffekt (Gesamtwerte) ist die Änderung im  $R^2$  durch einen quadratischen Reaktionszeit-Term signifikant ( $F_{1, 47}=5.47$ ,  $p=.02$ ) und beträgt .10 (Abb. 4-10).



**Abbildung 4-10**, Nichtlinearer Zusammenhang zwischen Reaktionszeit- und Fehlereffekte des EAST

#### 4.4.2.3.3 *Bedeutung der Farb-Valenz-Zuordnung*

Um die Effekte der Farb-Valenz-Zuordnung zu prüfen, wird eine multivariate Varianzanalyse getrennt für reaktionszeit- und fehlerbasierte EAST-Effekte mit der Farb-Valenz-Zuordnung als festem zwei-stufigen Faktor analysiert. Die Mittelwerte und Statistiken enthält Tabelle 4-15.

Block	Farb-Valenz-Variation	Reaktionszeiten			Fehlerraten in %		
		M (SD)	T	p	M (SD)	T	p
1	1	.02 (.08)	1.68	.201	7.00 (11.46)	6.21	.016
	2	.05 (.09)			-3.00 (16.47)		
2	1	-.01 (.10)	.010	.931	-4.00 (13.20)	1.65	.205
	2	-.01 (.11)			1.67 (17.68)		
3	1	-.01 (.09)	10.68	.002	10.00 (15.59)	14.28	.000
	2	.08 (.10)			-4.67 (11.56)		
4	1	-.01 (.10)	2.25	.140	2.67 (15.91)	.32	.573
	2	.04 (.10)			.33 (13.06)		
<b>Gesamt</b>	1	.00 (.05)	2.85	.034	3.92 (6.37)	4.90	.002
	2	.04 (.05)			-1.42 (10.43)		

**Tabelle 4-15**, T-Test der EAST-Scores in Abhängigkeit von der Farb-Valenz-Zuordnung, 1 (N=25): blau=“positiv“, grün=“negativ“, 2 (N=25): blau=“negativ“ grün=“positiv“

In Abhängigkeit von der Zwischensubjekt-Variation der Farb-Valenz-Zuordnung bestehen zum Teil signifikante Unterschiede sowohl für die Reaktionszeit-, als auch die Fehlereffekte. Die Kombination von blau mit positiver Valenz und grün mit negativer Valenz (Variation 1) resultiert für die fehlerbasierten EAST-Werte des ersten und dritten Blocks in signifikant *größeren* Effekten. Diese Unterschiede schlagen sich ebenfalls im Gesamtwert des EAST nieder. Dagegen liefert Variation 2 für die Reaktionszeiten im dritten Block einen signifikant größeren Effekt. Auch hier wirkt sich der Unterschied in diesem einen Block auf den Gesamtwert aus.

Zusammenfassend erbringt der Vergleich beider Teilstichproben keine systematischen Unterschiede der extrinischen Farb-Valenz-Zuordnung für die EAST-Effekte. Ob Blau die „positive“ und Grün die „negative Farbe“ für die Probanden darstellte oder umgekehrt, liefert weder ein einheitliches, noch ein durchgängig unterscheidbares Muster. In den Reaktionszeitdaten resultiert Variation 2 für einen Block in größeren Effekten, während bei Betrachtung der Fehler Variation 1 in zwei Blöcken größere Effekte liefert. Ebenso gegenläufige Ergebnisse erbringt der Gruppenvergleich für die Gesamteffekte. Ein Reaktionszeiteffekt von Null wird für Gruppe 1 beobachtet, wogegen in Gruppe 2 ein hypothesenkonformer positiver Haupteffekt beobachtet wird. Innerhalb der Fehlerdaten weist Variation 2 sogar einen negativen Haupteffekt auf, während die Gruppe mit Variation 1 den erwartungsgemäß positiven

Effekt erbringt. Die Stabilität der EAST-Effekte muss auch entsprechend dieser Analyse als gering bewertet werden, wie bereits die Reliabilitätsanalysen offen legen.

#### 4.4.2.3.4 Validität

Um die Validität der EAST-Effekte zu prüfen, werden die Zusammenhänge zwischen EAST-Effekte und den Fragebogenmaßen nicht nur in der Gesamtstichprobe, sondern auch getrennt für beide Farb-Valenz-Gruppen betrachtet (Tab. 4-16).

	<b>Farb-Valenz-Variation 1</b>		<b>Farb-Valenz-Variation 2</b>		<b>Gesamtstichprobe</b>	
	HOKO	HOAF	HOKO	HOAF	HOKO	HOAF
<b>Reaktionszeiten</b>						
EAST 1	.49	.49	.29	.42	.36	<b>.41</b>
EAST 2	.15	.16	-.07	-.01	.05	.08
EAST 3	-.03	-.08	.19	.24	.03	.01
EAST 4	-.15	.05	.34	-.20	.04	-.09
Gesamt	.18	.28	.33	.20	.19	.16
<b>Fehlerraten</b>						
EAST 1	.05	.11	.29	-.01	.19	.09
EAST 2	.13	.13	.05	.09	.07	.07
EAST 3	<b>.43</b>	.38	-.11	.14	.25	<b>.32</b>
EAST 4	.04	.12	.10	.27	.07	.19
Gesamt	.37	<b>.43</b>	.14	.16	.24	<b>.29</b>

**Tabelle 4-16**, Zusammenhang EAST-Werte (pro Block und gesamt) mit Fragebogenskalen in Abhängigkeit von der Farb-Valenz-Variation, 1 (N=25): blau=“positiv“, grün=“negativ“, 2 (N=25): blau=“negativ“ grün=“positiv“; HOKO=kognitive Einstellungen zu Homosexualität, HOAF=affektive Einstellungen zu Homosexualität

Für beide Farb-Valenz-Variationen steht der Reaktionszeiteffekt des ersten EAST-Blocks in erwartungskonformen Zusammenhang mit den Vorurteilsmaßen des Fragebogens, dies trifft auch für die Gesamtstichprobe zu (HOKO:  $r=.36$ ,  $p<.05$ ; HOAF:  $r=.41$ ,  $p<.01$ ). Block zwei bis vier stehen in keinem statistisch bedeutsamen Zusammenhang zu den Vorurteilsskalen.

Zwischen den Fehlereffekten und den expliziten Einstellungsmaßen wird für die Blöcke ein weniger konsistentes Zusammenhangsmuster beobachtet. Der EAST-Effekt des dritten Blocks korreliert positiv mit beiden Vorurteilsmaßen und erreicht statistische Bedeutsamkeit für kognitive Vorurteile, jedoch nur für Gruppe 1 der Farb-Valenz-Variation ( $r=.43$ ,  $p<.05$ ). Für die Gesamtstichprobe besteht ein positiver Zusammenhang zwischen Block 3 und insbesondere den affektiven Vorurteilen gegenüber Homosexualität ( $r=.32$ ,  $p<.05$ ). Auch der über die vier Blöcke gemittelte Gesamtwert des Fehler-

effekts steht in positiver Beziehung zu affektiven Vorurteilen in der Gesamtstichprobe ( $r=.29, p<.05$ ), was speziell durch einen hohen Zusammenhang innerhalb der Farb-Valenzgruppe 1 zustande kommt.

Zwischen den EAST-Effekten und den expliziten Prädiktormäße bestehen deutlich weniger statistisch bedeutsame Korrelationen. Von den 110 betrachteten Korrelationen, erreichen lediglich 6 Zusammenhänge statistische Bedeutsamkeit, was einem kaum überzufälligen Signifikanzanteil von 5.45% entspricht (Tab. 4-17). Das Zusammenhangsmuster erscheint zudem teilweise nicht sinnvoll interpretierbar – so korreliert der Fehlereffekt des zweiten Blocks negativ mit der Motivation zu vorurteilsfreiem Verhalten ( $r=-.33, p<.05$ ), während der dritte Block zu MVV in positiver Beziehung steht ( $r=.33, p<.05$ ). Hypothesenkonträr korreliert der reaktionszeitbasierte EAST-Effekt des dritten Blocks negativ mit Autoritarismus ( $r=-.29, p<.01$ ), je autoritärer eine Versuchsperson ist, desto weniger Unterschiede bestehen für sie zwischen den valenzbezogenen Assoziationen für homo- und heterosexuellen Stimuli. Der signifikant positive Zusammenhang zwischen fehlerbasiertem Gesamtwert und Modern Racism ( $r=.31, p<.05$ ) ist dagegen inhaltlich sinnvoll. Substanzielle Beziehungen zwischen impliziten Maßen und motivationalen Variablen sollten, mit Ausnahme der Internalen Motivation zur Vorurteilsfreiheit (IMS, siehe Abschnitt 2.2.2), aufgrund der relativen Unbeeinflussbarkeit der indirekten Verfahren nicht bestehen. Geschlecht und Alter korrelieren nicht signifikant mit den EAST-Effekten, ebenso wenig wie die Anzahl homo- oder bisexueller Freunde, die ein Proband angibt zu haben. Die sexuelle Orientierung korreliert signifikant negativ zum Reaktionszeiteffekt des zweiten Blocks ( $r=-.37, p<.01$ ). Das bedeutet, in den Reaktionszeiten weisen die beiden Probanden, die angaben, homosexuell orientiert zu sein, deutlich geringere EAST-Effekte auf. Dieser erwartungskonforme Befund wird jedoch nicht stabil über alle Blöcke und beide Datenformate beobachtet.

	<b>RWA</b>	<b>SDO</b>	<b>AIT</b>	<b>MRS</b>	<b>IMS</b>	<b>EMS</b>	<b>MVV</b>	<b>Geschl.</b>	<b>Alter</b>	<b>Sex-O</b>	<b>FreundH</b>
<b>Reaktionszeiten</b>											
EAST 1	.04	.14	.01	.05	-.14	-.06	-.24	-.09	.04	-.14	.08
EAST 2	-.01	-.10	.07	.01	.00	-.17	.11	-.06	.08	<b>-.37</b>	.01
EAST 3	-.29	.17	-.17	-.02	.21	-.20	-.05	.15	.02	.19	.11
EAST 4	.09	-.07	.17	.20	.04	.06	.04	-.18	-.11	-.20	-.01
Gesamt	-.09	.05	.07	.11	.07	-.18	-.05	.14	.02	.09	.08
<b>Fehlerraten</b>											
EAST 1	.24	.17	.06	.13	-.13	-.06	-.12	.08	-.03	-.07	-.12
EAST 2	.05	.24	.07	.25	-.19	-.16	<b>-.33</b>	.16	.04	.12	.12
EAST 3	.09	-.12	-.03	.11	.18	.07	<b>.31</b>	.10	.04	-.18	-.27
EAST 4	.03	.12	.15	.25	<b>-.30</b>	-.18	-.23	-.02	.07	-.16	-.09
Gesamt	.17	.18	.10	<b>.31</b>	-.18	-.14	-.16	.15	.06	-.13	-.15

**Tabelle 4-17,** Zusammenhang EAST-Werte (pro Block und gesamt) mit Fragebogenskalen für die Gesamtstichprobe; Abkürzungen siehe Tab. 4-2 bzw. 4-3



Da hierin ein Vorteil des EAST gegenüber dem IAT besteht (vgl. Kapitel 1.6), werden im Folgenden die EAST-Effekte getrennt für homosexuelle und heterosexuelle Bilder<sup>3</sup> betrachtet (Tabelle 4-18).

Für die Reaktionszeitdaten ergibt sich folgendes Bild: Erster und zweiter Block liefern insbesondere für die homosexuellen Stimuluseffekte erwartungsgetreue Zusammenhänge mit expliziten Vorurteilen; die Korrelationen der heterosexuellen Bilder mit den Fragebogenmaßen erreichen zwar nicht das 5%-Signifikanzniveau, sind jedoch betragsmäßig richtig gerichtet. Für die letzten beiden Blöcke zeigt sich ein nichtsignifikantes und zudem inkonsistentes Zusammenhangsmuster.

In den Fehlerdaten werden keine sinnvollen Korrelationen zwischen den homosexuellen Stimuluseffekten und den Fragebogenmaßen erhalten. Dagegen sind alle Zusammenhänge für die heterosexuellen Stimuluseffekte, wie zu erwarten ist, positiv. Die Beträge der Korrelationen erreichen jedoch nicht statistische Bedeutsamkeit, im Falle des Gesamtwertes, der mit kognitiven und affektiven Einstellungen zu Homosexualität zu jeweils .27 korreliert, wird das Alpha-Niveau knapp verfehlt ( $p=.06$ ).

	Reaktionszeiten		Fehlerraten	
	HOKO	HOAF	HOKO	HOAF
<b>Homosexuelle Stimuli</b>				
EAST 1	.07	.31	.04	-.02
EAST 2	.28	.14	.06	-.02
EAST 3	.13	.10	-.17	-.22
EAST 4	-.05	-.04	-.01	-.12
EAST Gesamt	.23	.27	-.06	-.22
<b>Heterosexuelle Stimuli</b>				
EAST 1	.27	.14	.19	.13
EAST 2	.20	.10	.18	.16
EAST 3	.02	-.04	.16	.21
EAST 4	-.03	-.11	.09	.14
EAST Gesamt	.21	.06	.27	.27

**Tabelle 4-18.** Korrelationen zwischen stimulusabhängigen EAST-Effekten und expliziten Einstellungen

#### 4.4.2.4 Zusammenhänge zwischen IAT und EAST

Zur Untersuchung der Konvergenz beider impliziter Maße werden die Beziehungen zwischen den vier Blöcken des EAST und den vier IAT-Teileffekten auf der Basis der vier unabhängigen Stimulussequenzen in kompatibel und inkompatiblen Block für Reaktionszeit- und Fehlerdaten betrachtet (Tab. 4-19).

<sup>3</sup> EAST für homosexuelle Stimuli: [(Stimulusart *homosexuell* x Valenz *positiv*) – (Stimulusart *homosexuell* x Valenz *negativ*)];  
EAST für heterosexuelle Stimuli = [(Stimulusart *heterosexuell* x Valenz *negativ*) – (Stimulusart *heterosexuell* x Valenz *positiv*)]

Hierfür bestehen substantielle positive Zusammenhänge lediglich zwischen dem reaktionszeitbasierten EAST-Effekt des ersten Blocks und drei der vier IAT-Teileffekte, sowie den Gesamt-IAT. Das bedeutet hohe EAST-Effekte im ersten Block gehen hypothesenkonform mit hohen IAT-Effekten einher. Darüber hinaus liegen für die restlichen EAST-Blöcke sowie den Gesamt-EAST keine statistisch bedeutsamen Korrespondenzen zum IAT vor.

	<b>IAT</b>				
	1	2	3	4	Gesamt
<b>Reaktionszeiten</b>					
EAST 1	.19	.34	<b>.40</b>	<b>.39</b>	.34
EAST 2	-.06	-.05	.02	.19	-.00
EAST 3	-.11	-.18	-.15	.00	-.13
EAST 4	-.19	-.22	-.14	-.03	-.19
EAST Gesamt	-.10	-.08	.03	.23	-.02
<b>Fehler</b>					
EAST 1	-.10	-.05	.05	-.02	-.02
EAST 2	.03	-.07	.05	.08	.01
EAST 3	.11	.20	.11	.08	.12
EAST 4	.06	.07	.09	.13	.22
EAST Gesamt	-.20	-.13	-.05	-.20	-.16

**Tabelle 4-19,** Zusammenhänge zwischen IAT (Sequenzen) und EAST (Blöcke)

#### 4.4.3 Gesamtbetrachtung von IAT, EAST und Fragebogenmaßen

Für eine gemeinsame Vorhersage expliziter Vorurteile durch IAT und EAST muss berücksichtigt werden, dass weder die Reaktionszeit- noch die Fehlereffekte des EAST als Gesamtwert geltend gemacht werden können, da zwischen den Blöcken zu große Divergenzen bestehen. Aufgrund seiner Relevanz für die explizite Einstellungsmessung und die erwartungskonforme Korrespondenz zum IAT wird daher lediglich der *erste* EAST-Block gemeinsam mit dem Gesamt-IAT als Prädiktor expliziter Einstellungen gemäß Fragebogenmaßen betrachtet.

In einer ersten Regressionsanalyse werden die beiden impliziten Maße als Prädiktoren zur Vorhersage kognitiver und affektiver Vorurteile aufgenommen (Modell A, Tab. 4-20). Hierbei zeigt sich zunächst die Überlegenheit des IAT zur Vorhersage kognitiver expliziter Einstellungen ( $\text{Beta}_{\text{kog}}=.41, p<.01$ ) und die vergleichsweise größere Bedeutung des EAST für affektive Einstellungen ( $\text{Beta}_{\text{kog}}=.30, p<.0$ , Tab. 4-20).

Prädiktoren	kognitive Einstellungen (HOKO)			affektive Einstellungen (HOAF)		
	Beta	<i>p</i>	Modellstatistik	Beta	<i>p</i>	Modellstatistik
<b>IAT</b>	.35	.017	R <sup>2</sup> =.23	.26	.077	R <sup>2</sup> =.22
<b>EAST</b>	.22	.128	F (2, 46)=7.00, <i>p</i> <.01	.29	.048	F (2, 46) =6.52, <i>p</i> <.01

**Tabelle 4-20**, Vorhersage expliziter Einstellungen durch IAT und EAST (Modell A)

In einem zweiten Schritt wird die maximale Vorhersagekraft der expliziten Prädiktormäße für die kognitiven und affektiven Einstellungen untersucht (Modell B, Tab. 4-21). Gemeinsam erklären die Prädiktoren 62% der Varianz kognitiver Einstellungen zu Homosexualität (korrigiertes R<sup>2</sup>=.50). Signifikante Beta-Gewichte erzielen dabei RWA, das Alter und die Anzahl homo- oder bisexueller Freunde. Die Varianz affektiver Einstellungen zu Homosexualität kann zu 49% durch die Prädiktormäße aufgeklärt werden (korrigiertes R<sup>2</sup>=.33). Daran sind nunmehr mit signifikanten Regressionsgewichten EMS und wiederum das Alter der Probanden beteiligt.

Prädiktoren	kognitive Einstellungen (HOKO)			affektive Einstellungen (HOAF)		
	Beta	<i>p</i>	Modellstatistik	Beta	<i>p</i>	Modellstatistik
<b>RWA</b>	.46	.001		.22	.137	
<b>SDO</b>	-.00	.999		-.14	.408	
<b>AIT</b>	-.01	.968		.30	.065	
<b>MRS</b>	.15	.246		.19	.193	
<b>IMS</b>	-.25	.086	R <sup>2</sup> =.62	-.15	.387	R <sup>2</sup> =.49
<b>EMS</b>	-.13	.283	F (11, 38) =5.51	-.37	.017	F (11, 37) =3.19
<b>MVV</b>	.04	.778	<i>p</i> <.001	.02	.894	<i>p</i> <.01
<b>Geschlecht</b>	.04	.759		.21	.156	
<b>Alter</b>	.33	.006		.30	.030	
<b>Sex-O</b>	.03	.766		-.03	.816	
<b>FreundH</b>	-.37	.002		-.26	.058	

**Tabelle 4-21**, Vorhersage expliziter Einstellungen durch expliziten Prädiktoren, (Modell B)

Abschließend wird geprüft, inwieweit die bereits relativ hohe Varianzaufklärung an expliziten Einstellungen durch die Prädiktormäße durch die Hinzunahme der impliziten Einstellungsmaße gesteigert werden kann. Zu diesem Zweck werden hierarchische lineare Regressionen berechnet, mit denen die Änderung in der Varianzaufklärung expliziter Einstellungen durch die zusätzlich einbezogenen impliziten Maße untersucht werden kann (Modell C, Tab. 4-22).

Das Hinzunehmen von IAT und EAST zur Vorhersage kognitiver Einstellungen zu Homosexualität resultiert in einer signifikant gesteigerten Varianzaufklärung von 11.5% (Änderung in F(2, 36)=7.61,

$p < .001$ ), wobei es nunmehr der EAST-Effekt ist, der neben RWA, dem Alter und der Anzahl homo- oder bisexueller Freunde der ein signifikantes Regressionsgewicht erhält. Bei der Vorhersage affektiver Einstellungen zu Homosexualität führt der zusätzliche Einbezug von IAT und EAST zu einer Steigerung der Varianzaufklärung um 14.9% (Änderung in  $F(2, 35)=7.14$ ,  $p < .01$ ).

Prädiktoren	kognitive Einstellungen (HOKO)			affektive Einstellungen (HOAF)		
	Beta	<i>p</i>	Modellstatistik	Beta	<i>p</i>	Modellstatistik
<b>RWA</b>	.42	.001		.13	.369	
<b>SDO</b>	-.04	.716		-.18	.218	
<b>AIT</b>	-.05	.661		.28	.060	
<b>MRS</b>	.15	.196		.17	.216	
<b>IMS</b>	-.24	.053		-.14	.345	
<b>EMS</b>	-.09	.398	$R^2=.73$	-.32	.017	$R^2=.64$
<b>MVV</b>	.10	.464	$F(13, 36)=7.45$	.08	.626	$F(13, 35)=4.69$
<b>Geschlecht</b>	-.01	.898	$p < .001$	.15	.253	$p < .001$
<b>Alter</b>	.36	.001		.35	.006	
<b>Sex-O</b>	.03	.747		-.04	.735	
<b>FreundH</b>	-.38	.001		-.26	.031	
<b>IAT</b>	.12	.303		.24	.106	
<b>EAST</b>	.29	.007		.26	.042	

**Tabelle 4-22**, Vorhersage expliziter Vorurteile durch explizite Prädiktoren und IAT und EAST (Modell C)

Die Vorhersage der expliziten Einstellungen zu Homosexualität mit den Fehlereffekten von IAT und EAST (Block 1) schlägt fehl. An kognitiven Einstellungen klären beide impliziten Fehlereffekte 9.3% der Varianz auf ( $F(2, 47)=2.41$ ,  $p=.101$ ), an affektiven Einstellungen 6.6% ( $F(2, 46)=1.62$ ,  $p=.210$ ). Die hierarchischen Regressionen liefern entsprechend keine signifikanten Steigerungen der Varianzaufklärung durch die Hinzunahme der impliziten Fehlermaße.

#### 4.4.4 Moderation durch Motivation zur Vorurteilsfreiheit

Zur Untersuchung von Moderatoreffekten im Zusammenhang von impliziten und expliziten Vorurteilen durch die Motivation zur Vorurteilsfreiheit werden die drei via Fragebogen erhobene Skalen *Externale* und *Internale Motivation to Respond Without Prejudice* (EMS und IMS) sowie *Motivation zu vorurteilsfreiem Verhalten* (MVV) mittels allgemeinem linearem Modell auf ihre Moderatorwirkung hin untersucht. Zur besseren Vergleichbarkeit werden alle betrachteten Maße z-standardisiert. Anschließend werden jeweils für IAT und EAST drei Produktterme aus dem Motivationsmaß und dem impliziten Maß gebil-

det (z.B. IAT\*IMS). Diese Moderatorterme werden in einer hierarchischen Regression auf ihre inkrementelle Relevanz zur Vorhersage von kognitiven und affektiven Einstellungen zu Homosexualität durch IAT bzw. EAST und Motivation zur Vorurteilsfreiheit geprüft. Aus diesen möglichen Kombinationen ergeben sich insgesamt 12 Einzelanalysen. An dieser Stelle finden ausschließlich die Reaktionszeiteffekte Beachtung.

Eine Analyse der Regressionsmodelle erbringt keine, auf dem 5%-Niveau signifikanten, Moderatoreffekte seitens der verschiedenen Motivationsmaße. Auch in den beiden Gesamtmodellen zur Vorhersage kognitiver und affektiver Einstellungen zu Homosexualität mit IAT, EAST, den Motivationsmaßen und den Moderatortermen wird keine statistisch bedeutsame Steigerung der Varianzaufklärung durch die Moderatorterme erhalten (Tab. 4-23).

	kognitive Einstellungen (HOKO)						affektive Einstellungen (HOAF)							
	R <sup>2</sup>	ΔR <sup>2</sup>	p	Beta	T	p	R <sup>2</sup>	ΔR <sup>2</sup>	p	Beta	T	p		
<b>IAT IMS</b>	.34	.01	.554	I	.39	2.93	.005	.19	.01	.414	I	.40	2.62	.012
				Mo	-.41	-3.16	.003				Mo	-.22	-1.51	.139
				I*Mo	.08	.60	.554				I*Mo	.13	.82	.414
<b>IAT EMS</b>	.25	.05	.086	I	.51	3.81	.000	.18	.01	.539	I	.41	2.89	.006
				Mo	.14	1.06	.297				Mo	-.11	-.78	.441
				I*Mo	.24	1.75	.086				I*Mo	.09	.62	.539
<b>IAT MVV</b>	.20	.01	.554	I	.43	3.17	.003	.21	.05	.089	I	.42	3.03	.004
				Mo	-.07	-.54	.589				Mo	-.03	-.24	.813
				I*Mo	.08	.60	.554				I*Mo	.24	1.74	.089
<b>EAST IMS</b>	.31	.01	.409	I	.30	2.38	.022	.21	.00	.624	I	.38	2.87	.006
				Mo	-.40	-3.16	.003				Mo	-.22	-1.66	.104
				I*Mo	-.10	-.83	.409				I*Mo	.07	.49	.624
<b>EAST EMS</b>	.15	.01	.483	I	.36	2.66	.011	.18	.00	.646	I	.40	2.96	.005
				Mo	.09	.62	.537				Mo	-.12	-.92	.365
				I*Mo	.10	.71	.483				I*Mo	-.06	-.46	.646
<b>EAST MVV</b>	.15	.01	.455	I	.34	2.43	.019	.17	.00	.757	I	.40	2.85	.007
				Mo	-.07	-.51	.614				Mo	-.03	-.20	.845
				I*Mo	.10	.75	.455				I*Mo	.04	.31	.757

**Tabelle 4-23**, Moderatorwirkung von Motivation zur Vorurteilsfreiheit auf implizit-explizit-Korrespondenz; I= implizites Maß, Mo=Moderator IVMF, I\*Mo= Produktterm der Moderation; ΔR<sup>2</sup>= Änderung in R<sup>2</sup> durch Moderatorterm I\*M

Den Regressionsanalysen gemäß sind die impliziten Maße durchweg signifikante und in der Mehrzahl der Fälle auch alleinige Prädiktoren expliziter Einstellungen. Die Zusammenhänge zwischen impliziten und expliziten Einstellungen sind dabei stabil positiv gerichtet. Die Internale Motivation zur Vorurteilsfreiheit (IMS) ist das einzige Motivationsmaß, für das in der Vorhersage kognitiver Einstellungen zu Homosexualität ein signifikantes Regressionsgewicht geschätzt wird, und zwar unabhängig davon, ob der IAT oder der EAST als impliziter Prädiktor betrachtet wird. Die negative Pfadladung zeigt an,

dass mit zunehmender Internaler Motivation zur Vorurteilsfreiheit positivere kognitive Einstellungen zu Homosexualität bestehen. Auch wenn die Motivationsmaße nur in diesen beiden beschriebenen Fällen statistisch bedeutsam sind, werden in zehn der zwölf Regressionsmodelle erwartungskonform negative Vorzeichen für die Motivation zur Vorurteilsfreiheit geschätzt. Das bedeutet, stärker vorurteilsfrei motivierte Probanden kontrollieren ihre expliziten Einstellungen mehr, so dass sie weniger Vorurteile als weniger motivierte Probanden darin zum Ausdruck bringen. Die Moderatorterme sollten entsprechend negative Vorzeichen haben, da Probanden, die wenig motiviert sind, eine höhere explizit-implizit-Korrespondenz aufweisen sollten, als Probanden, die stark motiviert sind und entsprechend in ihren expliziten von ihren impliziten Einstellungen abweichen. Dies ist jedoch nur in 2 von 12 Modellen der Fall. Auch hier muss betont werden, dass die Moderatorterme durchweg nicht signifikant sind, dennoch weisen sie zum Teil recht hohe Beta-Gewichte auf. Im Modell mit IAT und EMS zur Vorhersage kognitiver Einstellungen beträgt das standardisierte Regressionsgewicht für den Moderatorterm  $IAT*EMS$  .24 ( $p < .09$ ). Ähnliches gilt für die Vorhersage affektiver Einstellungen durch IAT und MVV, hier wird ebenfalls ein Regressionsgewicht von .24 ( $p < .09$ ) geschätzt.

## 4.5 Diskussion

Studie I war angelegt, die Effekte einer konstanten Stimulusabfolge für alle Probanden auf Reliabilität und Validität von IAT und EAST zu testen. Den Basisanalysen zufolge liegen deutliche Unterschiede in den psychometrischen Eigenschaften beider impliziter Einstellungsmessungen vor. Für den IAT-Effekt wird mit einem Cronbach's Alpha von .87 eine vergleichsweise hohe interne Konsistenz erhalten. Die Zusammenhänge zwischen IAT und expliziten Vorurteilen sind signifikant und erwartungskonform positiv. Höhere Reaktionszeiteffekte im IAT gehen mit negativeren expliziten Einstellungen zu Homosexualität einher sowie mit höheren Autoritarismus- und Modern-Racism-Werten. Für die Fehlerbasierten IAT-Effekte sind die Korrelationen zu kognitiven und affektiven Einstellungen ebenfalls positiv. Sie erreichen jedoch nicht das 5%-Signifikanzniveau. Für den IAT kann also zusammenfassend festgehalten werden, dass die Differenz zwischen den Reaktionszeiten des inkompatiblen und des kompatiblen eine intern konsistente und nach den verfügbaren Kriterien zu urteilen ebenfalls valide implizite Einstellungsmessung liefert.

Weniger widerspruchsfrei ist die Ergebnislage für die Daten des EAST. Die vier Messblöcke verhalten sich sowohl untereinander, als auch zu den Außenkriterien nicht konsistent. Die interne Konsistenz beträgt lediglich .20 (Reaktionszeiten) und .38 (Fehlerraten). Das bedeutet, der EAST liefert keinen reliablen Gesamtscore, sondern sollte auf der Basis der Blockeffekte analysiert werden. In vorliegender Studie scheint der erste Messblock aufgrund seines systematischen Zusammenhangsmusters zur Ge-

genüberstellung mit IAT und Selbstbericht ein valides Effektmaß zu liefern. Wobei einerseits die allgemein akzeptable Lösung nicht in einem post-hoc-Vorgehen zur Auswahl eines „funktionierenden“ EAST-Effekts liegen kann und andererseits nicht klar ist, ob ein Zusammenhang zu expliziten Einstellungen das Kriterium sein sollte. Im Hinblick auf die geringe Reliabilität des EAST muss an dieser Stelle jedoch beachtet werden, dass sich die Probanden hinsichtlich der Bedeutung des relevanten Merkmals der Zielkonzeptstimuli (Farbe) in Studie I unterscheiden. Während für die Hälfte der Probanden Blau die „positive“ und Grün die „negative Farbe“ darstellte, verhielt es sich für die andere Hälfte der Probanden umgekehrt. Die resultierenden Stichproben sind mit  $N=25$  relativ klein und anfälliger für Zufallsvariationen. Auch wenn keine systematischen Unterschiede zwischen beiden Farb-Valenz-Gruppen beobachtet wurden, mindert diese Variation die uneingeschränkte Vergleichbarkeit der Probanden, was sich ebenfalls negativ auf die Reliabilität auswirken könnte.

Wie der Zusammenhangsanalyse zwischen Reaktionszeit- und Fehlerdaten des EAST abzulesen ist, besteht hier kein linearer, sondern bestenfalls ein umgekehrt u-förmiger Zusammenhang. Das bedeutet offensichtlich, dass Fehler- und Reaktionszeiteffekte im EAST nicht gleichermaßen die impliziten Valenzassoziationen zu Homo- bzw. Heterosexualität widerspiegeln. Entsprechend ist bei der Betrachtung der Reaktionszeiteffekte der *erste* EAST-Block als Maß mit statistisch bedeutsamen und erwartungskonform positiven Zusammenhängen zu beiden expliziten Einstellungsmaßen (kognitiv und affektiv) einzustufen. Für die Fehlerraten trifft dies jedoch auf den *dritten* Block zu, wobei die ebenfalls positive Korrelation zu den kognitiven Einstellungen zu Homosexualität nicht signifikant ist. Eine Ursache für die Divergenz von Reaktionszeit- und Fehlereffekten könnte die Schwierigkeit der Aufgabenstellung im EAST sein. De Houwer (2003) findet in einem Experiment (Blumen-Insekten-EAST) ebenfalls keine signifikante Korrelation zwischen Fehler- und Reaktionszeiteffekten des EAST, für ein zweites Experiment (Selbstkonzept-EAST) dagegen berichtet der Autor positive Korrelationen für drei von fünf Stimulusarten. Diese relative Unkorreliertheit von Fehler- und Reaktionszeiteffekten trifft, um es noch mal direkt zu vergleichen, für die IAT-Daten nicht zu, diese stehen in linear positivem Zusammenhang ( $r=.35$ ). Ein höherer IAT-Effekt in den Reaktionszeitdaten geht entsprechend mit einem höheren IAT-Effekt in den Fehlerdaten einher. Der IAT ist also insgesamt deutlich reliabler, als der EAST. Hierbei ist jedoch zu beachten, dass der IAT-Effekt jeweils auf Blockmittelwerten beruht, während für die EAST-Effekte lediglich die inhaltlich relevanten Stimuli verrechnet werden. Damit umfasst der EAST-Effekt weitaus weniger Trials, als der IAT-Effekt, was bereits Teige et al. (2004) als geringes Aggregationsniveau im EAST kritisieren. Eine augenscheinlich einfache Lösung wäre die Durchführung mehrerer Trials. Um jedoch auf ein mit dem IAT vergleichbares Aggregationslevel zu gelangen, müssten in einem EAST-Block doppelt so viele Trials erhoben werden, wie in kompatibel und inkompatibel Block des IAT zusammen. Dies sei kurz an einem Rechenbeispiel veranschaulicht: Ein IAT mit jeweils 40 Trials in den beiden kritischen Messblöcken liefert ein auf

insgesamt 80 Trials basierendes Effektmaß. Um eine ebensolche Anzahl an Einzelmessungen für den EAST-Effekt verrechnen zu können, müssten neben den 80 Zielkonzepttrials weitere 80 evaluative Trials für die Sicherung der extrinsischen Valenzbelegung der Antworttasten dargeboten werden, die nicht in den EAST-Effekt eingehen, insgesamt also 160 Trials. Die Beantwortung von 160 Trials in einem Block, das heißt ohne Pause, wird jedoch einerseits für die Probanden als unzumutbar eingeschätzt und aus diesem Grund auch inhaltlich angezweifelt.

Von einer durchgängigen Korrespondenz zwischen den beiden impliziten Maßen kann nicht ausgegangen werden. Jedoch bestehen zwischen dem Reaktionszeiteffekt des ersten EAST-Blocks und dem IAT-Effekt eine signifikant positive Korrelation, was mit Vorsicht zu interpretierende Hinweise auf Zusammenhänge beider impliziter Maße liefert. Für den ersten EAST-Block sind nicht nur in Bezug auf den IAT, sondern auch die expliziten Einstellungen erwartungskonforme Ergebnisse zu registrieren (siehe unten), während die restlichen drei EAST-Blöcke keine konsistenten Beziehungsmuster liefern. Mögliche Ursachen für die Überlegenheit des ersten EAST-Blocks können allgemein verminderte Einflüsse von Lerneffekten, Strategienwechsel, Ermüdungserscheinungen oder auch spezifische Effekte der Stimulusabfolge sein. Die Anzahl an Blöcken sollte demzufolge nach nicht das ausschlaggebende Kriterium für hohe interne Konsistenz sein. Die Inkonsistenz der Messblöcke veranschaulicht zugleich, wie sensibel bzw. instabil der EAST im Ganzen ist.

Die gemeinsame Betrachtung der Reaktionszeiteffekte von IAT und EAST, wobei hier wiederum ausschließlich der erste Teileffekt einbezogen werden kann, zeigt, dass der IAT vornehmlich zur Vorhersage kognitiver Einstellungen geeignet ist, während der EAST insbesondere die affektiven Einstellungen vorhersagt. Das deutet darauf hin, dass vor allem der IAT kognitive Assoziationsstärken erfasst, der EAST dagegen möglicherweise stärker valenzbezogene Assoziationen. Mit den impliziten Maßen werden 23% bzw. 22% Varianz kognitiver bzw. affektiver Einstellungen aufgeklärt. Werden die expliziten Einstellungsprädiktoren einbezogen, die allein 62% an kognitiven und 49% an affektiven Einstellungen zu Homosexualität aufklären, erweist sich allein der EAST auch für kognitive Einstellungen als inkrementell valide. Die zusätzliche Varianzaufklärung beträgt 11% für kognitive und 15% für affektive Einstellungen zu Homosexualität. Der IAT sagt dabei jedoch zusätzlich zu den expliziten Prädiktoren und dem EAST keinen signifikanten Anteil Varianz an den expliziten Einstellungen vorher. Dies ist wahrscheinlich auf die gemeinsame Varianz, die IAT und Autoritarismus teilen, zurückzuführen. Autoritarismus stellt den stärksten Prädiktor kognitiver Einstellungen zu Homosexualität dar, steht in positiver Beziehung zum IAT ( $r=.40$ ), ist jedoch unkorreliert mit dem EAST ( $r=.04$ ). Die standardisierten Regressionsgewichte von IAT und EAST zur Vorhersage affektiver Einstellungen zu Homosexualität unterscheiden sich dagegen betragsmäßig nicht sehr stark voneinander.

Moderatoreffekte der Beziehungen zwischen impliziten und expliziten Einstellungen durch die Motivation zur Vorurteilsfreiheit bestehen im vorliegenden Fall weder für den IAT, noch für den EAST.



Die interne Motivation, ohne Vorurteile zu sein, stellt von den drei betrachteten Einstellungsmaßen den stärksten Prädiktor expliziter, insbesondere kognitiver Einstellungen dar, wobei durchweg eine negative Beziehung zwischen beiden Maßen besteht. Möglicherweise repräsentiert das Motivationsmaß bereits selbst eine Art Einstellungsmaß, das aufgrund der gemeinsamen methodischen Varianz des Selbstberichts in engerer Beziehung zu expliziten als zu impliziten Einstellungen steht. Die Zusammenhänge zwischen dem Motivationsmaßen und den beiden IAT-Effekten (Reaktionszeiten und Fehler) sind bspw. allesamt negativ und rangieren von  $-.09$  (EMS) bis  $-.23$  (MVV). Bereits Devine (1989) und Devine et al. (2002) zeigen, dass insbesondere eine interne Motivation zur Vorurteilsfreiheit Einfluss auf den IAT-Effekt hat. Zwischen dem EAST und den Motivationsmaßen bestehen dagegen weniger konsistente Zusammenhänge, so dass sowohl negative, als auch positive Vorzeichen die drei signifikanten Korrelationen kennzeichnen. Da der EAST als Messung weniger durchschaubar und zugleich schwieriger als der IAT ist, hat die Motivation zur Vorurteilkontrolle hierfür möglicherweise weniger Bedeutung.

Auch wenn in den hier verfolgten Fragestellungen ein interindividueller Fokus anliegt, sind einige der Ergebnisse auf der Ebene von Stichprobenmittelwerten zu diskutieren. Der Befund abnehmender Reaktionszeiten über die EAST-Blöcke hinweg für die Reaktion auf heterosexuelle Paarbilder (für den IAT ist eine solche stimulusspezifische Analyse unüblich) könnte darauf hinweisen, dass gemischtgeschlechtliche Paare prinzipiell leichter zu entdecken sind, als die gleichgeschlechtlichen männlichen Paare. Madigan (1983) zeigt, dass die notwendige Zeitdauer zum Erkennen eines Bildes im zweistelligen Millisekundenbereich liegt und damit deutlich unter der normalen Fixationsdauer eines Bildes (150 bis 300ms) bzw. der Darbietungszeit der Stimuli im Experiment (ca. 600ms). Zwar ist gemäß EAST-Instruktion keine inhaltliche Verarbeitung der Zielkonzeptstimuli vonnöten, da diese allein auf der Basis ihrer Färbung beantwortet werden sollen, dennoch zeugen die signifikant kürzeren Reaktionszeiten für heterosexuelle Stimuli davon, dass eine reine Oberflächenauswertung nach Farbe (perzeptuelle Verarbeitung) nicht ohne inhaltliches Erkennen (konzeptuelle Verarbeitung) stattfindet (Theorie der dualen Kodierung, siehe Paivio, 1991). Auf diesem Konzept basiert bspw. auch die *Affective Simon Task* (De Houwer & Eelen, 1998), mit der der EAST strukturell verwandt ist (vgl. De Houwer, 2003a).

Ein weiterer Unterschied zwischen homo- und heterosexuellen Stimuli liegt in der kritischen Interaktion zwischen Stimulusart und extrinsischer Valenz. Diese wurde nur für die heterosexuellen Stimuli gefunden. Für gemischtgeschlechtliche Paare machte es einen entscheidenden Unterschied für die Reaktionszeiten, ob gleichgeschlechtliche Paare in der positiven oder der negativen Farbe dargeboten wurden. Erwartungsgemäß wurden längere Reaktionszeiten bei den einstellungsinkongruenten Stimuli, also heterosexuell und negative Farbe, erhalten. Für die überwiegende Mehrzahl der Untersuchungsteilnehmer stellen heterosexuelle Paare die Eigengruppe hinsichtlich sexueller Orientierung dar. Es handelt es sich bei diesem Befund also primär um die Bevorzugung der Eigengruppe Heterosexueller,

als eine Ablehnung der Fremdgruppe Homosexueller in der Stichprobe. Zwar steht die fehlende Interaktion für homosexuelle Stimuli der Intention des EAST auf Stichprobenebene entgegen, im vorliegenden Fall sind jedoch weniger die Mittelwerte von Interesse, als vielmehr die individuell unterschiedlichen Varianzen und Kovarianzen der Einstellungsmaße. Hinsichtlich dieser differentiellen Validität konnte zumindest für den ersten EAST-Block eine Bedeutung gezeigt werden, wobei eine Grundvoraussetzung für die Untersuchung interindividueller Differenzen jedoch eine hinreichend gute Reliabilität ist.

Die Ergebnisse von Studie I resümierend, resultiert aus der fixierten Stimulusabfolge für den IAT eine relativ hohe interne Konsistenz, wobei der EAST nicht deutlich reliabler ist, als bei De Houwer (2003a). Dies könnte mit an der experimentell variierten Farbbedeutung für die Probanden liegen. Da hierfür keine bedeutenden Unterschiede gefunden wurden, ist im Weiteren von einer solchen Variation abzusehen. Die geringe Reliabilität des EAST ist recht unbefriedigend, jedoch wirkt das Potential des *ersten* EAST-Blocks zur Vorhersage kognitiver und affektiver Einstellungen zu Homosexualität ermutigend. Da das EAST-Maß basierend auf dem ersten Messblock weniger von Lerneffekten, Ermüdungserscheinungen oder veränderten Antwortstrategien beeinflusst sein sollte, erscheint dieser Befund sinnvoll. Dies kann jedoch nur dann verallgemeinert werden, wenn sich die Bedeutung des ersten Blocks im EAST replizieren lässt (siehe Studie III).

Die alleinige Erfassung eines Einstellungsinhaltes mit impliziten und expliziten Maßen ermöglicht keine Trennung der Varianzanteile, die auf die betrachtete Einstellungsdisposition bzw. die Messmethodik zurückzuführen ist. Mit 50 Untersuchungsteilnehmern umfasst Studie I zudem zwar eine vergleichsweise akzeptable Stichprobengröße, dennoch erlaubt ein N von 50 keine messfehlerbereinigte Analyse der Daten mit latenter Strukturgleichungsmodellierung. Um eine latente Konstruktvalidierung, also die Untersuchung konvergenter und diskriminanter Validität, expliziter und impliziter Einstellungen vornehmen zu können, ist einerseits die Erhebung von mindestens zwei inhaltlichen Einstellungsdispositionen und andererseits eine genügend große Stichprobe ( $N > 150$ ) vonnöten. Zur Untersuchung des Zusammenhangs expliziter und impliziter Einstellungen im Sinne einer Explizit-implizit-Dimension wird eine weitere Messung entwickelt, die sowohl spontane, unbewusste ergo implizite, als auch reflektierte, bewusste ergo explizite Einstellungsanteile erfasst. Die sinnbildliche Zwischenstellung dieser Messung sollte sich in jeweils stärkeren Beziehungen zu expliziten sowie impliziten Einstellungen äußern, als explizite und implizite Einstellungen untereinander unterhalten. Zu diesem Zweck wird eine offene Assoziativmessung mit verbalem Antwortformat gewählt, die in der Methodik an den Thematischen Apperzeptionstest (TAT, Murray, 1943) angelehnt ist. Der Entwicklung dieses Messverfahrens ist die im Folgenden berichtete Studie II gewidmet.

## **5 Studie II: Entwicklung einer offenen Assoziativmessung zur Erfassung expliziter & impliziter Einstellungsanteile**

### **5.1 Theoretische Grundlagen und Anliegen von Studie II**

Die Einstellungsforschung widmet sich nach wie vor der Untersuchung expliziter und impliziter Einstellungen hinsichtlich ihrer Bedeutung und ihrer Konzeptualisierung. Als ein Bindeglied der Korrespondenz expliziter und impliziter Einstellungen sind u.a. einstellungsbezogene Motivationen, zu denen auch die allgemeine soziale Erwünschtheit zu zählen ist, angeführt worden. Diese beeinflussen insbesondere bewusste, kontrollierbare Kognitionen, also expliziten Einstellungen. Es konnte gezeigt werden, dass eine hohe Motivation zur Vorurteilsfreiheit mit einer geringeren Korrespondenz von Selbstberichten und Assoziationsstärkemaßen in Beziehung steht (Banse & Gawronski, 2003; Crandall, Eshelman & O'Brien, 2002; Devine, 1989; Devine et al., 2002; Dunton & Fazio, 1995). Problematisch dabei ist die explizite Messung der Motivation zur Vorurteilsfreiheit bzw. sozialen Erwünschtheit. McClelland (1985) kritisiert die, im Zug der allgemeinen Betonung kognitiver Prozesse in der Psychologie vorgenommene kognitive Prägung des Motivationsbegriffs bspw. durch Atkinson & Feather (1966), was dazu geführt hat, dass unter Motivation schlichtweg alles gefasst wurde, was eine Reaktions tendenz beeinflusst, die Erwartungen, der Anreiz, die Valenz, sogar die Fähigkeit. Daraus resultierte eine undifferenzierte und ungerechtfertigte Verallgemeinerung des Motivationskonzepts. Gleichsam fand in der Motivmessung eine Wende weg vom Einsatz indirekter Messmethoden, wie dem Thematischen Apperzeptionstest (TAT), hin zur direkten Erfassung von Motiven, bspw. über Selbstberichte statt (z.B. Weiner, 1980). McClelland zeigt jedoch eindrucksvoll die notwendige Differenzierung zwischen der eigentlichen Motivstärke und dem Anreizwert, anhand der Vorhersage von Verhalten. Der Autor stellte die Bedeutung mit dem TAT gemessener, sozialer Motive, neben biologischen Trieben wie Hunger für spezifische physiologische Prozesse heraus, während er zugleich zeigte, dass selbst berichteten Motive an dieser Stelle keinerlei Bedeutung zukam. Zudem weist er auf die inkrementelle Validität von Motiven (nach McClelland via TAT erfasst), Werten und Fähigkeiten für reflektiertes vs. unreflektiertes Verhalten (z.B. sozialer Kontakt) nach (McClelland, 1985).

Hier lassen sich deutliche Parallelen zu den Konzepten expliziter und impliziter Einstellungen erkennen. Die Unterscheidung zwischen spontanen, unbewussten, unreflektierten einerseits und überlegten, bewussten und reflektierten Einstellungs- oder Motivationsinhalten andererseits ermöglicht eine bedeutend präzisere und zudem differenziertere Verhaltensvorhersage. Es liegt nahe, die TAT-Methodik zur Messung einstellungsrelevanter Motive und Kognitionen einzusetzen und entsprechende Maße den expliziten und impliziten Einstellungen sowie der expliziten Motivation zur Vorurteilsfreiheit gegenüber zustellen. Dazu bedarf es einer Adaptation des klassischen Verfahrens.

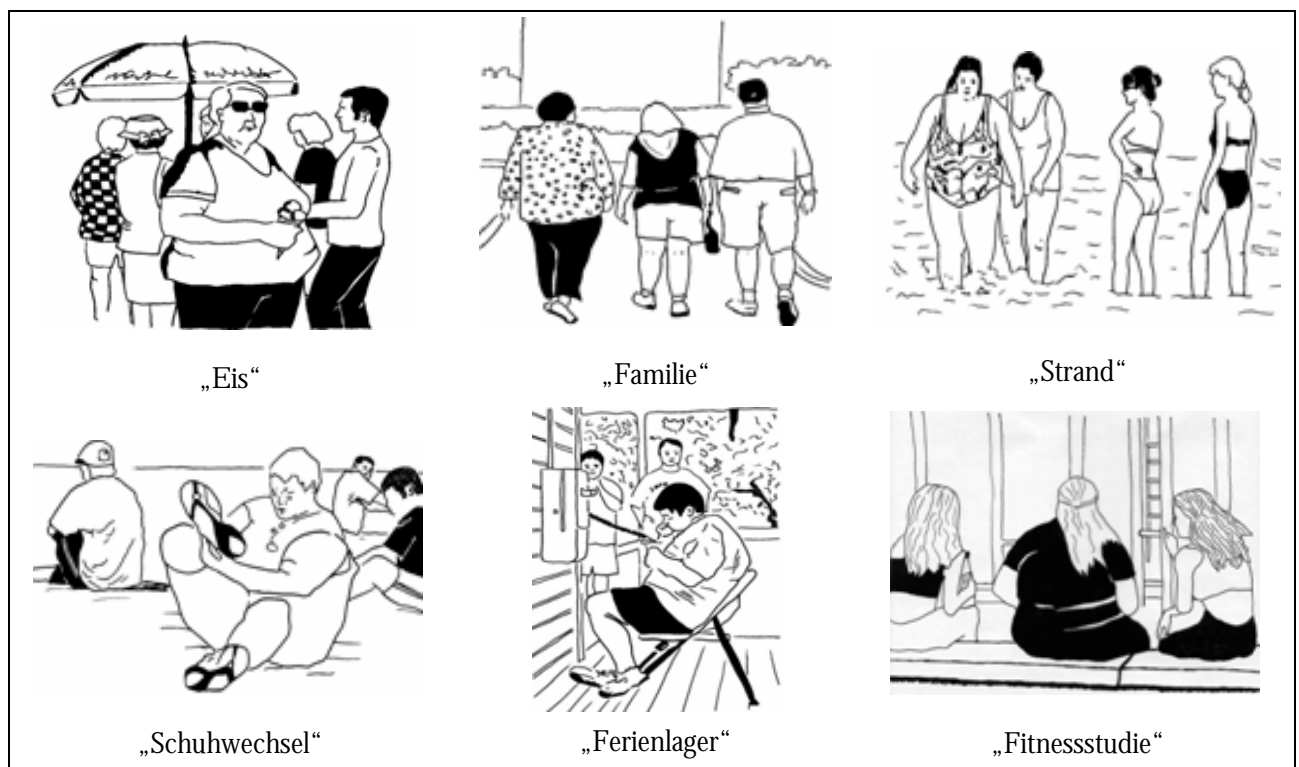
Studie II dient entsprechend der Entwicklung und Erprobung eines Verfahrens zur Messung expliziter und impliziter Einstellungen, das einige messmethodische Aspekte des TATs aufgreift. In Anlehnung an den TAT zur Messung impliziter Motive wird für die Einstellungsmessung das mehrdeutige bildhafte Stimulusformat und die Verfahrensinstruktion einer zum Bildstimulus passenden, jedoch völlig frei fantasierten Geschichte übernommen. Für die beiden ausgewählten Einstellungsinhalte Homosexualität und Übergewicht relevante Bildmotive sollten spontane Assoziationen induzieren, die in eine selbst verfasste Geschichte ebenso wie überlegte und bewusst kontrollierte Kognitionen einfließen. Dies wird durch die Uneindeutigkeit der Messintention und das offene, verbale Antwortformat unterstützt. Für eine entsprechende Messung bedarf es zunächst der Neuentwicklung geeigneter Bildmotive, wobei es insbesondere darauf ankommt, bei maximal sich öffnendem Interpretationsspielraum zugleich die sichere Aktivierung des betreffenden Stereotyps bei den Probanden zu gewährleisten. Dies ist insbesondere für das Einstellungsobjekt Homosexualität von großer Bedeutung, da die sexuelle Orientierung der Menschen nicht äußerlich erkennbar ist. Die Darstellung von Personen mit stereotypem „schwulen“ Äußeren würde bspw. lediglich die Aktivierung *eines* spezifischen Substereotypen (Bsp. feminine Schwule) provozieren, was eine Einschränkung des Einstellungsspektrums zur Folge hätte. Eine andere Variante ist die Abbildung eines gleichgeschlechtlichen romantischen Paares. Auch die Formulierung der Bildkarteninstruktion könnte einer Aktivierung des Stereotyps dienen. Prinzipiell ist eine möglichst abstrakte Abbildung von Personen zu wählen, da die Individualität von Gruppenmitgliedern bestimmte Bewertungen über den Stereotyp hinaus evozieren kann (Blair, 2002). Neben dem zu erstellenden Bildmaterial sind spezifische Kodierschemata für die Geschichten zu entwickeln, da die Messung zunächst qualitatives Datenmaterial liefert. Für die Inhaltsanalyse der Geschichten sind einstellungs- und motivationsrelevante Variablen- und Kategoriensysteme festzulegen, die es ermöglichen aus der Vielfalt an Informationen kognitive, emotionale und motivationale Aspekte von Einstellungen herauszufiltern.

## **5.2 Entwicklung der offenen Assoziativmessung**

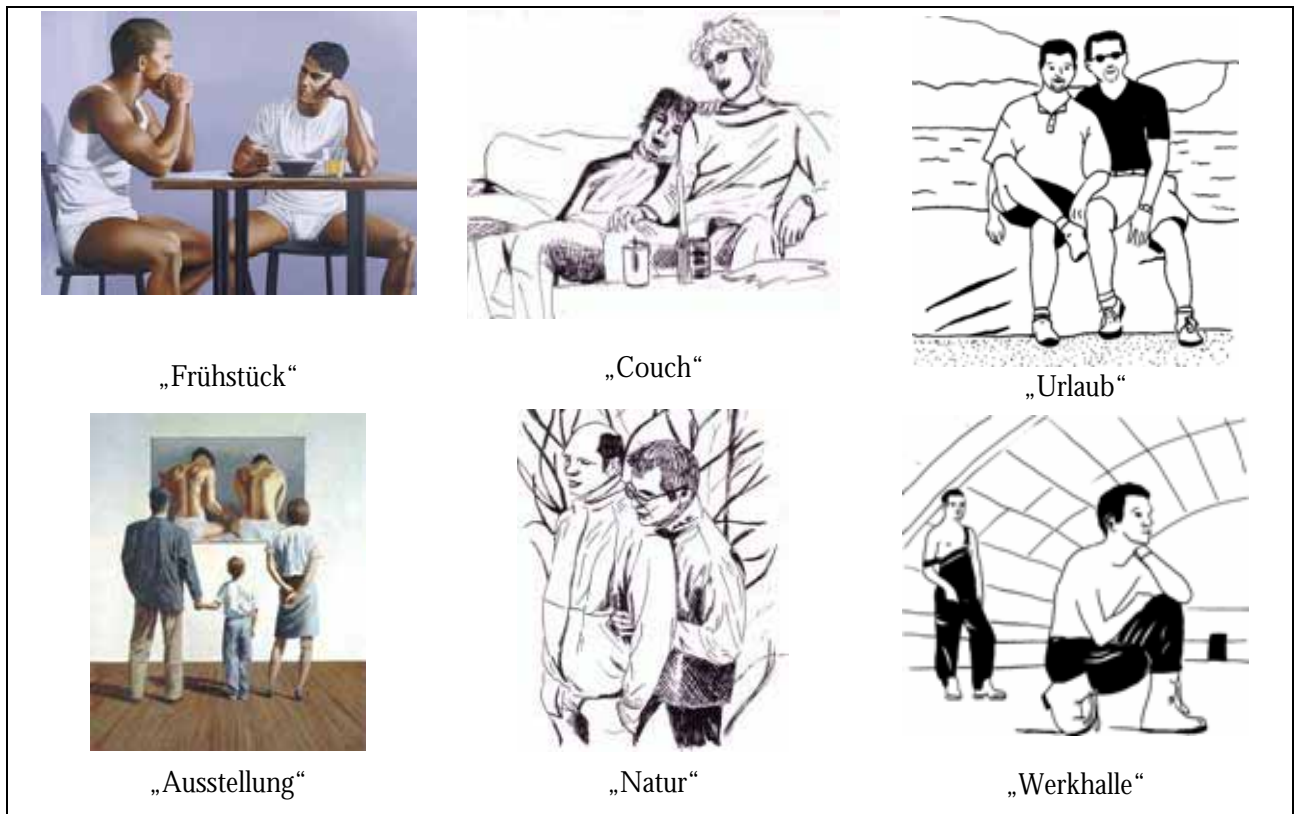
### **5.2.1 Bildmaterial**

Grundvoraussetzung eines TAT-ähnlichen Verfahrens zur Erfassung von Einstellungsaspekten und einstellungsbezogener Motivation sind geeignete Bildmotive. Als Bildmaterial für die Messung sollten schematische Darstellungen eingesetzt werden, deren Komplexitätsgrad gering genug ist, genügend Interpretationsspielraum zu bieten. Besonderes Gewicht bei der Entwicklung von Bildstimuli lag auf dem Einstellungsobjekt *Homosexualität*, da sich homosexuelle Menschen im Gegensatz zu übergewichtigen nicht zwangsläufig durch eindeutige äußere Merkmale auszeichnen. Zur Entwicklung geeigneter

Bildstimuli wurde zunächst ein Pool an Bildmaterial für beide Einstellungsobjekte mit Hilfe des Internets und privater Fotografien zusammengestellt. Diese Grundgesamtheit an Bildmotiven wurde in einer Expertengruppe auf 11 Bildern für das Einstellungsobjekt „Homosexualität“ und 9 Bilder für „Übergewicht“ reduziert. Die ausgewählten Bilder wurden anschließend einer ersten Voruntersuchung unterzogen. Zu jedem Bild wurden spontane Assoziationen verschiedener Personen eingeholt. Gemäß der Sammlung dieser offenen Antworten wurde in einem weiteren Schritt je Einstellungsobjekt eine Rangliste der Bilder nach ihrer Eignung für die Erfassung von Einstellungen erstellen. Ein Kriterium war die Häufigkeit von stereotypen Nennungen, woraus insbesondere für den Stereotyp Homosexualität die kritische Fähigkeit, den Stereotypen zu aktivieren, abgeleitet wurde. Darüber hinaus wurden Bilder, die sehr verschiedene Assoziationen induzierten, besser bewertet, als Bilder, für die jede befragte Person ähnliche Assoziationen hervorbrachte. Schließlich wurden die jeweils sechs ranghöchsten Bilder pro Einstellungsobjekt ausgewählt und in Anlehnung an den klassischen TAT bearbeitet und somit einander etwas angeglichen (siehe Abb. 5-1, 5-2). Die Originalfotografien wurden in einfache Zeichnungen schematisiert und wenn nötig auf eine Schwarzweißtönung reduziert, um für alle Bilder eine Reduzierung des Informationsgehalts zu gewährleisten.



**Abbildung 5-1,** Bildmotive und Bildbezeichnungen für das Einstellungsobjekt Übergewicht



**Abbildung 5-2,** Bildmotive und Bildbezeichnungen für das Einstellungsobjekt Homosexualität

### 5.2.2 Aufbau der Kodierschemata

Tomkins (1947) fordert für die Auswertung des klassischen TATs: „Das Kategoriensystem müsse als Ganzes auch die Totalität des untersuchten Gegenstandes zugänglich machen“ (S. 26). Der untersuchte Gegenstand im vorliegenden Fall sind Einstellungen und einstellungsrelevante Motive zu homosexuellen und übergewichtigen Menschen, die ihren Niederschlag in selbst verfassten Geschichten zu den ausgewählten Bildmotiven finden sollten. Eine umfassende Betrachtung des Gegenstandes schließt also stereotypische Kognitionsinhalte und Assoziationen, emotionale, wertbezogene und motivationale Aspekte ein. Um das verbale Datenmaterial im Sinne der empirischen Fragestellung verwerten zu können, wird für jedes Bild zunächst ein Kodierschema entwickelt, das sich aus einem bildspezifischen und einem für alle Motive allgemeingültigen Teil zusammensetzt (vgl. Tab. 5-1). Spezifisch für jedes Bildmotiv sind die durch die aktivierten Stereotypeninhalte. Hierbei variiert die Anzahl kodierter Inhaltsthemen für die insgesamt zwölf Bilder zwischen vier und sieben.

Der für alle Bilder gültige Teil eines Kodierschemas umfasst die fünf formellen Variablen *Valenz*, *Grundstimmung*, *Ernsthaftigkeit*, *Stereotypenerkennung*, *Wortanzahl*. Mit *Valenz* wird die affektive Konnotation, also die Wertigkeit der Geschichte, die sich beispielsweise in der Wortwahl widerspiegelt, erfasst. Die Variable rangiert von positiv (1) bis negativ (5). Die *Grundstimmung* ermöglicht eine von der allge-

meinen Valenz gesonderte Betrachtung des Ausgangs der Geschichte, wobei als Kategorien optimistische (1), ausgewogene (2) und pessimistische (3) Stimmung festgelegt wurden. Zwar gehen Valenz und Grundstimmung theoretisch eng miteinander einher, dennoch kann eine Geschichte als neutral oder gar positiv valent zu bewerten sein und dabei mit einer einzelnen Wendung einen negativen Ausblick erhalten. Die Variable *Ernsthaftigkeit* erfasst den Grad der der Ernstgemeinheit einer geschriebenen Geschichte, also inwiefern die Geschichte Ernst oder Ironie, Zynismus, Spott bzw. Sarkasmus widerspiegelt. Diese Variable wird als motivationaler Indikator im Sinne einer Abwehr der Einstellungsthematik Übergewicht bzw. Homosexualität interpretiert. Ähnlich verhält es sich mit der Variablen *Stereotypenerkennung*, die von „nicht erkannt“ (0) über „implizit erkannt“ (allgemeine stereotypenrelevante Themen/ Bezeichnungen verwendet) bis zu „explizit erkannt“ (Stereotypenlabel verwendet) variiert. Hiermit wird erfasst, ob eine Geschichte Hinweise darauf enthält, dass das entsprechende Einstellungsobjekt vom Probanden sicher erkannt wurde. Besondere Relevanz hat diese Variable für den Stereotyp Homosexualität, da es hierbei unvergleichbar schwerer ist, eine Art der bildhaften Darstellung zu finden, die bei eindeutiger Stereotypenaktivierung noch genügend Interpretationsspielraum lässt. Im Gegensatz zu übergewichtigen sind homosexuelle Menschen durch keine besonderen äußerlichen Merkmale charakterisiert. Zugleich sollte sich in dieser Variablen eine Motivation zur Vorurteilsfreiheit ausdrücken. Mit *Wortanzahl* wird für den Umfang einer Geschichte, der maßgeblichen Einfluss auf die Anzahl genannter Themen haben sollte, und zudem einen weiteren motivationalen Aspekt aufgreift, kontrolliert.

<b>Variablenlabel</b>	<b>Wertelabel</b>
<i>Inhalt</i> Spezifische Themen (siehe Tab. 5-2, 5-3)	0= ‚Thema wird genannt‘ 1= ‚Thema wird nicht genannt‘
<i>Valenz</i>	1= ‚positiv‘ 2= ‚tendenziell positiv‘ 3= ‚neutral‘ 4= ‚tendenziell negativ‘ 5= ‚negativ‘
<i>Grundstimmung</i>	1= ‚optimistisch/positiv‘ 2= ‚ausgewogen/neutral‘ 3= ‚pessimistisch/negativ‘
<i>Ernsthaftigkeit</i>	0= ‚Geschichte ist spöttisch, ironisch, zynisch oder sarkastisch‘ 1= ‚Geschichte ist ernst zu nehmen‘
<i>Stereotypenerkennung</i>	0= ‚nicht erkannt‘ 1= ‚implizit erkannt (Verwendung allg. stereotypenrelevanter Themen)‘ 2= ‚explizit erkannt (Verwendung des Stereotypenlabel oder Synonyme)‘
<i>Wortanzahl</i>	

**Tabelle 5-1**, Allgemeiner Aufbau eines Kodierschematas

Die spezifischen Inhaltsvariablen für jedes Bilder werden indikatorcodiert (1= 'Thema genannt', 0= 'Thema nicht genannt'). Sie dienen einerseits als Kontrollvariablen für die gelungene Aktivierung des Stereotyps, z.B. bei der Verwendung vieler stereotyper Themen in der Geschichte, bzw. für eine Abwehr der Thematik bei entsprechend ausbleibender Einbeziehung stereotyper Themen. Allein mittels dieser Inhaltsvariablen kann jedoch nicht eindeutig festgestellt werden kann, ob ein minimales Vorkommen stereotypenrelevanter Themen in einer Geschichte auf die Abwehr des Stereotypen oder eine Nichterkennung zurückzuführen ist. Es gilt also zu prüfen, welche formellen Variablen zur Klärung dieser Frage beitragen. Andererseits reflektieren die Inhaltsvariablen in gewissem Maße die Kreativität eine Probanden. Die Inhaltsvariablen werden je Bild zu einem Inhaltsscore gemittelt. Entsprechend zeugen hohe Werte von einer Verwendung vieler Stereotypeninhalte in einer Geschichte. Niedrige Werte können sowohl die misslungene Stereotypenaktivierung, als auch die bewusste oder unbewusste Abwehr der Vorurteilsthematik reflektieren. Die Variable „Inhalt“ kann als Kriterium der stereotypenrelevanten Reichhaltigkeit einer Geschichte interpretiert werden. Die folgenden beiden Tabellen veranschaulichen die bildspezifischen Abschnitte der Kodierschemata für Übergewicht (Tab. 5-2) und für Homosexualität (Tab. 5-3).

<b>Bild</b>	<b>Umschreibung des spezifischen Themas</b>
„Eis“	1 Anonymität, sich hinter der Sonnenbrille verstecken wollen
	2 Schamgefühl, unter den Blicken anderer leiden, sich für die Figur schämen
	3 Denken an Essen, inneres Abwägen, wie viele Kalorien das Eis hat und ob es gegessen werden soll oder nicht
	4 Frustessen, Essen als Kompensation/ Ablenkung, obwohl es gar keinen Spaß macht
	5 Einsamkeit, allein auf dem Fest, keine Freunde
„Famili- lie“	1 Gedanken ans Essen
	2 Nationalität Amerikaner
	3 Abnehmen wollen, Diät
	4 Zusammengehörigkeitsgefühl, fühlen sich nur untereinander wohl und stark
	5 Trägheit, Bequemlichkeit, kürzeste Weg zum Essen/ Ausruhen nehmen
„Strand“	1 Diskriminierung der Übergewichtigen, abfällige Bezeichnungen wie "Wal", abwertende Gedanken/ Äußerungen (der Schlanken) zu den Übergewichtigen
	2 Kategorisierung der sozialen Welt in die Personengruppen "Dicke" und "Dünne"
	3 Scham gegenüber Anderen aufgrund der eigenen Figur, eigenen Körper am liebsten verstecken, sich in Gegenwart anderer (Schlanker) unwohl fühlen
	4 Unzufriedenheit mit der eigenen Figur ohne Bezugnahme auf andere Personen, dick sein = unglücklich sein
„Schuh- wechsel“	1 Müdigkeit, Trägheit, Faulheit, Bequemheit
	2 gesundheitliche Probleme (Kondition, Schmerzen)
	3 Schwitzen
	4 Unsportlichkeit
	5 Abnehmen, Diät
	6 Ausgrenzung dicker Menschen, Hänseleien
	7 Misserfolg

**Tabelle 5-2,** Bildspezifische Inhaltsvariablen der Bilder für das Einstellungsobjekt Übergewicht, *Fortsetzung nächste Seite*



<b>Bild</b>	<b>Umschreibung des spezifischen Themas</b>
„Ferien-lager“	1 Einsamkeit, Ausgegrenzt sein
	2 Unselbstständigkeit, Abhängigkeit, Heimweh, Anruf bei de Eltern, Muttersöhnchen
	3 Probleme in der Gesellschaft, Heimlichkeit, Fassungslosigkeit, Diskriminierung
	4 Bequem, muss sich hinsetzen beim Telefonieren, nicht leistungsfähig
	5 Gesundheitliche Probleme, unsportlich
	6 Unangenehmer Körpergeruch, Schwitzen
„Fitness-studio“	1 Willensstärke bezüglich Sport, Essen
	2 Neid auf die Schlanken, heimlicher Vergleich der Übergewichtigen mit ihren schlanken Freundinnen
	3 geringer Selbstwert der Übergewichtigen, schlank sein = glücklich sein/ dick sein = unglücklich sein
	4 Friedfertigkeit/ Vermittlerfunktion der Übergewichtigen
	5 Ausgrenzung der Übergewichtigen durch andere

**Fortsetzung Tabelle 5-2**

<b>Bild</b>	<b>Umschreibung des spezifischen Themas</b>
Frühs-tück	1 Beziehungsprobleme
	2 Sex, One-Night-Stand
	3 Probleme in der Gesellschaft, Heimlichkeit, Fassungslosigkeit, Diskriminierung
	4 Vertrautheit, Nähe, Bindung, Liebe
	5 anderweitige Probleme, z.B. finanzielle
Natur	1 Verständnis, Nähe, Zusammenhalt, Zärtlichkeit, Trostspenden
	2 Probleme in der Gesellschaft, Diskriminierung, ausgestoßen sein, Randgruppe sein, verbale/ physische Gewalt erleben
	3 Eigene Probleme: Trauer, Selbstzweifel, niedriges Selbstbewusstsein, Gesundheit
	4 Familiäre Bande, Vater-Sohn-Beziehung
Couch	1 Liebe, nicht zwangsläufig nur zwischen dargestellten Personen, sondern auch allgemein
	2 Heterosexuelles Pärchen erkannt, statt homosexuelles
	3 Beziehungsprobleme, Streit, Trennungsthematik
	4 Konsum von Drogen/ Alkohol, Betrunkensein, (Party-) Exzesse
	5 Gemütliches stressfreies Zusammensitzen
Aus-stellung	1 Konflikte in der Familie
	2 Harmonie in der Familie
	3 Inhalt des Ausstellungsbildes
	4 Toleranz
	5 (unterschwellig) auf Erotik angesprochen
	6 Auf Beziehung der Eltern abgezielt
Urlaub	1 Erinnerungsfoto an einen Urlaub
	2 Zeit füreinander haben, diese genießen
	3 Glücklich, zufrieden, entspannt sein
	4 Große Vertrautheit, kennen sich schon lange
Werks-halle	1 Kontext von Arbeit
	2 Dreckig, Kräfte zehrend, schwitzen
	3 Kollegenverhältnis, wenig vertraut
	4 Kontext von Sport
	5 Auf die eigenen Gedanken fixiert

**Tabelle 5-3** Bildspezifische Inhaltsvariablen der Bilder für das Einstellungsobjekt Homosexualität

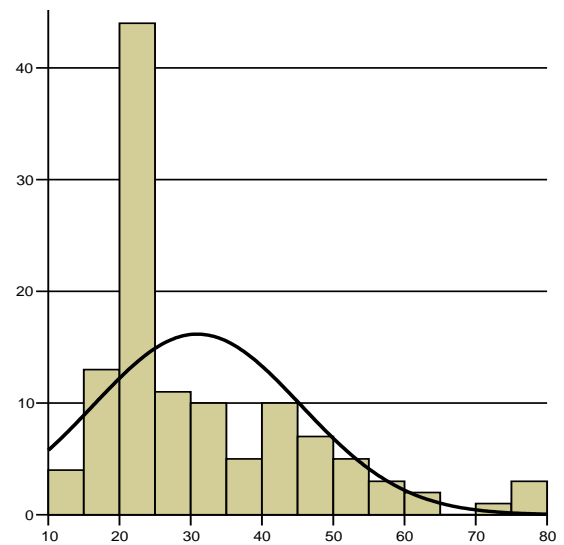
Zur Überprüfung der Eignung von Bildmaterial und Kodierschemata für die Einstellungs- und einstellungsbezogene Motivmessung wird eine empirische Untersuchung durchgeführt.<sup>4</sup> Ziel dieser Untersuchung ist zudem die Erstellung einer Rangliste der Bildmotive bzw. die Auswahl der zwei jeweils am besten geeigneten Bilder pro Einstellungsobjekt.

## 5.3 Methode

### 5.3.1 Stichprobe

Die aus der Normalbevölkerung rekrutierte Stichprobe umfasst insgesamt  $N=119$  Probanden. Die Geschlechterverteilung ist mit 55,5% Frauen ( $N=66$ ) relativ ausgewogen. Das Alter ist dagegen deutlich schief verteilt mit einem Mittelwert von 30,9 Jahren ( $SD=14,6$ ) und einem Median von 24,0 Jahren. Das liegt an der relativ großen Gruppe 20- bis 25-Jähriger bei einer Spannweite von 12 bis 77 Jahren (vgl. Abb. 5-1). Mit 46,6 % ist die Allgemeine Hochschulreife der am häufigsten erreichte höchste Bildungsabschluss. Jeweils eine Person gibt an, homo- bzw. bisexuell orientiert zu sein, während sechs weitere Probanden keine Angabe zu ihrer sexuellen Orientierung machen.

Da jeweils sechs Bildmotive pro Einstellungsobjekt getestet werden sollen, jedoch das Schreiben von insgesamt 12 Geschichten die Chancen auf eine hinreichend große Probandenzahl deutlich mindern würde, bearbeitet jeder Untersuchungsteilnehmer ein Testheft mit jeweils vier Geschichten. Die Gesamtstichprobe gliedert sich gemäß den drei Testheften in drei Teilstichproben, wobei die Zuordnung von Probanden zu Testheften zufällig stattfindet. Dennoch sind bei Stichprobenumfängen dieser Größenordnung ( $N=36$  bis  $43$ ) keine homogenen Merkmalscharakteristika zu erwarten. So weisen die drei Teilstichproben Unterschiede hinsichtlich Alters- und Geschlechtsverteilung auf. Die Altersspannweite variiert bspw. zwischen 63 Jahren in Stichprobe 2 (Testheft 2) und 39 Jahren in Stichprobe 3 (Testheft 3). Daraus resultiert in Stichprobe 2 ein vergleichsweise hoher Altersmittelwert (33,7 Jahre) und eine bedeutend höhere Standardabweichung (19,5 Jahre), als in den beiden anderen Stichproben. Die Charakteristika der Teilstichproben veranschaulicht Tabelle 5-4.



**Abbildung 5-3,** Altersverteilung in der Gesamtstichprobe

<sup>4</sup> Die empirische Untersuchung fand im Rahmen zweier Empiriepraktika (Wintersemester 2003/2004 und Sommersemester 2004) mit Psychologiestudenten des Grundstudiums an der Friedrich-Schiller-Universität Jena statt.

		<b>Testheft 1</b>	<b>Testheft 2</b>	<b>Testheft 3</b>
<b>Stichprobenumfang</b>		N=38	N=36	N=43
<b>Alter in-Jahren</b>	M (SD)	31.1 (13.2)	33.7 (19.5)	28.3 (10.4)
	Median	24	23	24
	Range (Spannweite)	12-62 (50)	14-77 (63)	12-51 (39)
<b>Geschlecht</b>	weiblich	20 (52,6%)	19 (52,8%)	27 (62,8%)
	männlich	18 (47,4%)	17 (47,2%)	16 (37,2%)
<b>höchster Bildungsabschluss</b>	ohne	1 (2,7%)	1 (2,8%)	2 (4,7%)
	Hauptschule	2 (5,4%)	1 (2,8%)	1 (2,3%)
	Realschule	4 (10,8%)	2 (5,6%)	5 (11,6%)
	Abitur	20 (54,1%)	18 (50,0%)	24 (55,8%)
	Berufsausbildung	4 (10,8%)	8 (22,2%)	6 (14,0%)
	Hochschule	6 (16,2%)	6 (16,7%)	5 (11,6%)
<b>Herkunft</b>	Neue Bundesländer	20 (52,6%)	10 (27,8%)	26 (60,5%)
	Alte Bundesländer	16 (42,1%)	24 (66,6%)	14 (32,5%)
	nicht in Deutschland	2 (5,3%)	2 (5,6%)	3 (7,0%)

**Tabelle 5-4**, Stichprobencharakteristika für die drei Testhefte

Als qualitatives Datenmaterial liegen insgesamt 455 Geschichten vor (Tab. 5-5). Bis auf wenige Ausnahmen haben alle Probanden vier Geschichten geschrieben. Zwei Teilnehmer bearbeiteten lediglich den Fragebogen, ohne eine Geschichte zu verfassen.

		<b>Testheft 1</b>		<b>Testheft 2</b>		<b>Testheft 3</b>			
1	Frühstück	H	35	Strand	Ü	36	Ferienlager	Ü	41
2	Eis	Ü	38	Couch	H	34	Urlaub	H	43
3	Natur	H	38	Schuhwechsel	Ü	35	Fitnessstudio	Ü	43
4	Familie	Ü	37	Ausstellung	H	34	Werkhalle	H	41
<i>gesamt</i>		<i>N=148</i>		<i>N=139</i>		<i>N=168</i>			

**Tabelle 5-5**, Anzahl vorliegender Geschichten je Bildmotiv, H= Einstellungsobjekt Homosexualität, Ü=~ Übergewicht

### 5.3.2 Material

Das Untersuchungsmaterial besteht aus drei unterschiedlichen Testheften. Jedes Testheft beginnt mit einer Coverstory, in die die Instruktionen zum Schreiben der Geschichten eingebettet sind. Daran schließt sich die Geschichtenmessung an, wobei die Geschichten beider Einstellungsobjekte im Wechsel geschrieben werden. Für jede Geschichte wird ein separates Blatt vorgesehen, auf deren linken Hälften die jeweiligen Bildmotive präsentiert sind. Unter den Bildern befinden sich zur Erinnerung die bereits in der Instruktion eingeführten anleitenden Fragen für das Schreiben der Geschichten, die aus dem klassischen TAT übernommen werden:

- *Was stellt die Illustration dar?*
- *Was denken und fühlen die dargestellten Personen?*
- *Wie kam es zu der abgebildeten Szene?*
- *Wie könnte sich die Situation weiter entwickeln?*

Die rechten Hälften der Blätter sind jeweils liniert und bezeichnen den Platz zum Niederschreiben der Geschichten. Jedes Testheft schließt mit einem Fragebogen ab, der zur expliziten Messung von Einstellungen gegenüber übergewichtigen und homosexuellen Menschen, zur Erfassung vorurteilsbezogener Motivation sowie ausgewählter Korrelate sozialer Einstellungen dient. Der Selbstbericht umfasst folgende Skalen:

<b>Messung</b>	<b>Abkürzung</b>	<b>N Items</b>
<i>Einstellungen zu übergewichtigen Menschen</i> (Degner, in prep)	ÜBKO	34
<i>Kognitive Einstellungen zur Homosexualität</i> (Seise, Banse & Neyer, 2002; 6 zusätzliche Items von Kämpfe, 2002)	HOKO	15
<i>Affektive Einstellungen zur Homosexualität</i> (Seise, Banse & Neyer, 2002)	HOAF	9
<i>Modern Racism</i> (McConahay, 1986; dt. Übersetzung der Autorin)	MRS	8
<i>Personal Need for Structure</i> (Neuberg & Newsom, 1993; deutschsprachige Fassung von Hänze, 2002)	PNS	15
<i>Motivation zu vorurteilsfreiem Verhalten</i> (Banse & Gawronski, 2003)	MVV	16
<i>Internal- and External Motivation to Respond Without Prejudice</i> (Plant & Devine, 1998, übersetzte und psychometrisch getestete Fassung)	IMS/EMS	8
<i>Anzahl übergewichtiger/ homosexueller Freunde</i>	FreundH/Ü	2
<i>Sexuelle Orientierung</i>	Sex-O	1
<i>Körpergröße/ Gewicht</i>	BMI	2
<i>Geschlecht, Alter, höchster Bildungsabschluss, Bundesland</i>		4

**Tabelle 5-6,** Skalen und Items der expliziten Fragebogenmessung

Alle Items sind auf einem fünfstufigen Antwortformat von 1 „lehne vollkommen ab“ bis 5 „stimme vollkommen zu“ zu beantworten, sodass höhere Werte eine stärkere Ausprägung im Sinne der Konstruktformulierung anzeigen. Das Geschlecht ist wie folgt kodiert: 1= weiblich, 2=männlich. Für die sexuelle Orientierung stehen drei Antwortmöglichkeiten zur Verfügung: 1= heterosexuell, 2= bisexuell, 3= homosexuell. Die beiden Items zur Erfassung der Anzahl homo-/bisexueller und übergewichtige Freunde und Bekannte sind vierfach gestuft, von „keine“ (1), „etwa 1 bis 4“ (2), „etwa 5 bis 9“ (3) und „mehr als 10“ (4).

### 5.3.3 Durchführung

Untersuchungsteilnehmern aus der Allgemeinbevölkerung wurde zufällig eines von drei Testheften zugewiesen und zur Bearbeitung vorgelegt. Die Probanden wurden aus dem Bekanntenkreis der Studenten rekrutiert. Die Bearbeitung der Testhefte erfolgte größtenteils in Eigenverantwortung. In der Coverstory werden die Untersuchungsteilnehmer gebeten, die Auswahl von Illustrationen für einen Kurzgeschichtenband zu unterstützen.

Nach Abschluss der empirischen Untersuchung wurden zwei unabhängige Expertenratings zur Augescheinvalidität der Bildmotive eingeholt, die im Anschluss an die Auswertung der empirischen Daten zur Geschichtenmessung dokumentiert sind.

### 5.3.4 Kodierung

Das umfangreiche qualitative Datenmaterial von 455 Geschichten wurde zur Kodierung so aufgeteilt, dass jede Versuchsperson komplett, d.h. alle vier Geschichten, von jeweils vier unabhängigen Ratern kodiert werden konnten. Insgesamt waren 18 geschulte Rater an der Überführung der Geschichten in quantitatives Datenmaterial gemäß den Kodierschemata beteiligt, woraus resultiert, dass nicht jeder Rater alle Bilder kodiert hat. Aufgrund organisatorischer Hindernisse konnte lediglich sichergestellt werden, dass pro Geschichte vier Ratings vorliegen und die vier Geschichten einer Versuchsperson von den jeweils gleichen vier Ratern kodiert wurden. Das bedeutet andererseits, dass alle Daten für eine spezifische Bild-Kodierschemata-Einheit nicht von den gleichen vier Ratern stammen. Dies stellt eine zusätzliche Quelle von Fehlervarianz dar, wobei die Interrater-Reliabilitäten im TAT ohnehin bereits als zu gering kritisiert werden (vgl. Hofer & Chasiotis, 2004). Die Berechnung der Interrater-Reliabilitäten erfolgt mittels Ein-Weg-Modell (Shrout & Fleiss, 1979).

Die vier Ratings je Kodiervariable wurden zur Auswertung aggregiert. Für die Analyse der Inhaltsvariablen, die mit 0 und 1 indikatorkodiert sind, wurden die aggregierten Variablen zusätzlich gerundet.

Da unter Umständen bei vier Ratern keine eindeutige Entscheidung für eine von zwei Kategorien abgeleitet werden kann, wurden, wenn notwendig, zusätzliche Entscheidungsratings eingeholt. Dies war in 191 einzelnen Fällen vonnöten, was bei insgesamt 455 Geschichten, 61 Inhaltsvariablen und 18 Ratern durchaus vertretbar ist (vgl. Tab. 5-7).

	<b>Bild 1</b>	<b>Bild 2</b>	<b>Bild 3</b>	<b>Bild 4</b>
<b>Testheft 1</b>	12% (Frühstück)	17% (Eis)	7% (Natur)	10% (Familie)
<b>Testheft 2</b>	10% (Strand)	5% (Couch)	6% (Schuhwechsel)	6% (Ausstellung)
<b>Testheft 3</b>	5% (Ferienlager)	10% (Urlaub)	2% (Fitnessstudio)	8% (Werkshalle)

**Tabelle 5-7,** Anteil notwendiger Zusatzratings zur Entscheidung über die Inhaltsvariablen je Bild

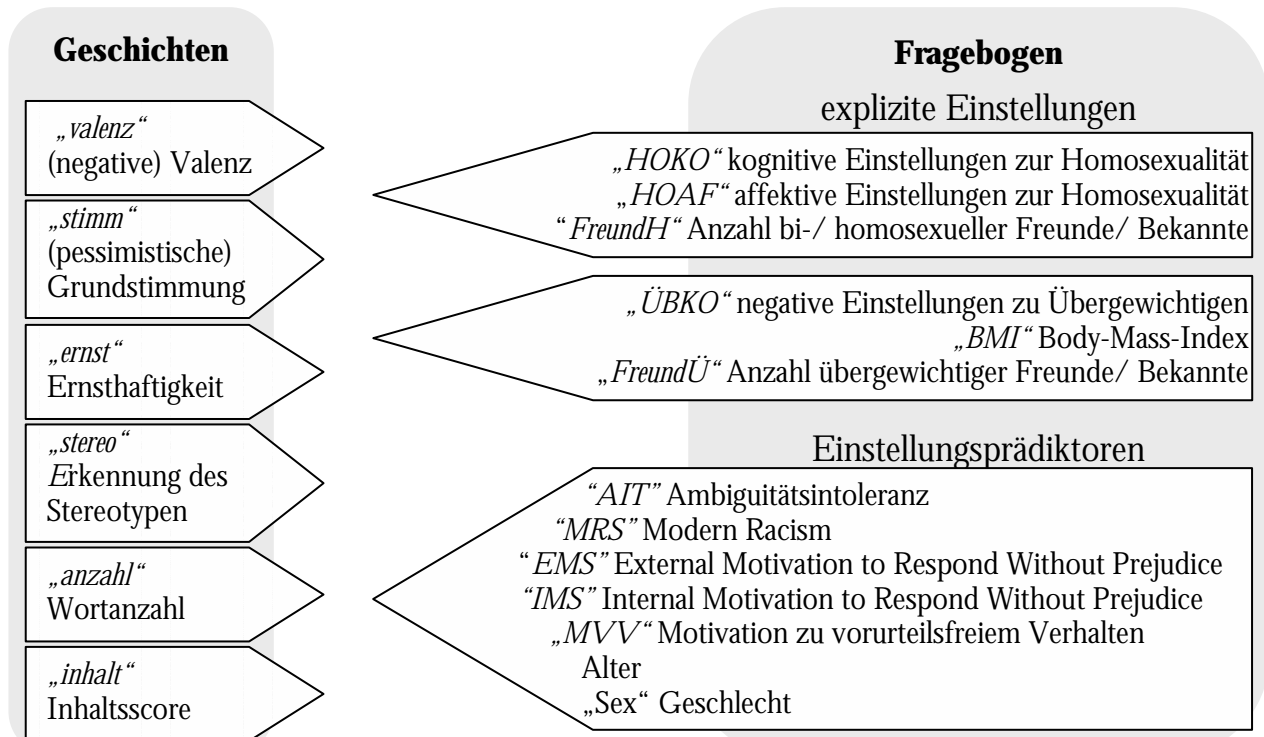
## 5.4 Auswertung

### 5.4.1 Allgemeines Vorgehen

Zur Beurteilung der offenen Assoziativmessung werden verschiedene Kriterien betrachtet:

- 1.) Interne Evaluationskriterien:
  - a) Deskriptive Statistiken der kodierten Variablen
  - b) Zusammenhänge zwischen den Variablen
- 2.) Externe Evaluationskriterien:
  - a) Zusammenhänge mit expliziten Einstellungen und einstellungsbezogener Motivation
  - b) Zusammenhänge mit Einstellungsprädiktoren und demografischen Variablen
  - c) Expertenrating des Bildmaterials

Ausschlaggebend für die Messung von Einstellungen und einstellungsbezogener Motivation über die selbst verfassten Geschichten sind insbesondere die für alle Geschichten kodierten *formellen* Variablen (Valenz, Stimmung, Stereotyperkennung, Ernsthaftigkeit und Wortanzahl). Die bildspezifischen Inhaltsvariablen werden als Kontrollvariablen herangezogen, um zu überprüfen, einerseits *inwieweit* die Darstellung eine Aktivierung des entsprechenden Stereotypen leistet und andererseits, ob der angestrebte vielfältige Interpretationsspielraum gegeben ist. Einen Überblick der für die Evaluation der offenen Assoziativmessung betrachteten Variablen gibt Abbildung 5-4.



**Abbildung 5-4:** Evaluation der offenen Assoziativmessung. Variablen in Konstruktichtung kodiert

### Exkurs 5-1 (Fortsetzung nächste Seite)

#### – Alphafehler-Kumulation –

Für die verschiedenen Geschichten-Variablen werden sowohl deren interne Beziehungen sowie deren Zusammenhang mit expliziten Einstellungs- und Motivationsvariablen überprüft. Aufgrund der sich daraus ergebenden Vielzahl notwendiger Signifikanztests besteht ein erhöhtes Risiko für Alpha-Fehler-Kumulierung, das sich wie folgt berechnet:

$$P(\text{korrekte } H_0 \text{ Entscheidung}) = (1-0.05)^{N \text{ Tests}}$$

Die Analyse der internen Zusammenhänge sowie der Beziehungen zwischen den Geschichtenvariablen und den Fragebogenkriterien bedeutet die Durchführung von 75 Signifikanztests pro Bild. Daraus ergibt sich eine Wahrscheinlichkeit, die Nullhypothese korrekterweise anzunehmen, von 2.1%, was eine annähernde Sicherheit darstellt, mindestens einen Alpha-Fehler einzugehen.

Dem Problem der Alphafehler-Kumulierung muss vorausschauend begegnet werden, da Korrekturmöglichkeiten für das Signifikanzniveau, wie die Bonferroni-Korrektur (adjustiertes  $p = p / N$  Tests), in einem  $p$ -Wert resultieren würden, der bei den vorliegenden Stichprobengrößen nicht eine Korrelation als signifikant ausweisen würde. Zunächst kann die Richtung sinnvoller Zusammenhänge ad hoc bestimmt werden, wobei auf einseitiges Testen verzichtet wird, um das Alphaniveau nicht zu senken. Es ist außerdem zu beachten, dass das Ziel der Zusammenhangsanalysen im vorliegenden Fall eine Materialauswahl ist, also ein Vergleich der sechs Bilder je Stereotyp, um eine Entscheidung für die zwei besten Bilder je Stereotyp treffen zu können. Entscheidend für die Auswahl sind dabei vorrangig die Beziehungen zwischen der *Valenz* der Geschichten, da sich hierin affektive, und somit vorurteilsrelevante Einstellungsaspekte niederschlagen können, und der expliziten Messung. Im Fall einer Korrespondenz zwischen offener Assoziativmessung und expliziter Einstellungsmessung ist mit positiven Zusammenhängen zwischen der Valenz der Geschichten und den expliziten Einstellungen zu rechnen, da hohe Werte jeweils negative Valenz und negative Einstellungen reflektieren. Da den Analysen ohnehin recht kleine Stichproben zugrunde liegen, erscheint der Alpha-Fehler vergleichsweise unkritisch.

#### 5.4.2 Hypothesen

Bei der Zusammenhangsanalyse der Geschichtenvariablen mit den Fragebogenmaßen, stellen die messspezifischen Unterschiede beider Verfahren eine Quelle für Interpretationsfehler dar. So sind fehlende Zusammenhänge zwischen den unterschiedlichen Einstellungsmaßen kein eindeutiges Zeichen mangelnder Validität z.B. der offenen Assoziativmessung (da der Fragebogen als Referenz für die Geschichtenmessung dient). Mögliche Ursachen für eine eventuelle Unabhängigkeit zwischen expliziten Einstellungen und den Geschichtenvariablen können zwar einerseits mangelnde Validität der Bildmotive bzw. der Inhaltskodierung reflektieren. Andererseits ist jedoch auch mit unterschiedlicher Methodenvarianz und Messfehlerbehaftung beider Messungen zu rechnen, was dazu führen kann, dass wahre Zusammenhänge verdeckt bleiben. Da an dieser Stelle keine hinreichende statistische Kontrolle von Methoden- und Messfehlervarianz erfolgen kann, ist mit geringen bis moderaten Korrelationsbeträgen zwischen den Geschichtenvariablen und den expliziten Einstellungsmaßen zu rechnen.



Zur Richtung der Zusammenhänge zwischen den formellen Variablen der offenen Assoziativmessung und den selbstberichteten Einstellungen können jedoch bereits im Vorfeld einige Hypothesen aufgestellt werden:

- Für *Valenz* und *Grundstimmung* sind positive Beziehungen zu expliziten Einstellungen zu erwarten
- Von *Ernsthaftigkeit* ist ein negativer Zusammenhang zu erwarten, da Zynismus, Spott, Sarkasmus als Indikatoren einer Abwehr der Thematik interpretiert werden können
- Die *Stereotypenerkennung* dient als Kontrollvariable insbesondere für das Einstellungsobjekt Homosexualität. Sie kann ebenfalls als Motivation betrachtet werden, die Stereotypen nicht anzusprechen und sollte insofern negative Beziehungen zu expliziten Einstellungen unterhalten
- Für die *Wortanzahl* können vorab keine eindeutigen Zusammenhangsannahmen abgeleitet werden. Im Sinne einer motivierten Vermeidung der Thematik wären eher negative Korrelationen zwischen der Geschichtenlänge und Fragebogenmaßen zu erwarten. Es ist jedoch ebenso denkbar, dass die Wortanzahl in keinerlei Beziehung zu expliziten Einstellungen steht

## 5.5 Ergebnisse

### 5.5.1 Offene Assoziativmessung

#### 5.5.1.1 Basisanalyse

Die Interrater-Reliabilitäten (Ein-Weg, zufällig) sind im Mittel gut bis exzellent. Für die formelle Variable *Ernsthaftigkeit* wird ein sehr heterogenes Reliabilitätsmuster über die 12 Einzelmessungen hinweg beobachtet. In vier von zwölf Fällen kann aufgrund ungenügender Varianz gar keine Interraterreliabilität berechnet werden (Bilder „Frühstück“, „Natur“, „Familie“, „Ausstellung“) und in 2 Fällen wird ein Intraklassen-Korrelationskoeffizient von nahe Null ermittelt („Couch“: ICC=.04; „Schuhwechsel“: ICC=.00). Die mittlere Interraterreliabilität über die bildspezifischen Inhaltsvariablen beträgt .79 und über die allgemeinen formellen Variablen .73.

Einer deskriptiven Analyse zufolge weisen die Bilder deutliche Unterschiede in nahezu allen kodierten Merkmalen auf (siehe Tab. 5-8). Auf einem globalen Level zeigen sich auch zwischen beiden Einstellungsobjekten Differenzen auf. Werden die Variablen der jeweils sechs Einzelmessungen für Homosexualität und Übergewicht zusammengefasst, fällt die deutlich positivere Valenz in den Geschichten, die zu den Homosexualitätsmotiven (H-Geschichten) verfasst wurden, auf (z.B. Urlaub: 2.02; Werkshalle: 2.51). Das mag daran liegen, dass gleichzeitig – und wie zu erwarten war – die Erkennung des thematisierten Stereotypen Homosexualität in den H-Geschichten nicht zum Ausdruck kommt. Am deutlichsten fällt die Stereotypenerkennung für die Bilder „Urlaub“, „Natur“ und „Frühstück“ aus, was für „Urlaub“ und „Natur“ auch die Inhaltsscores reflektieren. Nur ein sehr geringer Anteil an

Probanden erkennt im Bild „Couch“ den Stereotypen Homosexualität – aufgrund der Darstellung sahen viele Probanden statt ein homosexuelles- vielmehr ein heterosexuelles Pärchen. Bereits an dieser Stelle muss das Bild daher als fraglich für die vorgesehene Vorurteilmessung bezeichnet werden.

Die Geschichten zu den Übergewichtsmotiven (Ü-Geschichten) lassen oft sogar eine *explizite* Erkennung des Stereotypen seitens der Probanden zu, in jedem Fall jedoch eine implizit zum Ausdruck kommende Stereotypenaktivierung. Die Valenz der Ü-Geschichten liegt im Durchschnitt über dem neutralen Skalenmittel von 3, einzige Ausnahme ist das Bild „Fitnessstudio“ mit 2.79.

Ebenso wie die Valenz ist die Grundstimmung, also der Ausgang der Geschichten, tendenziell positiver in den H-Geschichten. Diese im Mittel optimistischere Stimmung rührt hauptsächlich von 2, im Vergleich zu den anderen Motiven, deutlich positiver aufgefassten Bildern her, und zwar „Urlaub“ (1.47) und „Natur“ (1.71). Bei den Ü-Bildern fällt „Ferienlager“ durch die mit Abstand pessimistischste Grundstimmung auf.

	Bild	Valenz	Stimmung	Ernsthaftigkeit	Stereotyp-erkennung	Wortanzahl	Inhalt
		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Übergewicht	Eis	3.34 (.82)	2.18 (.61)	.87 (.34)	1.58 (.68)	74.45 (39.36)	.26 (.22)
	Fam	3.24 (.80)	2.08 (.83)	.95 (.23)	.89 (.81)	64.82 (40.08)	.24 (.19)
	Stra	3.44 (.65)	2.28 (.66)	.97 (.17)	1.75 (.60)	78.00 (38.29)	.55 (.29)
	Sch	3.20 (.68)	2.20 (.58)	1.00 (.00)	1.63 (.65)	69.94 (38.28)	.27 (.22)
	Fer	3.17 (.80)	2.60 (.54)	.93 (.26)	.95 (.91)	75.25 (37.87)	.25 (.21)
	Fit	2.79 (.77)	2.16 (.79)	.91 (.29)	1.23 (.90)	60.58 (28.35)	.12 (.16)
<i>gesamt</i>		<i>3.14 (.58)</i>	<i>2.22 (.41)</i>	<i>.93 (.18)</i>	<i>1.25 (.60)</i>	<i>69.96 (33.56)</i>	<i>.28 (.18)</i>
Homosexuell	Früh	3.17 (.89)	2.40 (.60)	.89 (.32)	.63 (.81)	72.94 (42.27)	.25 (.17)
	Nat	2.58 (.92)	1.71 (.77)	.95 (.23)	.71 (.80)	65.18 (33.85)	.39 (.21)
	Cou	2.91 (1.06)	2.06 (.65)	.97 (.17)	.15 (.44)	64.06 (35.38)	.30 (.16)
	Aus	2.79 (.73)	2.00 (.55)	.97 (.17)	.50 (.83)	66.23 (37.38)	.20 (.18)
	Url	2.02 (.83)	1.47 (.55)	.91 (.29)	1.05 (.98)	58.85 (27.18)	.45 (.28)
	Werk	2.51 (.64)	2.15 (.62)	.93 (.26)	.52 (.86)	56.94 (28.72)	.29 (.20)
<i>gesamt</i>		<i>2.63 (.66)</i>	<i>1.95 (.44)</i>	<i>.93 (.20)</i>	<i>.61 (.59)</i>	<i>63.95 (31.56)</i>	<i>.31 (.17)</i>

**Tabelle 5-8**, Deskriptive Statistiken der Geschichtenvariablen, Inhalt: Mittel der bildspezifischen Inhaltsvariablen (siehe Tab. 5-2 und 5-3); Bildkürzel: Fam=“Familie“, Stra=“Strand“, Sch=“Schuhwechsel“, Fer=“Ferienlager“, Fit=“Fitnessstudio“; Früh=“Frühstück“, Nat=“Natur“, Cou=“Couch“, Aus=“Ausstellung“, Url=“Urlaub“, Werk=“Werkhalle“ (siehe Abb. 5-1 und 5-2)

### 5.5.1.2 Zusammenhänge der kodierten Variablen

Die internen Zusammenhänge der Kodiervariablen stellen insbesondere eine Kontrollmöglichkeit für die Kodierung selbst dar. Einerseits sollte ein Kodierschemata möglichst wenig inhaltliche Überlappungen aufweisen, andererseits untermauern selektive und erwartungsgemäß gerichtete Zusammenhänge die Validität der Kodierung. Die Korrelationen der Geschichtenvariablen werden im Folgenden für beide Einstellungsobjekte gesondert betrachtet (vgl. Tab. 5-9, 5-10). Dabei ist zu beachten, dass es sich an dieser Stelle um Zusammenhänge handelt, die jeweils nur zwischen den beiden Geschichtenmessungen je Einstellungsobjekt innerhalb der testheftspezifischen Stichproben gelten. Die Zusammenhänge zwischen den Geschichtenvariablen von z.B. Testheft 1 und 2 können nicht geprüft werden, da den Messungen verschiedene Stichprobe zugrunde liegen.

In Abschnitt 5.5.3.3 erfolgt eine generalisierte Zusammenhangsanalyse für die Geschichtenvariablen und die expliziten Einstellungsmaße innerhalb der Gesamtstichprobe.

Für die Geschichtenvariablen zum Einstellungsobjekt Übergewicht bestehen deutliche und durchgängige Überlappungen zwischen den formellen Variablen *Valenz* und *Stimmung* (von  $r=.54$  für „Eis“ bis  $r=.87$  für „Urlaub“). Diese wertigkeitsanzeigenden Variablen sind offenbar nicht hinreichend voneinander zu unterscheiden. Bei den Ü-Geschichten zu „Ferienlager“ und „Fitness“ sowie den H-Geschichten zu „Natur“, „Urlaub“ und „Werkshalle“ stehen Valenz und Stimmung zudem in positivem Zusammenhang mit der *Ernsthaftigkeit*. In diesen Fällen zeichnen sich Geschichten, die als ernst zu nehmend eingestuft werden, durch signifikant negativere Valenz und pessimistischere Stimmung aus. Ausschließlich für „Couch“ geht eine Erkennung des Stereotypen mit positiver- ( $r=-.45$ ,  $p<.01$ ), für „Eis“ und „Fitnessstudio“ mit negativer Valenz einher ( $r=.46$  und  $r=.47$ ,  $p<.01$ ). Der gemittelte Inhaltsscore steht bis auf eine Ausnahme (Ferienlager“) in allen Fällen in positiver Beziehung zur Wortanzahl der Geschichten. Für „Familie“ ( $r=.61$ ), „Schuhwechsel“ ( $r=.47$ ), „Ausstellung“ ( $r=.49$ ), „Urlaub“ ( $r=.48$ ) und „Werkshalle“ ( $r=.33$ ) erreichen die Korrelationen statistische Bedeutsamkeit. Dieser Zusammenhang erscheint nahezu notwendig, zugleich wird damit jedoch auch die *inhaltliche* Relevanz der Kontrollvariablen „Wortanzahl“ veranschaulicht.

<b>Testheft 1</b>						<b>„Eis“</b>					<b>„Familie“</b>				
N=38	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt
Valenz	<b>.54</b>	-.11	<b>.46</b>	.26	.26	<b>.81</b>	-.17	.22	-.02	.02					
Stimm		-.03	.17	.05	<b>.46</b>		.13	-.07	-.20	-.27					
Ernst			.10	.03	-.15			-.28	-.19	-.26					
Stereo				.32	.26				.23	.18					
Anzahl					.32					<b>.61</b>					
<b>Testheft 2</b>						<b>„Strand“</b>					<b>„Schuhwechsel“</b>				
N=36	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt
Valenz	<b>.82</b>	-.24	-.08	.01	.26	<b>.77</b>	-.04	.25	.22	.32					
Stimm		-.22	.06	.09	.40		.08	.27	.34	<b>.47</b>					
Ernst			-.06	-.10	-.06			.20	.31	.21					
Stereo				.10	<b>.55</b>				-.06	.27					
Anzahl					.22					<b>.46</b>					
<b>Testheft 3</b>						<b>„Ferienlager“</b>					<b>„Fitnessstudio“</b>				
N=43	Stimm.	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt
Valenz	<b>.64</b>	.38	.30	-.15	.11	<b>.74</b>	<b>.62</b>	<b>.47</b>	-.23	.16					
Stimm		<b>.58</b>	.19	-.23	-.17		.35	.20	-.24	-.10					
Ernst			.13	.02	-.16			<b>.52</b>	.03	-.04					
Stereo				-.04	<b>.50</b>				.07	.30					
Anzahl					.01					.20					

**Tabelle 5-9**, Interne Zusammenhänge der Geschichtenmessung für **Einstellungsobjekt Übergewicht**:  
Stimm=Stimmung, Ernst=Ernsthaftigkeit, Stereo=Stereotyp erkannt

<b>Testheft 1</b>						<b>„Frühstück“</b>					<b>„Natur“</b>				
N=38	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt
Valenz	<b>.61</b>	.17	.21	-.35	-.16	<b>.78</b>	<b>.44</b>	-.05	-.34	.20					
Stimm		.12	.18	-.17	-.03		<b>.66</b>	-.04	-.33	<b>.45</b>					
Ernst			-.17	-.01	-.16			.05	-.27	.08					
Stereo				-.08	.23				-.20	-.12					
Anzahl					.17					.25					
<b>Testheft 2</b>						<b>„Couch“</b>					<b>„Ausstellung“</b>				
N=36	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt
Valenz	<b>.84</b>	-.20	<b>-.45</b>	-.21	-.31	<b>.65</b>	.01	-.14	-.16	-.31					
Stimm		-.18	<b>-.48</b>	-.02	-.33		-.00	-.16	.14	.27					
Ernst			-.13	.18	.19			.15	.41	.20					
Stereo				.05	.28				.07	.14					
Anzahl					.24					<b>.49</b>					
<b>Testheft 3</b>						<b>„Urlaub“</b>					<b>„Werkshalle“</b>				
N=43	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt	Stimm	Ernst	Stereo	Anzahl	Inhalt
Valenz	<b>.87</b>	<b>.57</b>	.07	-.20	-.36	<b>.78</b>	<b>.53</b>	.04	-.21	-.09					
Stimm		<b>.49</b>	-.01	-.09	-.28		<b>.47</b>	.14	-.27	-.10					
Ernst			-.07	-.17	-.12			-.01	-.11	.06					
Stereo				-.29	-.23				-.12	-.29					
Anzahl					<b>.48</b>					.33					

**Tabelle 5-10**, Interne Zusammenhänge der Geschichtenmessung für **Einstellungsobjekt Homosexualität**:  
Stimm=Stimmung, Ernst=Ernsthaftigkeit, Stereo=Stereotyp erkannt

## 5.5.2 Fragebogen

### 5.5.2.1 Basisanalyse

Die Reliabilitäten der Fragebogenskalen sind befriedigend bis sehr gut. Die Skala Internale Motivation zur Vorurteilsfreiheit (IMS) weist mit .65 die geringste interne Konsistenz auf, während für die Skala Kognitive Einstellungen zu Homosexualität die höchste Reliabilität erhalten wird (Tab. 5-11).

Die Skalenmittelwerte zeugen von ausgewogenen Einstellungen zu Übergewicht, leicht positiven kognitiven und ausgewogenen bis leicht negativen affektiven Einstellungen zu Homosexualität in der Gesamtstichprobe. Für die Motivation zu vorurteilsfreiem Verhalten (MVV), sowie die Internale Motivation zur Vorurteilsfreiheit (IMS) werden hohe Mittelwerte deutlich über dem theoretischen Skalenmittel von 3 ermittelt. Die Stichprobenmittelwerte für Modern Racism (MRS), Personal Need for Structure (PNS) und Externale Motivation zur Vorurteilsfreiheit (EMS) liegen relativ nah am theoretischen Skalenmittel. Die Versuchspersonen geben an, mehr übergewichtige, als homo- oder bisexuelle Personen im Freundes- oder Bekanntenkreis zu haben. Mit einem mittleren BMI von 23.06 liegt die Stichprobe im oberen Normalgewichtsbereich, jedoch besteht auch eine große Streubreite, was zum Teil durch den großen Altersrange zu begründen ist. Die Sexuelle Orientierung eignet sich aufgrund der geringen Varianz über die drei Teilstichproben hinweg nicht als Korrelat der Geschichtenmaße.

Die Analyse der einstellungsrelevanten Fragebogenmaße mittels univariater Varianzanalysen (ANOVA) erbringt keinerlei signifikante Unterschiede zwischen den Teilstichproben.

		<b>M</b>	<b>(SD)</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b>Cronbach's Alpha</b>
1.	<b>ÜBKO</b>	2.90	(.59)	1.68	4.53	2.85	.91
2.	<b>HOKO</b>	1.97	(.93)	1.00	5.00	3.40	.93
3.	<b>HOAF</b>	3.16	(.73)	1.44	4.40	3.56	.90
4.	<b>MRS</b>	2.75	(.85)	1.00	5.00	4.00	.72
5.	<b>PNS</b>	3.28	(.49)	2.07	4.86	2.79	.70
6.	<b>IMS</b>	4.05	(.80)	1.67	5.00	3.33	.65
7.	<b>EMS</b>	2.76	(.81)	1.00	5.00	4.00	.71
8.	<b>MVV</b>	3.68	(.51)	2.00	4.75	2.75	.80
9.	<b>FreundÜ</b>	2.15	(.60)	1	4	3	
10.	<b>FreundH</b>	1.66	(.81)	1	5	4	
11.	<b>BMI</b>	23.06	(3.19)	17.63	31.14	13.51	
12.	<b>Sex-O</b>	1.03	(.21)	1	3	2	

**Tabelle 5-11.** Deskriptive Statistiken und Reliabilität der Fragebogenmaße; ÜBKO= Einstellungen zu Übergewicht, HOKO/AF= kognitive/affektive Einstellungen zu Homosexualität, MRS= Modern Racism, PNS= Personal Need for Structure, MVV= Motivation zu vorurteilsfreiem Verhalten, IMS/ EMS= Internal/ External Motivation to Respond Without Prejudice, FreundÜ/H= Anzahl homosexueller/übergewichtiger Freunde und Bekannte, Sex-O= sexuelle Orientierung

### 5.5.2.2 Zusammenhänge der Fragebogenmaße

Moderate positive Korrelationen stehen zwischen den expliziten Einstellungsmaßen, wobei erwartungsgemäß die Zusammenhänge zwischen den beiden Maßen für das Einstellungsobjekt Homosexualität höher sind (HOKO, HOAF:  $r=.58$ ,  $p<.001$ ), als zwischen den Einstellungsobjekten (ÜBKO, HOKO:  $r=.33$ ,  $p<.01$ ; ÜBKO: HOAF:  $r=.21$ ,  $p<.05$ ; Tab. 5-12). Der BMI hat keine Bedeutung für Einstellungen oder Motivation und korreliert lediglich mit dem Alter ( $r=.50$ ,  $p<.001$ ). Die Sexuelle Orientierung dagegen korreliert negativ zu den expliziten Einstellungen zu Homosexualität, wobei der Zusammenhang zu affektiven Einstellungen (HOAF) signifikant ist ( $r=-.27$ ,  $p<.01$ ). Die beiden Probanden, die angaben, nicht heterosexuell orientiert zu sein, haben demzufolge nach eine positivere explizite Einstellung zu Homosexualität. Die Anzahl übergewichtiger Freunde unterhält keine Beziehungen zu anderen Variablen, dagegen haben Probanden mit mehr homo- oder bisexueller Personen im Freundes- und Bekanntenkreis signifikant positivere Einstellungen zu Homosexualität (HOKO:  $r=-.37$ , HOAF:  $r=-.41$ ,  $p<.001$ ). Außerdem zeigt die leicht positive Korrelation zur sexuellen Orientierung an ( $r=.20$ ,  $p<.05$ ), dass insbesondere die homosexuellen Probanden viele Freunde haben, die nicht heterosexuell orientiert sind. Die Kontaktvariable korreliert außerdem negativ mit PNS ( $r=-.19$ ,  $p<.05$ ), MRS ( $r=-.23$ ,  $p<.05$ ) und EMS ( $r=-.19$ ,  $p<.05$ ).

Erwartungskonform bestehen positive Zusammenhänge zwischen PNS, MRS und expliziten Einstellungen zu Übergewicht und Homosexualität. Internale und externale Motivation zur Vorurteilsfreiheit (IMS, EMS) sind, wie Devine (1989) und Devine et al. (2002) berichten, unkorreliert ( $r=-.04$ ,  $p>.05$ ). Davon abgesehen stehen die Motivationsvariablen untereinander in positivem Zusammenhang (EMS, MVV:  $r=.21$ ,  $p<.05$ ; IMS, MVV:  $r=.58$ ,  $p<.001$ ), unterscheiden sich allerdings in der Richtung ihrer Beziehung zu expliziten Einstellungen. Während externale Motivation zur Vorurteilsfreiheit in positivem Zusammenhang zu ÜBKO ( $r=.19$ ,  $p<.05$ ) und HOKO ( $r=.30$ ,  $p<.001$ ) steht, korrelieren internale Motivation zur Vorurteilsfreiheit (IMS) und die Motivation zur Vorurteilsfreiheit (MVV) durchweg negativ mit den expliziten Einstellungsmaßen.

Das Alter der Probanden steht in positiver Beziehung zu den expliziten Einstellungen, ältere Probanden haben signifikant negativer Einstellungen als jüngere (ÜBKO:  $r=.22$ , HOKO:  $r=.18$ ,  $p<.05$ , HOAF:  $r=.15$ ,  $p>.05$ ). Sie weisen zudem ein höheres Bedürfnis nach kognitiver Strukturiertheit ( $r=.40$ ,  $p<.001$ ) und stärkere Ausprägungen von Modern Racism ( $r=.44$ ,  $p<.001$ ) auf. Im Mittel sind ältere Probanden jedoch auch stärker external zur Vorurteilsfreiheit und zu vorurteilsfreiem Verhalten motiviert (EMS:  $r=.21$ ,  $p<.05$ ; MVV:  $r=.31$ ,  $p<.001$ ). Geschlechtsunterschiede werden für affektive Einstellungen zu Homosexualität ( $r=.32$ ,  $p<.001$ ) und die Motivation zu vorurteilsfreiem Verhalten ( $r=-.19$ ,  $p<.05$ ) beobachtet. Männer haben demzufolge nach negativere Einstellungen zu Homosexualität und sind weniger zur Vorurteilsfreiheit motiviert als Frauen.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. <b>ÜBKO</b>													
2. <b>HOKO</b>	<b>.33</b>												
3. <b>HOAF</b>	.21	<b>.58</b>											
4. <b>BMI</b>	.08	.13	.19										
5. <b>Sex-O</b>	-.01	.00	<b>-.27</b>	-.15									
6. <b>FreundÜ</b>	-.11	.04	.04	.05	.11								
7. <b>FreundH</b>	-.07	<b>-.37</b>	<b>-.41</b>	-.05	.21	-.02							
8. <b>PNS</b>	<b>.39</b>	<b>.32</b>	.23	.12	-.01	-.07	-.19						
9. <b>MRS</b>	<b>.50</b>	<b>.45</b>	<b>.26</b>	.13	-.16	-.03	-.23	<b>.44</b>					
10. <b>EMS</b>	.19	<b>.30</b>	.06	-.01	-.06	-.04	-.19	<b>.37</b>	<b>.25</b>				
11. <b>IMS</b>	<b>-.27</b>	<b>-.33</b>	<b>-.30</b>	.03	.03	-.02	.16	-.09	-.21	-.04			
12. <b>MVV</b>	-.21	-.17	-.13	.14	.12	.01	-.01	.13	-.12	.21	<b>.58</b>		
13. <b>Alter</b>	.22	.18	.15	<b>.50</b>	-.08	-.02	-.15	<b>.40</b>	<b>.44</b>	.21	.14	<b>.31</b>	
14. <b>Geschlecht</b>	.14	.16	<b>.32</b>	.09	.05	-.08	.00	-.08	.04	-.12	-.17	-.19	-.06

**Tabelle 5-12,** Korrelationen der Fragebogenmaße; Geschlecht: 1=weiblich, 2=männlich; Abkürzungen siehe Tab. 5-11

### 5.5.3 Gegenüberstellung der offenen Assoziativmessung mit expliziten Einstellungen

An dieser Stelle werden die Zusammenhänge zwischen den Variablen der Geschichtenmessung und den selbst berichteten Einstellungen, den einstellungsbezogenen Motivationsmaßen sowie den Einstellungsprädiktoren überprüft, um die Bedeutung der Geschichtenvariablen für explizite Maße untersuchen zu können. Zentral in der Gegenüberstellung der Bildmessung mit Fragebogenmaßen sind die Variable *Valenz* wie auch *Stimmung* und ihr Zusammenhang mit den Vorurteilsskalen des Fragebogens. Es ist jedoch an dieser Stelle nochmals zu berücksichtigen, dass den Geschichtenmaßen drei unterschiedliche Stichproben zugrunde liegen, was keine unbedingte Vergleichbarkeit der Zusammenhänge zu den Fragebogenmaßen erlaubt.

Von den 360 Korrelationen für das Einstellungsobjekt Übergewicht sind 39 signifikant, was eine Signifikanzrate von 10.8% bedeutet. Die Zusammenhangsanalyse für Homosexualität erbringt 43 statistisch bedeutsame Korrelationen (11.9%). Somit werden prinzipiell leicht überzufällig viele Zusammenhänge erhalten, die meisten signifikanten Korrelationen werden für „Strand“ (9), „Eis“ (8) und „Schuhwechsel“ (8) bzw. „Ausstellung“ (14) und „Couch“ (11) erhalten. Allgemein besteht ein relativ uneinheitliches Korrelationsmuster für die Bilder. Nur in wenigen Fällen werden ähnliche Zusammenhänge zwischen Geschichtenvariablen und Fragebogenmaßen über verschiedene Bilder hinweg beobachtet. Es muss jedoch berücksichtigt werden, dass der Gegenüberstellung der Geschichtenmessung mit den Fragebogenmaßen entsprechend der drei Testhefte drei unterschiedliche und zudem relativ kleine Stichproben zugrunde liegen. Diese Dreiteiltheit der Stichprobe erlaubt einen gültigen

Vergleich der Bildmotive nur innerhalb eines Testhefts. Über die Testhefte hinweg kann aufgrund der geringen Stichprobenumfänge keine Verallgemeinerung der Zusammenhangsmuster erfolgen.

### 5.5.3.1 *Einstellungsobjekt Übergewicht*

Für die Bildstimuli des Stereotypen Übergewicht können der Gegenüberstellung mit den Vorurteilsmaßen gemäß des Fragebogens wenige eindeutige Aussagen abgeleitet werden (Tab. 5-13). Weder für *Valenz* noch für *Stimmung* bestehen für eine der Ü-Geschichten signifikante Beziehungen zu „ÜBKO“ oder dem BMI. Für „Ferienlager“ und „Fitnessstudio“ sind die Zusammenhangsbeträge zumindest nicht nahe Null. Die Geschichtenlänge spielt dagegen eine bedeutsamere Rolle: Für „Strand“ und „Schuhwechsel“ stehen kürzere Geschichten mit negativeren Einstellungen zu Übergewicht in signifikanter Beziehung ( $r=-.33$ ,  $p<.05$  bzw.  $r=-.47$ ,  $p<.01$ ). Zudem werden durchgängig negative Korrelationen zwischen der Wortanzahl und dem BMI beobachtet, die für „Eis“ und „Strand“ signifikant sind ( $r=-.38$ ,  $p<.05$  bzw.  $r=-.53$ ,  $p<.01$ ). Personen mit höherem BMI schreiben demnach kürzere Ü-Geschichten. Die Anzahl übergewichtiger Freunde spielt nur für „Ferienlager“ und „Fitnessstudio“ eine Rolle. Für beide Ü-Geschichten bestehen negative Zusammenhänge zur Stereotyperkennung (bzw. -Ernennung,  $r=-.37$  bzw.  $r=-.32$ ,  $p<.05$ ). Für „Fitnessstudio“ wird zusätzlich eine negative Beziehung zwischen der Anzahl übergewichtiger Freunde und Valenz sowie Stimmung beobachtet ( $r=-.34$ ,  $p<.05$  bzw.  $r=-.42$ ,  $p<.01$ ). Diese Zusammenhänge könnten auf stärkere Aktivierung der Thematik Freundschaft in den beiden Darstellungen im Vergleich zu den anderen Bildmotiven zurückzuführen sein.

Personal Need for Structure (PNS) steht in zwei der Ü-Geschichten in positiver Beziehung zur Stereotyperkennung („Familie“:  $r=.35$ , „Schuhwechsel“:  $r=.41$ ,  $p<.05$ ) und in ebenfalls zwei Fällen in negativer Beziehung zur Wortanzahl („Strand“:  $r=-.54$ , „Schuhwechsel“:  $r=-.54$ ,  $p<.01$ ). Das bedeutet, Personen mit höherer PNS benennen in den genannten Geschichten weniger direkt den Stereotypen Übergewicht und schreiben weniger Wörter. Die drei Variablen, die eine Motivation zur Vorurteilsfreiheit bzw. -vermeidung anzeigen, haben eine nicht durchgängig gleichgerichtete Bedeutung für die Geschichtenvariablen. Die externale Motivation zur Vorurteilsfreiheit (EMS) korreliert negativ zur Geschichtenlänge bei „Strand“ ( $r=-.48$ ,  $p<.01$ ) und „Schuhwechsel“ ( $r=-.40$ ,  $p<.05$ ) sowie negativ zur Stereotyperkennung bei „Familie“ ( $r=-.35$ ). Die internale Motivation zur Vorurteilsfreiheit (IMS) steht bei „Eis“ in negativer-, bei „Ferienlager“ in positiver Beziehung zur Valenz ( $r=-.38$  bzw.  $r=.33$ ,  $p<.05$ ). Für die Motivation zu vorurteilsfreiem Verhalten wird, wie bereits für EMS, ein Zusammenhang zur Geschichtenlänge beobachtet, in drei Geschichten bestehen signifikant negative Zusammenhänge zur Wortanzahl („Strand“:  $r=-.33$ , „Schuhwechsel“:  $r=-.41$ , „Ferienlager“:  $r=-.33$ ,  $p<.05$ ).

Auch das Alter der Probanden korreliert negativ zur Geschichtenlänge, was in drei Fällen statistisch bedeutsam ist: „Strand“:  $r=-.65$ ,  $p<.001$ , „Schuhwechsel“:  $r=-.54$ ,  $p<.01$ , „Fitnessstudio“:  $r=-.37$ ,  $p<.05$ ).



Die Bedeutung des Geschlechts zeigt dich zum einen in kürzeren Geschichten für „Eis“ ( $r=-.33$ ,  $p<.05$ ), einer weniger direkten Stereotyperkennung für „Strand“ ( $r=-.38$ ,  $p<.05$ ) und einer direkteren Stereotyperkennung für „Fitnessstudio“ ( $r=.31$ ,  $p<.05$ ) sowie weniger inhaltlicher Vielfalt in „Schuhwechsel“ ( $r=-.43$ ,  $p<.05$ ) und „Ferienlager“ ( $r=-.38$ ,  $p<.05$ ) bei männlichen Probanden. Modern Racism (MRS) spielt eine vergleichsweise untergeordnete Rolle für die Ü-Geschichten, es wird lediglich eine signifikante Korrelation zur Wortanzahl bei „Strand“ ( $r=-.34$ ,  $p<.05$ ) erhalten.

Die Zusammenhänge für das Einstellungsobjekt Übergewicht resümierend wird eine konsistente Bedeutung der Wortanzahl für die expliziten Einstellungs- und Motivationsmaße deutlich. Eine negative Einstellung zu Übergewichtigen und eine höhere Motivation zur Vorurteilsfreiheit drücken sich korrelativ in kürzeren Geschichten aus. In zwei Fällen besteht zudem ein negativer Zusammenhang zwischen dem BMI der Probanden und der Wortanzahl, das heißt, schlankere Probanden schreiben längere Geschichten. Für „Ferienlager“ und „Fitnessstudio“ werden zwar erwartungskonforme Zusammenhangstendenzen zwischen den Einstellungsvariablen der Geschichtenmessung (Valenz und Stimmung) und den expliziten Einstellungen erhalten, dennoch erscheint dies nicht ausreichend, eine eindeutige Rangliste nach Eignung der Bildmotive zu erstellen, um eine Wahl der besten Geschichtenmessung(en) treffen zu können.

„Eis“	ÜBKO	BMI	FreundÜ	PNS	MRS	EMS	IMS	MVV	Alter	Geschl.
Valenz	.12	-.11	-.06	.05	-.09	.29	-.38	-.26	-.39	.09
Stimm	-.12	.10	.14	.11	.06	.01	<b>-.48</b>	-.32	-.32	.07
Stereo	.08	.00	.08	.05	-.11	.05	-.29	-.07	-.20	-.13
Ernst	-.19	-.04	.05	<b>-.35</b>	-.25	-.24	.11	-.03	-.09	.08
Anzahl	-.16	<b>-.38</b>	.11	-.12	-.24	-.11	.10	.02	-.15	<b>-.33</b>
Inhalt	-.07	.07	.25	-.03	-.07	-.06	<b>-.42</b>	-.09	-.35	-.31
<b>„Familie“</b>										
Valenz	.05	.11	-.03	.05	-.01	.23	-.10	-.18	.08	.02
Stimm	.06	.12	.01	-.12	-.07	.28	-.12	.28	.08	.00
Stereo	.02	.02	-.10	<b>.35</b>	.14	<b>-.35</b>	-.13	-.21	.17	-.24
Ernst	.18	<b>.42</b>	.16	.13	.16	-.07	-.21	.11	.03	-.08
Anzahl	-.03	-.29	.03	.03	-.14	-.14	.11	.00	-.01	-.26
Inhalt	-.04	<b>-.35</b>	-.31	.30	.14	.23	.08	.00	-.08	.15
<b>„Strand“</b>										
Valenz	.02	.07	.10	-.03	.09	-.01	-.05	.11	-.07	.23
Stimm	-.22	-.03	.16	-.22	-.06	-.14	-.04	.08	-.21	.19
Stereo	-.20	-.09	-.05	.04	-.13	.17	.24	<b>.34</b>	-.03	<b>-.38</b>
Ernst	.18	.18	-.12	-.03	.01	-.05	.16	-.03	.20	-.13
Anzahl	<b>-.33</b>	<b>-.53</b>	-.06	<b>-.54</b>	<b>-.34</b>	<b>-.48</b>	-.07	<b>-.33</b>	<b>-.65</b>	.28
Inhalt	-.28	-.02	.02	-.16	-.10	-.02	.19	.13	-.21	-.13

**Tabelle 5-13.** Zusammenhang zwischen Geschichtenmessung und Fragebogenmaßen für Einstellungsobjekt Übergewicht;  
Fortsetzung nächste Seite

„Schuhwechsel“	ÜBKO	BMI	FreundÜ	PNS	MRS	EMS	IMS	MVV	Alter	Geschl.
Valenz	.01	-.03	.11	-.03	.04	-.25	-.23	-.04	-.12	.14
Stimm	-.05	-.10	.14	-.12	-.12	-.25	-.24	-.09	-.25	.18
Stereo	.27	.11	.17	.41	.29	.05	-.12	.08	.10	-.02
Ernst	-.17	-.33	.12	-.09	.16	-.22	.10	-.29	-.17	-.06
Anzahl	<b>-.47</b>	-.32	.12	<b>-.54</b>	-.26	<b>-.40</b>	-.32	<b>-.41</b>	<b>-.54</b>	.17
Inhalt	-.40	-.25	.24	-.22	-.18	-.19	-.26	-.18	.09	-.43
<b>„Ferienlager“</b>										
Valenz	.11	.26	.24	-.06	.03	.03	<b>.33</b>	.22	.02	-.15
Stimm	.15	.17	-.10	.12	.27	.04	-.02	-.02	.30	.04
Stereo	.26	-.08	<b>-.37</b>	.07	.03	.18	-.06	-.15	-.18	-.11
Ernst	.02	.20	-.12	-.02	.25	.02	-.09	-.16	.15	.25
Anzahl	.12	-.06	-.13	-.17	-.20	-.14	-.17	<b>-.33</b>	-.27	.18
Inhalt	.11	-.03	-.15	-.12	-.08	.15	-.15	-.21	-.21	<b>-.38</b>
<b>„Fitnessstudio“</b>										
Valenz	.20	.16	<b>-.34</b>	.09	.27	.08	.16	.00	.24	.17
Stimm	.01	.02	<b>-.42</b>	.05	.12	.05	.11	-.02	.22	.16
Stereo	<b>.42</b>	.14	<b>-.32</b>	.05	.17	.15	-.12	-.08	.27	<b>.31</b>
Ernst	.08	.15	.16	.14	.06	.01	-.21	-.20	.21	.08
Anzahl	.05	-.17	.02	-.12	-.13	.08	-.02	.03	<b>-.37</b>	-.04
Inhalt	.27	.12	-.06	.22	-.05	.21	.08	-.12	-.23	-.03

**Fortsetzung Tabelle 5-13**, Testheft 1: „Eis“, „Familie“, N=38; Testheft 2: „Strand“, „Schuhwechsel“, N=36; Testheft 3: „Ferienlager“, „Fitnessstudio“, N=43; Abkürzungen siehe Tab. 5-12

### 5.5.3.2 Einstellungsobjekt Homosexualität

Im Gegensatz zu den Ü-Geschichten stehen die einstellungsrelevanten Geschichtenvariablen Valenz und Stimmung in zwei Fällen in engerer positiver Beziehung zu den expliziten Einstellungen zu Homosexualität (Tab. 5-14). Bei „Frühstück“ und „Natur“ werden positive Korrelationen erhalten, die deutlich von Null verschieden sind, jedoch lediglich in einem Fall statistisch bedeutsam sind, und zwar für die Korrelation zwischen *Valenz* und den affektiven Einstellungen zu Homosexualität bei „Natur“ ( $r=.35$ ,  $p<.05$ ). Die Geschichtenlänge zeigt für zwei weitere Geschichtsmessungen die gleiche Relevanz, wie bereits in den Ü-Geschichten herausgestellt werden konnte. Negativere explizite Einstellungen stehen zu kürzeren Geschichten in Beziehung („Couch“:  $r=-.45$ , „Ausstellung“:  $r=-.51$ ,  $p<.01$ ). Für „Couch“ ist, wie bereits die deskriptive Auswertung (Tab. 5-8) ersichtlich macht, die Aktivierung Stereotypen kritisch. Dass es sich jedoch um eine einstellungsrelevante Variable insbesondere in dieser H-Geschichte handelt, zeigen die signifikant negativen Beziehungen zu beiden expliziten Einstellungsmaßen (HOKO:  $r=-.37$ , HOAF:  $r=-.35$ ,  $p<.05$ ) sowie der positive Zusammenhang zur Anzahl homo- oder bisexueller Freunde und Bekannte ( $r=.36$ ,  $p<.05$ ). Eine deutliche Erkennung bzw. Thematisierung des Stereotypen in „Couch“ geht mit einer positiveren Einstellung zu Homosexualität gemäß Fragebogen und einer höheren Anzahl nicht heterosexuell orientierter Personen im Freundes- und Bekanntenkreis einher. Für Ausstellung steht zudem die Ernsthaftigkeit in entscheidendem Zu-

sammenhang zu affektiven Einstellungen ( $r=.37$ ,  $p<.05$ ), dergestalt, dass ernst zu nehmende Geschichten insbesondere von Probanden mit negativeren affektiven Einstellungen zu Homosexualität verfasst werden.

Die Anzahl homo- oder bisexueller Freunde und Bekannte korreliert für „Couch“ und „Ausstellung“ gleichermaßen positiv mit der Wortanzahl ( $r=.47$ ,  $p<.01$  bzw.  $r=.62$ ,  $p<.001$ ). Dies trifft auch für „Natur“ zu, wobei die Korrelation nicht statistisch bedeutsam ist ( $r=.29$ ,  $p>.05$ ). Für „Werkshalle“ wird dagegen ein negativer, jedoch ebenfalls nicht signifikanter Zusammenhang zwischen der Geschichtenlänge und der Anzahl homo- oder bisexueller Freunde beobachtet ( $r=-.28$ ,  $p>.05$ ). Zudem steht hier die Stimmung in erwartungskonträr negativer Beziehung zur Kontaktvariablen *FreundH*, dieser Zusammenhang ist signifikant ( $r=.46$ ,  $p<.01$ ) und zeigt an, dass Probanden, die angeben, mehr homo- oder bisexuelle Freunde zu haben, Geschichten zum Bildmotiv „Werkhalle“ schreiben, die eine negativere Grundstimmung bzw. ein pessimistischer Ausgang haben.

Das Bedürfnis nach kognitiver Struktur (PNS) weist sinnvolle Zusammenhänge insbesondere zur Geschichtenlänge („Couch“:  $r=-.66$ ,  $p<.001$ , „Ausstellung“:  $r=-.56$ ,  $p<.01$ , Urlaub“:  $r=-.32$ ,  $p<.05$ ) auf. Für „Frühstück“ korreliert PNS positiv ( $r=.38$ ,  $p<.05$ ), für „Ausstellung“ dagegen negativ mit dem Inhaltsreichtum ( $r=-.43$ ,  $p<.05$ ). Eine etwas höhere Bedeutung als für Übergewicht weist Modern Racism (MRS) innerhalb der H-Geschichten auf. Die Variable steht in negativem Zusammenhang zur Wortanzahl bei „Natur“ ( $r=-.34$ ,  $p<.05$ ) sowie in negativer Beziehung zur Stereotyperkennung bei „Ausstellung“ ( $r=-.39$ ,  $p<.05$ ) und zum Inhaltsreichtum bei „Urlaub“ ( $r=-.37$ ,  $p<.05$ ).

Wie bereits für die Ü-Geschichten, zeigt sich hier die Relevanz der Motivationsvariablen für die Geschichtenlänge. In drei Fällen werden signifikant negative Zusammenhänge zwischen EMS bzw. MVV und der Wortanzahl erhalten (EMS: „Couch“:  $r=-.47$ ,  $p<.01$ , „Ausstellung“:  $r=-.41$ ,  $p<.01$ ; MVV: „Ausstellung“:  $r=-.39$ ,  $p<.05$ ). Eine höhere externale Motivation zur Vorurteilsfreiheit geht mit negativerer Valenz bei „Couch“ ( $r=.34$ ,  $p<.05$ ) und „Werkshalle“ ( $r=.35$ ,  $p<.05$ ) einher, während für „Frühstück“ der erwartungsgemäße Zusammenhang zwischen hoher externaler Motivation zu Vorurteilsfreiheit und positiverer Grundstimmung besteht ( $r=-.36$ ,  $p<.05$ ). Die internale Motivation zur Vorurteilsfreiheit (IMS) hat kaum Bedeutung für die H-Geschichten, sie korreliert ebenso wie EMS positiv zur Valenz bei „Werkhalle“ ( $r=.36$ ,  $p<.05$ ). Auch MVV steht bei „Werkhalle“ in positiver Beziehung zur Valenz ( $r=.42$ ,  $p<.01$ ).

Analog zu den Ü-Geschichten werden auch für Homosexualität je kürzere Geschichten geschrieben, desto älter die Probanden sind, für „Couch“ und „Ausstellung“ sind die Korrelation signifikant ( $r=-.49$ ,  $p<.01$  bzw.  $r=-.60$ ,  $p<.001$ ). Ältere Versuchsteilnehmer erkennen bzw. thematisieren darüber hinaus den Stereotyp bei „Frühstück“ weniger deutlich ( $r=-.39$ ,  $p<.05$ ), schreiben eine negativere Geschichte zu „Couch“ ( $r=.37$ ,  $p<.05$ ) und eine spöttischere, sarkastischere bzw. zynischere Geschichte zu „Ausstellung“ ( $r=-.38$ ,  $p<.05$ ). Das Geschlecht hat für „Frühstück“ und „Natur“ vergleichbare Be-

deutung zu den Ü-Geschichten. Frauen schreiben hier längere Geschichten als Männer ( $r=-.41$  bzw.  $r=-.39$ ,  $p<.05$ ). Auch für die Valenz bei „Frühstück“ sowie die Stimmung für „Werkhalle“ bestehen Geschlechtsunterschiede, Frauen schreiben positivere „Frühstück“-Geschichten ( $r=.40$ ,  $p<.05$ ) dagegen aber pessimistischere „Werkhalle“-Geschichten ( $r=-.35$ ,  $p<.05$ ). Darüber hinaus verfassen die Frauen inhaltsreichere Geschichten zu „Couch“ und „Ausstellung“ und schreiben weniger ernst zu nehmend zu „Urlaub“ ( $r=.32$ ,  $p<.05$ ).

Auch für die H-Geschichten zeichnet sich die Wortanzahl als ein stabiles und zudem stimmig interpretierbares Korrelat expliziter Einstellungen und vorurteilsbezogener Motivation aus. Zwar weisen „Couch“ und „Ausstellung“ signifikante Korrelationen zu vielen Fragebogenmaße auf, dennoch bestehen hier nicht, wie für „Frühstück“ und „Natur“ der Fall, erwartungskonforme Zusammenhänge zwischen den affektiven Aspekten der Geschichten, also den Variablen Valenz und Stimmung, und den expliziten Einstellungen zu Homosexualität gemäß Fragebogen. Die interne Motivation zur Vorurteilsfreiheit spielt innerhalb der H-Geschichten eine untergeordnete Rolle, wogegen die externe Motivation zur Vorurteilsfreiheit wie auch die Motivation zu vorurteilsfreiem Verhalten einige signifikante Zusammenhänge zu diversen Geschichtenvariablen unterhalten. Hierbei sind jedoch deutliche Unterschiede zwischen den Bildmotiven bzw. den Stichproben zu beobachten.

<b>„Frühstück“</b>	<b>HOKO</b>	<b>HOAF</b>	<b>FreundH</b>	<b>PNS</b>	<b>MRS</b>	<b>EMS</b>	<b>IMS</b>	<b>MVV</b>	<b>Alter</b>	<b>Geschl.</b>
Valenz	.21	.30	.04	-.29	-.17	-.21	-.09	-.18	.17	.40
Stimm	.20	.27	.02	-.28	-.18	-.36	-.02	-.10	-.06	.19
Stereo	.03	-.11	.08	-.17	-.11	.01	-.03	-.14	-.39	.26
Ernst	.04	.17	-.04	.10	.04	-.06	-.10	.02	.09	-.21
Anzahl	-.17	-.17	.10	-.03	-.02	-.05	.00	-.12	-.24	-.41
Inhalt	-.14	-.14	.13	.38	.25	.25	-.07	.13	.09	-.05
<b>„Natur“</b>										
Valenz	.24	.35	-.02	.09	-.10	-.04	-.18	-.28	-.07	.13
Stimm	.27	.21	-.11	.10	.04	-.07	-.24	-.21	-.03	.08
Stereo	.04	-.06	.07	.08	.01	-.24	-.16	-.18	-.05	.20
Ernst	-.11	.02	-.02	.11	.08	-.22	-.09	.24	.15	-.23
Anzahl	-.22	-.30	.29	-.17	-.34	-.13	.16	.16	-.04	-.39
Inhalt	-.21	-.03	-.04	-.11	-.04	-.17	-.31	-.01	-.25	-.13
<b>„Couch“</b>										
Valenz	.16	.02	-.18	.13	.33	.34	.03	.06	.37	.07
Stimm	.03	.03	-.08	-.04	.06	.21	.10	.20	.20	.09
Stereo	-.37	-.35	.36	-.20	-.27	-.34	-.10	-.20	-.18	.00
Ernst	-.08	-.10	.15	-.14	-.10	.01	.07	-.01	-.19	-.19
Anzahl	-.45	-.18	.47	-.66	-.29	-.47	-.14	-.27	-.49	.08
Inhalt	.04	.24	.00	-.02	.00	-.01	-.06	-.01	.08	-.37

**Tabelle 5-14**, Zusammenhang zwischen Geschichtenmessung und Fragebogenmaßen für Einstellungsobjekt Homosexualität; *Fortsetzung nächste Seite*

„Ausstellung“	HOKO	HOAF	FreundH	PNS	MRS	EMS	IMS	MVV	Alter	Geschl.
Valenz	.03	-.06	-.02	.24	.22	.09	-.07	.05	.16	-.26
Stimm	-.07	-.07	.13	.05	.13	-.17	-.03	-.18	.02	.08
Stereo	-.13	-.23	.15	-.25	-.39	-.27	.16	-.20	-.19	.12
Ernst	.01	.37	.15	.03	-.29	-.08	-.23	-.14	-.38	.16
Anzahl	<b>-.51</b>	-.27	<b>.62</b>	<b>-.56</b>	-.33	-.41	-.22	-.39	<b>-.60</b>	.27
Inhalt	-.39	-.16	.42	-.43	-.24	<b>-.48</b>	.16	-.30	.21	-.34
<b>„Urlaub“</b>										
Valenz	.03	.17	-.01	-.25	.08	.01	.19	.02	.16	.22
Stimm	-.01	.06	-.02	-.10	.09	.24	-.23	.03	.20	.14
Stereo	-.02	.04	.08	.13	.06	-.13	.21	<b>.35</b>	.15	-.25
Ernst	.14	.23	.00	-.05	.03	.08	-.13	-.11	.15	<b>.32</b>
Anzahl	-.28	-.20	-.03	<b>-.32</b>	-.15	-.07	-.07	.28	-.16	.11
Inhalt	<b>-.38</b>	-.15	-.03	-.21	-.37	-.01	.09	.25	.02	-.22
<b>„Werkshalle“</b>										
Valenz	.06	.08	.23	-.10	.12	<b>.35</b>	<b>.36</b>	<b>.42</b>	.17	.02
Stimm	.03	-.28	<b>.46</b>	.24	.04	.18	.25	.01	.08	<b>-.35</b>
Stereo	-.21	.08	.21	-.18	-.10	-.02	-.03	-.13	-.22	.19
Ernst	.05	.10	-.01	.05	.06	-.09	-.09	.02	.17	.08
Anzahl	-.03	-.02	-.28	-.23	-.13	-.01	-.08	-.11	-.26	.13
Inhalt	.10	.00	-.25	-.02	.07	.09	-.08	.14	.03	.16

**Fortsetzung Tabelle 5-14,** Testheft 1: „Frühstück“, „Natur“, N=38; Testheft 2: „Couch“, „Ausstellung“, N=36; Testheft 3: „Urlaub“, „Werkshalle“, N=43; Abkürzungen siehe Tab. 5-12

### 5.5.3.3 Befunde für die Gesamtstichprobe

Im Folgenden werden die drei Teilstichproben zusammengeführt, um eine über die einzelnen Ü- bzw. H-Geschichten der drei Testhefte generalisierende Gegenüberstellung von Geschichtenmessung und Fragebogenmaßen vornehmen zu können (Tab. 5-15).

Mit erstaunlicher Stabilität zeichnet sich die *Geschichtenlänge* als Einstellungs- bzw. Vorurteilsskorrelat aus. Sowohl die Ü-Geschichten, als auch die H-Geschichten fallen länger aus bei Personen, die positivere explizite Einstellungen haben, wobei der Zusammenhang nur für die H-Geschichten signifikant ist ( $r=-.26$ ,  $p<.01$ ). Außerdem bestehen über beide Einstellungsobjekte hinweg generalisierbare negative Zusammenhänge zwischen der Wortanzahl der Geschichten und PNS (Ü:  $r=-.27$ ,  $p<.01$ ; H:  $r=-.33$ ,  $p<.001$ ), MRS (Ü:  $r=-.21$ , H:  $r=-.19$ ,  $p<.05$ ), EMS (Ü:  $r=-.21$ , H:  $r=-.20$ ,  $p<.05$ ) und dem Alter (Ü:  $r=-.36$ , H:  $r=-.32$ ,  $p<.001$ ). Längere Geschichten werden nicht nur von Personen mit weniger Vorurteilen geschrieben, sondern zugleich von Probanden, die ein geringeres Bedürfnis nach kognitiver Strukturiertheit, geringere Ausprägungen in Modern-Racism und weniger externale Motivation zur Vorurteilsfreiheit ausweisen und zudem von jüngeren Probanden. Für das Einstellungsobjekt Homosexualität steht zudem eine höhere Anzahl homo- oder bisexueller Freunde und Bekannte mit längeren Geschichten in Beziehung ( $r=.21$ ,  $p<.05$ ). Die objektiv messbare Variable der Länge einer Geschichte könnte demnach als Einstellungs- und Motivationsindikator dienen.

Die *Valenz*, welche rein konzeptuell die größten inhaltlichen Ähnlichkeiten zu expliziten Einstellungen aufweist, steht für beide Einstellungsobjekte in erwartungsgemäß positivem Zusammenhang zu den Fragebogenskalen. Dabei ist zunächst nur die Korrelation zwischen der Valenz der Ü-Geschichten und der expliziten Einstellungen zu Übergewicht signifikant ( $r=.21, p<.05$ ). Werden kognitive und affektive Einstellungen zu Homosexualität in einem Score gemittelt, ist auch hierfür der Zusammenhang zur Valenz der H-Geschichten statistisch bedeutsam ( $r=.18, p<.05$ ).

Die *Stereotypenerkennung* ist, wie oben bereits angedeutet wurde, von besonderer Relevanz für die H-Geschichten, wobei hier offensichtlich weniger die vermutete Problematik gelungener Stereotypenerkennung, als vielmehr eine bewusste Motivation zum Erkennbar-werden-lassen angezeigt wird. Je stärker ein Proband external zur Vorurteilsfreiheit motiviert ist, desto weniger deutlich thematisiert er/ sie das Einstellungsobjekt ( $r=-.22, p<.05$ ). Zudem korrelieren die Stereotyperkennung und Modern Racism (MRS) negativ miteinander ( $r=-.22, p<.05$ ). Demzufolge nach geht das Nichterkennen der Thematik Homosexualität bzw. eine verdeckte Thematisierung mit höheren MRS-Ausprägungen einher.

<b>Übergewicht</b>										
	<b>ÜBKO</b>	<b>BMI</b>	<b>FreundÜ</b>	<b>PNS</b>	<b>MRS</b>	<b>EMS</b>	<b>IMS</b>	<b>MVV</b>	<b>Alter</b>	<b>Geschl.</b>
<b>Valenz</b>	.21	.13	.09	.03	.08	.12	.00	-.07	-.05	.15
<b>Stimm</b>	-.01	-.05	.00	-.05	-.06	-.03	-.03	-.01	-.18	.06
<b>Stereo</b>	.15	.10	.04	.15	.10	.14	-.08	-.03	-.01	-.09
<b>Ernst</b>	.04	.15	.16	.02	.08	-.13	-.08	-.05	.15	-.11
<b>Anzahl</b>	-.14	<b>-.29</b>	.03	<b>-.28</b>	-.21	-.21	-.07	-.20	<b>-.36</b>	.01

<b>Homosexualität</b>										
	<b>HOKO</b>	<b>HOAF</b>	<b>FreundH</b>	<b>PNS</b>	<b>MRS</b>	<b>EMS</b>	<b>IMS</b>	<b>MVV</b>	<b>Alter</b>	<b>Geschl.</b>
<b>Valenz</b>	.16	.17	-.05	-.05	.13	.16	-.02	-.11	.18	.12
<b>Stimm</b>	.00	.01	.06	-.04	.07	.03	.00	-.17	.05	-.06
<b>Stereo</b>	-.15	-.11	.17	-.04	-.22	-.22	.02	-.05	-.22	.01
<b>Ernst</b>	.04	.10	-.11	.10	-.01	.00	-.10	.03	.04	-.06
<b>Anzahl</b>	<b>-.26</b>	-.19	.21	<b>-.33</b>	-.19	-.20	-.06	-.12	<b>-.32</b>	-.02

**Tabelle 5-15,** Korrelationen zwischen formellen Kriterien und Fragebogenmaßen für die Gesamtstichprobe (N=119), Abkürzungen siehe Tab. 5-12

Abschließend werden für beide Einstellungsobjekte die einstellungsrelevanten Maße aus der Geschichtsmessung (Valenz, Anzahl) den expliziten Einstellungen gegenübergestellt. Die Korrelationsmatrix repräsentiert die Zusammenhänge für die drei unterschiedlichen Testhefte bzw. Stichproben und kann aufgrund dessen nicht wie eine Multitrait-Multimethod-Matrix interpretiert werden. Dennoch ermöglicht diese Gegenüberstellung eine Abschätzung der Beziehungen innerhalb und zwischen den Maßen einer Einstellungsdisposition (Tab 5-16).

	1.	2.	3.	4.	5.
1. <b>Valenz-Ü</b>					
2. <b>Anzahl-Ü</b>	.13				
3. <b>ÜBKO</b>	.22	-.14			
4. <b>Valenz-H</b>	<b>.38</b>	-.21	.18		
5. <b>Anzahl-H</b>	.06	<b>.85</b>	-.14	-.17	
6. <b>HOKOAF</b>	.16	<b>-.24</b>	<b>.32</b>	.18	<b>-.26</b>

**Tabelle 5-16**, Zusammenhänge zwischen Geschichtenmessung und expliziten Einstellungen über beide Einstellungsobjekte hinweg für Gesamtstichprobe  
 Valenz-Ü/H: Valenz der Ü- bzw. H-Geschichten, Anzahl-Ü/H= Wortanzahl der Ü- bzw. H-Geschichten, HOKOAF= Mittelwert kognitiver und affektiver Einstellungen zu Homosexualität (HOKO, HOAF)

Wie erwartet zeigt sich der leichte positive, jedoch signifikante Zusammenhang zwischen der Valenz der Geschichten und den expliziten Einstellungen innerhalb eines Einstellungsobjekts (Ü:  $r=.22$ , H:  $r=.18$ ,  $p<.05$ ), jedoch wird für eine der beiden Korrelationen *zwischen* den Einstellungsobjekten eine ebensolche Beziehung beobachtet, und zwar zwischen der Valenz der H-Geschichten und ÜBKO – den expliziten Einstellungen zu Übergewicht ( $r=.18$ ,  $p<.05$ ). Die Wortanzahl korreliert für Homosexualität signifikant mit den expliziten Einstellungen ( $r=-.26$ ,  $p<.01$ ), dies gilt nicht für das Einstellungsobjekt Übergewicht. Dagegen besteht jedoch eine ebenfalls signifikant negative Beziehung zwischen der Länge der Ü-Geschichten und den expliziten Einstellungen zu Homosexualität ( $r=-.24$ ,  $p<.01$ ). Die hohe Korrelation zwischen den Umfängen der Ü- und H-Geschichten ( $r=.85$ ,  $p<.001$ ) reflektiert sicherlich eine individuelle Schreibgewohnheit, mehr oder weniger Wort zu machen. Die Valenz von Ü- und H-Geschichten, ebenso wie explizite Einstellungen zu Übergewicht und Homosexualität stehen in vergleichsweise stärkerem Zusammenhang zueinander ( $r=.32$  bzw.  $r=.38$ ,  $p<.001$ ), was auf methodenspezifische Varianz in den Geschichten- wie auch den Fragebogenmaßen beruhen kann.

## 5.6 Expertenratings

Abschließend und unabhängig von der empirischen Untersuchung der Bildmotive werden zwei zusätzliche Expertenratings eingeholt. Diese dienen dazu, für die Bildmotive eine Rangliste von den besseren zu den schlechter geeigneten Bildern erstellen zu können.

### 5.6.1 Expertenrating I

Neben der Auswertung der Bildmotive anhand der empirischen Ergebnisse werden Expertenurteile zur Augenscheinvalidität des Bildmaterials für eine Einstellungsmessung eingeholt. Acht Experten beurteilten jedes der zwölf Bilder auf jeweils einer siebenstufigen Likert-Skala „Eignung zur Erfassung

von Vorurteilen gegenüber Homosexuellen bzw. Übergewichtigen“ (unabhängige Bewertung) und vergaben zusätzlich für die sechs Bildmotive eines Einstellungsobjektes Rangplätze von 1 „am besten geeignet“ bis 6 „am wenigsten geeignet (Rangvergabe). Werden die Rangplätze umkodiert, so dass hohe Werte einen unteren Rang anzeigen, können beide Experteneinschätzungen pro Bildmotiv zu einem multiplikativen Gesamtscore zusammengefasst werden (Gesamturteil). Tabelle 5-17 beinhaltet die Ergebnisse der Expertenbefragung, aufsteigend sortiert nach Expertenurteil.

		Unabhängige Bewertung			Rangplatz			Gesamt- Urteil
		Min	Max	M (SD)	Min	Max	M (SD)	M
Übergewicht	<b>Eis</b>	5	7	6.00 (.76)	5	6	5.63 (.52)	34.00
	<b>Sch</b>	5	6	5.50 (.54)	1	5	4.00 (1.41)	22.38
	<b>Stra</b>	4	6	5.13 (.84)	2	6	4.25 (1.39)	22.38
	<b>Fam</b>	2	6	4.63 (1.30)	1	6	3.00 (1.85)	15.50
	<b>Fit</b>	2	5	4.00 (1.20)	1	4	2.63 (.92)	11.38
	<b>Fer</b>	2	6	3.88 (1.13)	1	4	2.25 (1.04)	9.63
Homosexuell	<b>Früh</b>	2	7	5.25 (1.75)	1	6	4.75 (1.83)	27.38
	<b>Nat</b>	1	7	5.06 (1.94)	1	6	3.81 (2.03)	21.78
	<b>Url</b>	3	7	4.94 (1.47)	1	5	4.10 (1.15)	20.41
	<b>Werk</b>	3	6	3.87 (1.13)	2	5	3.38 (1.30)	13.88
	<b>Cou</b>	2	5	3.13 (.99)	1	3	2.25 (.89)	7.50
	<b>Aus</b>	2	5	2.75 (1.17)	1	4	1.75 (1.04)	5.75

**Tabelle 5-17**, Expertenrating der Bildmotive, N=8; Gesamturteil= Produkt aus unabhängiger Bewertung und Rangplatz;

Gemäß den Expertenratings können je Einstellungsobjekt drei relativ geeignete und drei weniger geeignete Bilder voneinander unterschieden werden. Für Übergewicht werden „Eis“, „Schuhwechsel“ und „Strand“ als die drei ranghöchsten Bildmotive eingeschätzt, wobei Eis mit weitem Anstand als bestes Bildmotiv beurteilt wird und die beiden letztgenannten exakt das gleiche Gesamturteil erzielen, da „Schuhwechsel“ zwar ein wenig besser in der unabhängigen Eignungseinschätzung abschneidet, jedoch einen niedrigeren Rangplatz als „Strand“ erhält. Für Homosexualität gelangen gemäß der Expertenratings „Frühstück“, „Natur“ und „Urlaub“ auf die ersten Ränge, wobei auch hier zwischen dem zweit- und drittplatzierten Motiv lediglich geringfügige Unterschiede im Gesamturteil bestehen.



### 5.6.2 Expertenrating II

Um eine differenziertere Rangliste der Bildmotiveignung erstellen zu können und gemäß dem ersten Expertenrating insbesondere Entscheidungen über die jeweils zweiten und dritten Rangplätze treffen zu können, wird ein zweites Expertenrating eingeholt. Einer anderen Gruppe von zehn unabhängigen Experten werden die beiden kritischen Bildpaare „Natur“ vs. „Urlaub“ und „Strand“ vs. „Schuh“ zur Auswahl des jeweils besten Bildes für eine Einstellungsmessung vorgelegt. Zur Auswertung der Urteile, werden die Stimmen für jedes Bild aufsummiert.

Das Ergebnis der Befragung spiegelt deutlich die Ähnlichkeit der konkurrierenden Bilder wider. Sechs Experten wählten „Natur“ als besseres Bild und entsprechend fallen vier Expertenurteile zugunsten von „Urlaub“ aus (Einstellungsobjekt Homosexualität), wogegen sich für die beiden Übergewichtsbilder „Strand“ und „Schuh“ die Einschätzung die Waage halten (5:5).

## 5.7 Diskussion

Mit Studie II wurde eine offene Assoziativmessung entwickelt, die methodisch an den Thematischen Apperzeptionstest angelehnt ist und sowohl bewusste, reflektierte, als auch unbewusste, spontane Einstellungen zu einerseits Übergewicht und andererseits Homosexualität erfasst. Dazu wurden jeweils sechs Bildmotive erstellt und Probanden vorgelegt mit der Aufgabenstellung, eine zur Darstellung assoziierte Geschichte zu verfassen. Neben den Bildmotiven bedurfte es der Entwicklung adäquater Kodierschemata für das qualitative Geschichtenmaterial. Die Kodierung setzte sich aus einem bildspezifischen inhaltlichen und einem allgemeingültigen Teil zusammen, wobei die für jedes Bildmotiv spezifischen Inhaltsvariablen aufgrund fehlender Vergleichsmöglichkeiten mehr der Kontrolle dienten. Der allgemeine Teil, den alle Kodierschemata gemein haben, fokussiert zum einen auf die Erfassung einstellungsrelevanter, emotionaler Aspekte mittels Variablen, die anzeigen, inwiefern eine Geschichte eine Wertigkeit ausdrückt (Valenz) und welche Ausrichtung einer Geschichte verliehen wird (Stimmung). Zum anderen konnten motivationale Aspekte berücksichtigt werden mit Variablen, die mehr oder weniger als Kontrolle angelegt waren, die Einschätzbarkeit einer Geschichte zu verbessern (Stereotyperkennung, Ernsthaftigkeit, Wortanzahl).

Ziel von Studie II ist es, die Bildmotive für jedes Einstellungsobjekt hinsichtlich ihrer Güte zur Erfassung von Einstellungen und einstellungsrelevanten Aspekten einschätzen zu können, um die am meisten geeigneten Bilder für eine Erhebung auswählen zu können. Als Kriterien für die Evaluation der Bildmotive wird insbesondere die Gegenüberstellung der Geschichtenmessung mit den ebenfalls erhobenen selbstberichteten Einstellungs- und Motivationsmaßen zu Rate gezogen. Hierbei besteht jedoch zunächst die Schwierigkeit der Vergleichbarkeit. Die drei Teilstichproben, in denen jeweils vier

Bildmotive getestet wurden, sind zu klein, um eine ausgewogene Merkmalsverteilung annehmen zu können. Zwar werden in keinem der einstellungsrelevanten Fragebogenmaße signifikante Unterschiede zwischen den drei Stichproben festgestellt, dennoch weisen sie zum Teil in einigen demografischen Merkmalen, wie dem Alter, deutliche Differenzen auf. Zudem muss die Methodenspezifität beider Messungen berücksichtigt werden, die zu systematischen Verzerrungen der Zusammenhänge zwischen den Geschichtenvariablen und den Fragebogenmaßen führen kann. Der Betrachtung der aggregierten Geschichten- und Fragebogenmaße zufolge, kann davon ausgegangen werden, dass die Maße aus offener Assoziativmessung und Selbstbericht neben inhaltlicher Einstellungsvarianz auch einen gewissen Anteil methodenspezifischer Varianz aufweisen. Diese Methodeneffekte sind zwar substantiell, jedoch nicht derart stark, dass sie die Varianz resultierend aus Unterschieden in den Einstellungsdispositionen völlig überlagern. Diese Befundlage im Sinne konvergenter und diskriminanter Validität zu interpretieren ist jedoch wiederum aufgrund der Stichprobencharakteristik nicht zulässig.

Trotz der Dreiteilung der Stichprobe und der jeweils relativ geringen Stichprobenumfänge zeigt die Zusammenhangsanalyse eine Reihe konsistenter Beziehungen zwischen offener Assoziations- und Fragebogenmessung auf. Dies gilt relativ ungeachtet des Bildmotivs insbesondere für die Geschichtenlänge gemessen in Worten. Hierbei findet sich eine stabile Beziehung, die umso weniger geschriebene Worte aufzeigt, desto negativer die expliziten Einstellungen via Fragebogen ausfallen. Die Geschichtenlänge reflektiert zudem motivationale Aspekte, da stabile negative Beziehungen zwischen der Wortanzahl und den einstellungs- bzw. vorurteilsbezogenen Motivationsmaßen des Fragebogens bestehen. Die trifft insbesondere für die externale Motivation zur Vorurteilsfreiheit und die Motivation zu vorurteilsfreiem Verhalten, beides also stärker extrinsisch orientierte Dispositionen, zu. Das bedeutet also, je stärker jemand nach außen hin motiviert ist, wenige Vorurteile zu haben, desto kürzer fällt die Geschichte zu einem entsprechenden Bildmotiv aus. Dahinter könnte sich bspw. eine Vorsichtsmaßnahme, nicht doch aus Versehen etwas Negatives zu schreiben, verbergen, was mit einer von Blair (2001) ausgewiesenen Strategie zur Vorurteilkontrolle – der Unterdrückung – in Zusammenhang gebracht werden kann.

Neben der konsistent negativen Beziehung zur Geschichtenlänge, ergibt sich für die selbstberichteten Motivationsmaße sich ein relativ inkonsistentes Beziehungsmuster zu den Geschichtenvariablen beider Einstellungsobjekte. Für Übergewicht wird in einem Fall ein negativer Zusammenhang zwischen internaler Motivation zur Vorurteilsfreiheit und negativer Geschichtensvalenz aufgedeckt, in einem anderen Fall dagegen ein positiver. Die externale Motivation zur Vorurteilsfreiheit ist ebenfalls zum Teil in positiven Zusammenhang zur Valenz zu bringen, insbesondere innerhalb der H-Geschichten, zum Teil werden hier aber ebenfalls negative Korrelationen aufgefunden. Es ist also nicht allgemeingültig und eindeutig festzulegen inwiefern die Motivation zur Vorurteilsfreiheit als Indikator geeigneter Bildmotive zu interpretieren ist.

Die Geschichtenvariable, von der eine enge theoretische Beziehung zu den selbstberichteten Einstellungen angenommen wird, ist die Valenz. Auch hierfür findet sich ein erwarteter Zusammenhang zu expliziten Einstellungen über alle Bildmotive eines Einstellungsobjektes hinweg. Dennoch kann aufgrund der Valenz keine eindeutige Aussage über die relative Eignung der Bildmotive beider Einstellungsobjekte getroffen werden. Für Übergewicht wären nach dem Ausschlussprinzip „Familie“ und „Ferienlager“ zunächst hintan zu stellen. Für beide Bildmotive werden nur vereinzelte, wenige aussagekräftige Zusammenhänge zu den Fragebogenmaßen beobachtet. Relevant für die Motivation zur Vorurteilkontrolle erweisen sich insbesondere die Bilder „Eis“, „Strand“ und „Schuhwechsel“, wobei insbesondere für „Eis“ und „Schuhwechsel“ auch die Valenz der Geschichten eine Tendenz aufweist, zu vorurteilsbezogener Motivation in Beziehung zu stehen, während für „Strand“ durchweg Korrelationsbeträge nahe Null erhalten werden. Für das Einstellungsobjekt Homosexualität sind es „Frühstück“ und „Natur“, deren Valenz zu den expliziten Einstellungen gemäß Fragebogen in stärkstem Zusammenhang stehen. Für alle anderen H-Geschichten weichen die Korrelationen zwischen Valenz und kognitiven sowie affektiven Einstellungen nicht deutlich von Null ab. Dagegen liefern die Variablen der Bildmotive „Couch“ und „Ausstellung“ viele erwartungskonsistente Zusammenhänge zu den Fragebogenmaßen der Motivation und der Anzahl homo- oder bisexueller Freunde und Bekannte. An dieser Stelle kann keine eindeutige Entscheidung über die am meisten geeigneten Motive für das Einstellungsobjekt Homosexualität getroffen werden.

Als eine zweite Evaluationsmöglichkeit der Bildmotive wurden zwei Gruppen unabhängiger Experten zur Beurteilung des Bildmaterials die nach Augenscheinvalidität befragt, wobei das zweite Expertenrating dazu diente, Entscheidungen über die für jedes Einstellungsobjekt strittigen zweiten und dritten Rangplätze der Bildmotive treffen zu können. Gemäß den beiden Expertenratings wird für das Einstellungsobjekt Übergewicht „Eis“ als bestgeeignetes und „Strand“ und „Schuhwechsel“ als jeweils zweitbeste Bilder ausgewählt. Die letzten drei Ränge bekleiden „Familie“, „Fitnessstudio“ und „Ferienlager“. Hier zeigt sich deutlich die Übereinstimmung mit der empirischen Evaluation der Bildmotive, deren zufolge „Eis“, „Strand“ und „Schuhwechsel“ als am beste geeignet eingeschätzt werden.

Für die Messung von Einstellungen zu Homosexualität parallelisieren die Expertenratings die valenzbasierte Evaluation der Bildmotive über die Zusammenhänge zum Fragebogen, indem sie „Frühstück“ und „Natur“ als die besten Bilder wählen. Die laut empirischer Untersuchung ebenfalls als geeignet zu bewertenden Bilder „Couch“ und „Ausstellung“ haben entsprechend der Expertenmeinung die letzten beiden Ränge inne. Hierin zeigt sich eine große Variabilität der Einschätzung. Aus der Sicht der Autorin lässt sich das Bild „Werkhalle“ aufgrund erwartungskonsistenter Beziehungsmuster zu den Fragebogenmaßen ausschließen, wobei die verbleibenden fünf Bildmotive mehr oder weniger alle gut geeignet scheinen, Einstellungs- und einstellungsbezogene Motivationsinhalte zu Übergewicht und Homosexualität zu erfassen.

Den überzufälligen, häufig erwartungskonformen und betragsmäßig moderaten Zusammenhängen zwischen expliziter Einstellungsmessung gemäß Fragebogen und den Geschichtenvariablen zufolge, kann davon ausgegangen werden, dass die offene Assoziativmessung einen Anteil expliziter Einstellungen zu Übergewicht und Homosexualität erfasst. Darüber hinaus reflektiert die Geschichtenmessung einstellungsrelevante Motivationen. Inwieweit die Befunde aus aktueller Untersuchung generalisierbar sind und mit der offenen Assoziativmessung auch implizite Einstellungsanteile gemessen werden können, wird mit Studie III prüfbar. Diese abschließende empirische Studie dient der integrativen Untersuchung einer Reihe globaler Fragestellungen, von denen der Konstruktvalidierung von Einstellungen die zentrale Bedeutung zukommt. Mit Studie III kann zugleich die Replizierbarkeit der Befunde aus Studie I und II geprüft werden, da als Messinstrumente für Einstellungen zu Übergewicht und Homosexualität explizite Fragebogenskalen, implizite Maße kognitiver Assoziationsstärke sowie die offene Assoziativmessung eingesetzt werden.

## 6 Studie III: Konstruktvalidierung von Einstellungsmaßen aus expliziter, impliziter und offener Assoziativmessung

### 6.1 Theoretische Grundlagen und Anliegen von Studie III

Die theoretischen Grundlagen von Studie III sind dem einleitenden Teil vorliegender Arbeit nahezu deckungsgleich, da mit dieser Untersuchung entscheidende Informationen zur Konzeptualisierung von expliziten und impliziten Einstellungen erbracht werden, was den Hauptfokus darstellt. Entsprechend dient Studie III der Untersuchung folgender zentraler Fragestellungen:

- (1) *Inwiefern kann für explizite und implizite Einstellungen zu Übergewicht und Homosexualität konvergente und diskriminante Validität angenommen werden?  
Wie bedeutsam sind Effekte der Einstellungsdisposition (Traitvarianz) im Verhältnis zu methodenspezifischen Effekten der expliziten, impliziten und offenen Assoziativmessung (Methodenvarianz)?*
- (2) *Inwieweit und durch welche der anderen Einstellungsmessungen können explizite Einstellungen vorhergesagt werden. Ist es zulässig, eine Explizit-Implizit-Dimension der Einstellungen anzunehmen?*
- (3) *Lassen sich, explizit erfasst, individuelle von gesellschaftlich antizipierten Vorurteilen trennen und welche Bedeutung hat eine solche Differenzierung für implizite Maße bzw. den Zusammenhang zwischen expliziten und impliziten Einstellungen?*
- (4) *Können die in der Literatur berichteten Moderatorwirkungen der Motivation zur Vorurteilsfreiheit in der Beziehung impliziter und expliziter Einstellungsmaßen auf vorliegende Fragestellung appliziert werden?*
- (5) *Welche Bedeutung haben explizite und implizite soziale Einstellungen (Nomologische Netzwerk)?*

Hauptziel ist die Konstruktvalidierung expliziter und impliziter Einstellungen. Dabei geht es weniger um eine allgemeingültig Entscheidung für entweder das Dissoziations- oder Dual-Attitude-Modell (Devine, 1989; Wilson, Lindsey & Schooler, 2000) oder den Same-Construct-Approach (Wittenbrink et al., 1997; vgl. Abschnitt 2.1.1), sondern zunächst um die klassische Untersuchung diskriminanter und konvergenter Validität expliziter und impliziter Einstellungen, die Abschätzung dispositioneller und methodischer Varianzanteile, unter Berücksichtigung des Messfehlers. Hierzu werden die in den vorangegangenen Studien isoliert untersuchten Einstellungsmessungen gemeinsam an einer ausreichend großen Stichprobe eingesetzt. Der daraus resultierende Datenpool umfasst selbstberichtete explizite Einstellungen, implizite Einstellungen über kognitive Assoziationsstärken, Einstellungen mit sowohl expliziten, als auch impliziten Anteilen gemäß der offenen Assoziativmessung sowie fremdbberichtete, gesellschaftlich antizipierte Einstellungen. Darüber hinaus werden für jede Person zusätzliche

einstellungs- und motivationsrelevante Dispositionen erfasst. Insbesondere mit den Variablen der Geschichtenmessung, von der angenommen werden kann, explizite und implizite Einstellungsinhalte zu erfassen, wird der Frage nachgegangen, inwieweit die Annahme einer Explizit-Implizit-Dimension von Einstellungen zulässig ist. Prinzipiell ermöglicht die Erhebung der verschiedenen Daten an einer großen Stichprobe die Analyse messfehlerbereinigter Variablen über latente Strukturgleichungsmodellierung. Die Anwendung latenter Datenmodellierung ist nicht nur prinzipiell von Vorteil, sondern insbesondere bei der Analyse impliziter Maße wünschenswert, für hier z.T. sehr geringe Reliabilitäten berichtet werden (vgl. Abschnitt 1.6.3).

Zur Untersuchung der Bedeutung impliziter Maße werden, neben einer Zusammenhanganalyse zu ausgewählten Prädiktoren des nomologischen Netzwerkes sozialer Einstellungen (siehe unten), individuelle und gesellschaftlich antizipierte explizite Einstellungen separat betrachtet. Karpinski & Hilton (2003) postulieren eine Art Umweltassoziationsmodell für den IAT, worin für den IAT angenommen wird, insbesondere kulturelle Assoziationen zu den entsprechenden Einstellungsobjekten zu erfassen (vgl. Olson & Fazio, 2004). Unter anderem zur Untersuchung dieser Annahme, wird der Selbstbericht derart gestaltet, dass explizite *eigene* getrennt von *gesellschaftlich antizipierten* Einstellungen erhoben werden. Über die Zusammenhangsmuster von IAT und einerseits individuellen Einstellungen sowie andererseits für die gesellschaftliche Umwelt antizipierten Einstellungen können weitere Aufschlüsse über die Bedeutung des IAT-Effekts gewonnen werden. Nosek (2004) zeigt, dass die wahrgenommene „Selbst-Gruppen-Diskrepanz“ den Zusammenhang zwischen impliziten und expliziten Einstellungen moderiert. Durch eine große wahrgenommene Abweichung zwischen der eigenen Einstellung und der innerhalb der Bezugsgruppe kann sich eine Person als einzigartig oder besonders auszeichnen. In diesem Fall, nimmt Nosek (2004) an, bestehen enge Beziehungen zwischen der Einstellung und dem Selbstkonzept, was zu einer elaborierteren Einstellung führen kann und einen bewussteren Zugang zu den Einstellungsinhalten ermöglicht. Entsprechend resultierte in der zitierten Untersuchung aus einer hohen Selbst-Gruppen-Diskrepanz eine höhere Korrelation zwischen expliziten und impliziten Einstellungsmaßen. Mit der Trennung zwischen individuellen und antizipierten Einstellungen über separate Antwortskalen kann außerdem eine Art sozialer Erwünschtheitskontrolle erfolgen. Die Einstellungen einer Person stehen häufig in Beziehung zu dem, was andere Personen bzw. die Gesellschaft über einen Gegenstand denken. Da mit dem Antwortskalendesign des Fragebogens nicht allein die eigene, sondern auch die wahrgenommene Gesellschaftsmeinung erfragt wird, wird zumindest ein bewusstes Reflektieren über Gemeinsamkeiten und Unterschiede von individuellen Einstellungen und wahrgenommenen, möglicherweise häufig auch übernommenen Einstellungen angeregt. Gegebenenfalls spiegeln die gesellschaftlich projizierten Einstellungen aber auch einen hohen individuellen Einstellungsanteil wider, da hier die Person nicht als beispielsweise Vorurteilsträger „verurteilt“ werden kann. Es ist anzunehmen, dass der Gesellschaft negativere Einstellungen zugeschrieben werden, als

der eigenen Person, da unserer Gesellschaft in vielerlei Hinsicht eine Konfrontation mit Vorurteilen und Diskriminierung für den Einzelnen bietet. Hieran ist ein weiteres Ziel vorliegender Arbeit geknüpft, und zwar die Untersuchung von Moderatoreffekten sozialer Erwünschtheit bzw. Motivation zur Vorurteilsfreiheit hinsichtlich der Beziehung zwischen den verschiedenen Einstellungsmaßen. Dabei wird eine zusammenhangsmindernde Wirkung von expliziter Motivation zu Vorurteilsfreiheit erwartet, da hochmotivierte Personen, systematisch geringere explizite Einstellungen aufzeigen, als niedrig motivierte, wogegen kein Unterschied in den impliziten Maßen in Abhängigkeit von der Motivation besteht

Über die Konstruktvalidität hinaus wird das nomologische Netzwerk expliziten und impliziten Einstellungen mittels etablierter und in Studie I und II bereits einbezogener Prädiktormäße untersucht. Die ausgewählten Prädiktoren sozialer Einstellungen sind Right-Wing-Autoritarismus (RWA, Altemeyer, 1988, 1996), Soziale Dominanzorientierung (SDO, Sidanius & Pratto, 1999), Konservatismus (Cloetta, 1983), soziale Erwünschtheit (SES, Stöber, 2001), Erwünschtheitsverhalten (BIDR, Selbst- und Fremdtäuschung, Musch, Brockhaus & Bröder, 2002), Need for Cognition (NFC, Cacioppo & Petty, 1982) und Ungewissheitstoleranz (UGTS, Dalbert, 1999). Die Prädiktoren sollten in mehr oder weniger engen Beziehungen zueinander stehen (z.B. Saucier, 2000), deren gemeinsames Substrat am ehesten als *kognitive Geschlossenheit* oder *Konservatismus* bezeichnet werden könnte. Auf einem etwas geringeren Aggregationslevel betrachtet, erfassen die Prädiktoren *Soziale Erwünschtheit* (SES, BIDR\_Selbsttäuschung, BIDR\_Fremdtäuschung), *kognitive Komplexität* (UGTS, NFC) und *Offenheit* bzw. *Konservatismus* (SDO, RWA, Konservatismus). Insofern kann die Relevanz der Dimensionen innerhalb der Prädiktoren für die Einstellungskriterien untersucht werden.

Darüber hinaus wird eine Reihe weitere relevante Variablen in Studie III einbezogen. Eine zentrale Rolle für Einstellungen zu sozialen Minderheiten spielt beispielsweise die eigene Gruppenzugehörigkeit bzw. *Selbstidentifikation* in Termini sozialer Gruppen (Selbstkategorisierungstheorie, Turner & Oakes, 1989). Für die Einstellungen zu Übergewicht bzw. Homosexualität vorliegenden Fall sind entsprechend körperliche Konstitution, sexuelle Orientierung und Identifikation mit dem eigenen Geschlecht von Bedeutung. Eine weitere, für die verschiedenen Einstellungsmessungen interessante Kontrollvariable stellt *Intelligenz* dar. Insbesondere aus Gründen der hohen Ökonomie und einer ermöglichten Erfassung des allgemeinen Intelligenzniveaus wird hierfür der Mehrfachwortwahl-Wortschatz-Intelligenztest (MWT-A, Lehrl, Merz, Burkard & Fischer, 1991) gewählt. Die Funktionsweise des MWT beruht im Wesentlichen auf Wiedererkennung bekannter Worte und deren Diskriminierung von Unbekanntem und stellt insofern insbesondere eine Wortschatzmessung dar. Hohe Korrelationen zu verschiedenen anderen, auch nicht-verbalen Intelligenztests (im Mittel von .71) weisen den MWT zum einen als „Indikator des Niveaus und Umfangs unseres verbalen Modells von Wirklichkeit“ (S.23) und zum anderen als Maß räumlich-figuraler Kompetenz aus.

Schließlich wird die individuelle Disposition einer Ekelempfindlichkeit erfasst, da diese insbesondere für die ausgewählten Einstellungsobjekte von Bedeutung sein kann. Sowohl homosexuelle, als auch übergewichtige Menschen können stärker emotionale Gefühle wie Ekel und Abscheu erregen, als andere häufig thematisierte soziale Minderheiten. Dieser Annahme kann mit Gegenüberstellung von Einstellungen und Ekelempfindlichkeit nachgegangen werden.

## 6.2 Methode

### 6.2.1 Stichprobe

Als Ausgangsstichprobe werden ausschließlich Probanden berücksichtigt, für die alle Daten des ersten und zweiten Messzeitpunkts vorliegen. Da zwei von 165 Probanden zu T2 nicht erschienen und eine weitere Person aufgrund einer anderen Muttersprache nach dem ersten Messzeitpunkt ausgeschlossen werden musste, besteht die für die Gesamtanalyse verfügbare Stichprobe aus  $N=162$  studentischen Probanden, von denen 121 Personen (74.4%) weiblich sind. Der Altersmittelwert beträgt 21.96 Jahre ( $SD=3.49$ , vgl. Abb. 6-1).

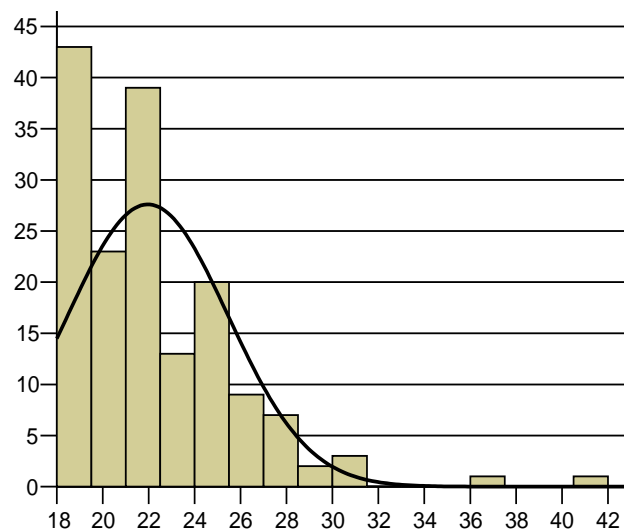
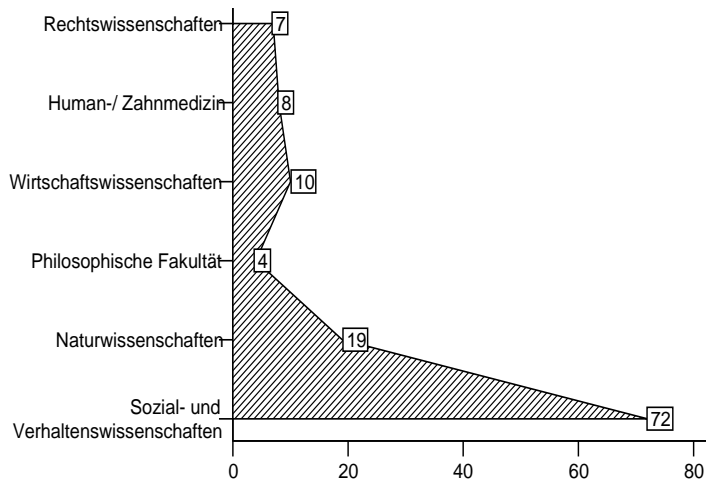


Abbildung 6-1, Altersverteilung

Mit 125 Personen (77.2%) ist die überragende Mehrheit an Untersuchungsteilnehmern in den Neuen Bundesländern aufgewachsen, wovon allein 96 Probanden aus Thüringen stammen. Drei Personen (1.9%) geben an, nicht in Deutschland aufgewachsen zu sein. Die überwiegende Anzahl Probanden ( $N=132$ , 82.0%) gibt als höchsten Bildungsabschluss die Allgemeine Hochschulreife (Abitur) an, 17 Personen (10.6%) verfügen zusätzlich über eine Berufsausbildung und 12 weitere Probanden (7.4%) über einen Hochschulabschluss. Zum Zeitpunkt der Untersuchung verfolgte der Hauptteil der Stichprobe einen Diplomstudiengang ( $N=120$ , 74.5%), wovon 68 Personen (42.2%) Psychologie studieren. Alle weiteren Studiengänge sind deutlich weniger stark vertreten, wie z.B. Rechtswissenschaften ( $N=7$ ), Humanmedizin ( $N=7$ ) und Betriebs- oder Volkswirtschaftslehre ( $N=8$ ). Etwa ein Viertel der Probanden verfolgt ein Magisterstudium ( $N=41$ , 25.5%), hier erlangt jedoch keine spezifische Fächerkombination besonderes Gewicht.





**Abbildung 6-2,** Studiengänge gemäß Fakultäten

Werden die Einzelstudiengänge nach Ausrichtung zusammengefasst, ist die Sozial- und Verhaltenswissenschaftliche Fakultät aufgrund des hohen Anteils an Psychologiestudenten mit 93 Personen (57.8%) am stärksten vertreten, gefolgt von der Naturwissenschaftlichen- (N=21, 13.0%), der Philosophischen- (N=18, 11.2%), der Wirtschaftswissenschaftlichen- (N=14, 8.7%), der Medizinischen (N=8, 5.0%) und schließlich der Rechtswissenschaftlichen Fakultät (N=7, 4.3%, Abb. 6-2).

Gemäß definierten Ausschlusskriterien für die impliziten Maße (Schwellenwerte für Reaktionszeiten und Fehlerraten) muss die Stichprobe ein weiteres Mal reduziert werden. Der Umfang der für alle Messungen validen Fälle beträgt schließlich N=156. Für die Basisanalyse der Einzelverfahren (Reliabilität, interne Zusammenhänge) wird die jeweils maximal verfügbare Probandenanzahl einbezogen.

## 6.2.2 Material

### 6.2.2.1 Implizite Messung

Zur Erfassung impliziter Einstellungen gegenüber Homosexualität und Übergewicht werden der Implizite Assoziationstest und der Extrinsic Simon Task eingesetzt. Die Entscheidung für diese beiden Verfahren resultierte aus einem Vergleich der Charakteristika ihrer Paradigmen, die größtmögliche methodische Ähnlichkeit aufweisen und dennoch spezifische Vorteile haben (siehe Kapitel 1.6).

Da der Fokus auf interindividuellen Differenzen liegt und jede Variation im Ablauf der Datenerhebung zu einer verminderten Vergleichbarkeit zwischen den Probanden führt (Banse, 2001; Teige et al., 2004), wurde sowohl die Stimulussequenz der Experimentalblöcke, als auch die Abfolge der Blöcke, die Belegung der Antworttasten und die Reihenfolge der Verfahren selbst für alle Probanden gleichsam festgelegt. Das Stimulusmaterial hatte einheitlich bildhaftes Format. Für die evaluativen Entscheidungsaufgaben (angenehm vs. unangenehm) aller vier Messungen wurden die gleichen Stimuli verwendet, dabei handelt es sich um 12 Landschaftsbilder, wovon die Hälfte eindeutig angenehme Umgebungen darstellen (z.B. Blumenwiese, Wald) und die andere Hälfte eindeutig unangenehme (z.B. Müllhalde, Hausruine). Entsprechend ihres Inhalts sollten diese Bilder mit „angenehm“ oder „unan-

„angenehm“ beantwortet werden. Die Abweichung von der in Studie I verwendeten Emotionsausdrücke und der evaluativen Entscheidung für „positiv“ oder „negativ“ hin zur Bewertung von Landschaftsbildern als „angenehm“ oder „unangenehm“ begründet sich in zweierlei Hinsicht: Um eine eindeutige Unterscheidbarkeit der Valenzstimuli von den Stimuli für die Zielkategorie „Übergewichtig“ (siehe unten) zu sichern und um eine stärker persönliche Relevanz mit den Kategorien angenehm vs. unangenehm anstelle der sehr allgemeinen Label positiv vs. negativ zu erzielen (vgl. Olson & Fazio, 2004).

Die Stimuli für die jeweiligen Klassifizierungsaufgaben von IAT und EAST sind innerhalb eines Einstellungsobjekts identisch. Das Stimulusmaterial für den Stereotyp Übergewicht besteht aus Portraits sechs übergewichtiger und sechs normalgewichtiger Frauen, für die im Vorfeld ein Matching für Alter und Attraktivität vorgenommen wurde. Für den Stereotyp Homosexualität werden die in Studie II bereits eingesetzten Bilder heterosexueller und männlicher homosexueller Liebespaare verwendet. Die exklusive Wahl von Frauen für das Einstellungsobjekt Übergewicht, sowie der von Männern für die Messung von Einstellungen gegenüber Homosexualität begründet sich ihrer Repräsentanz für das jeweilige Einstellungsobjekt. Die daraus resultierende inhaltliche Eingeschränktheit der impliziten Maße wird der unabschätzbaren Variabilität und den damit verbundenen Interpretationsschwierigkeiten, die ein Einschluss übergewichtiger Männer und homosexueller Frauen bedeuten würde, vorgezogen. Alle verwendeten Stimuli sind im Anhang aufgeführt.

Als Antworttasten fungieren das [A] und die [5] im Nummernblock. Über alle indirekten Maße hinweg sind unangenehme Stimuli mit [A] und angenehme Stimuli mit der [5] zu beantworten. Die Probanden wurden im Vorfeld der Messung dazu aufgefordert, sich während des gesamten Experimentes möglichst ruhig zu verhalten, gegebenenfalls ihr Mobiltelefon auszuschalten und wenn nötig die Brille aufzusetzen. Zu Beginn der Computereperimente wurden die Teilnehmer schließlich instruiert, einen Abstand von etwa 50 bis 60 cm zum Monitor einzunehmen und die Zeigefinder auf beide Antworttasten gelegt werden. Zwischen den einzelnen Blöcken bestand für die Teilnehmer jeweils die Möglichkeit, eine Pause einzulegen, um z.B. die Augen auszuruhen. Durch Tastendruck wurde der nächstfolgende Block gestartet. Den reibungslosen Ablauf der Experimente sicherten geschulte Versuchsleiter.

#### 6.2.2.1.1 IAT

Beide IATs sind, bis auf die Zielkonzeptstimuli und entsprechend die Instruktionen sowie die Bezeichnung der Reminder für die Tastenbelegung, völlig identisch aufgebaut. Für den Übergewichts-IAT bestand die Klassifizierungsaufgabe aus der Einordnung weiblicher Portraitbilder in eine der beiden Zielkonzeptkategorien, die mit „übergewichtig“ bzw. „normalgewichtig“ gekennzeichnet waren. Für den Homosexualitäts-IAT sollten Paarbilder in die Kategorien „homosexuell“ bzw. „heterosexuell“ eingeordnet werden. Beide IATs bestanden aus jeweils sieben Blöcken (Abb. 6-3).

Block	Aufgabe	Trials	Material	Tastenbelegung
1	Übung der evaluativen Entscheidungsaufgabe	20	evaluative Stimuli	
2	Übung der Klassifizierungsaufgabe	20	Zielkonzeptstimuli	
3	Simultane Übung beider Aufgaben aus 1 und 2	24	evaluative und Zielkonzeptstimuli	<i>kompatibel</i>
4	<b>kompatibler Experimentalblock</b>	48	Zielkonzeptstimuli	
5	Umlernen der Klassifizierungsaufgabe	20	Zielkonzeptstimuli	
6	Simultane Übung beider Aufgaben aus 1 und 5	24	evaluative und Zielkonzeptstimuli	<i>inkompatibel</i>
7	<b>inkompatibler Experimentalblock</b>	48	Zielkonzeptstimuli	

Abbildung 6-3, IAT-Aufbau

Im ersten Block werden 20 evaluative Trials präsentiert, die mit „angenehm“ oder „unangenehm“ zu beantworten sind. Der zweite Block dient zur Übung der Klassifizierungsaufgabe und besteht aus 20 Trials der Zielkonzeptkategorie. Auf die beiden Einzelblöcke folgt der erste gemischte Übungsblock mit 24 Trials (einmalige Darbietung jedes Stimulus) und vorurteils*kongruenter* Tastenbelegung ([5] unangenehm oder Übergewichtig/homosexuell vs. [A] angenehm oder normalgewichtig/heterosexuell). Nach der Übung der Aufgabe folgt der *kongruente Experimentalblock*, bestehend aus 48 Trials (zweimalige Darbietung jedes Stimulus). Im Anschluss erfolgt das Umlernen der Tastenbelegung für die Klassifikationsaufgabe mit 20 Trials (Block 5). Nun sollen Klassifizierungstrials der Kategorie „Übergewichtig“/“Homosexuell“ mit [5] und Trials der Kategorie „Normalgewichtig“/“Heterosexuell“ mit [A] beantwortet werden. Es folgt erneut ein Block mit 24 gemischten Trials zur Übung der neuen Tastenbelegung gemeinsam mit der evaluativen Entscheidungsaufgabe und nunmehr vorurteils*inkongruenter* Tastenbelegung (Block 6). Den Abschluss bildete der *inkongruente Experimentalblock* mit wiederum 48 Trials. Ein IAT dauerte in etwa 12 Minuten. Jeder IAT liefert aufgrund der zusätzlichen Verwertbarkeit der gemischten Übungsblöcke zwei reaktionszeit- und fehlerbasierte Effektmaße. Grundlage des IAT-Effekts ist die Differenz inkompatibler und kompatibler Blockmittelwerte für Latenzen bzw. Fehler.

#### 6.2.2.1.2    EAST

Abgesehen von den notwendigen Unterschieden im Stimulusmaterial und den Instruktionen sind die EASTs für beide Stereotypen ebenfalls identisch aufgebaut. Die Klassifizierungsaufgabe besteht darin, die grünlich und bläulich eingefärbten Zielkonzeptstimuli gemäß ihrer Farbe mit der „Angenehm-“ bzw. „Unangenehm-Taste“ ([5] bzw. [A]) zu beantworten. Zum Zweck einheitlicher Messbedingungen für alle Probanden (siehe oben) wurde auf eine Zwischensubjektvariation der Farbbedeutung verzichtet und somit stellt Blau für alle Versuchspersonen gleichermaßen die „Angenehme Farbe“ und Grün entsprechend die „unangenehme Farbe“ dar. Den Ergebnissen von Studie I zufolge, in denen lediglich der erste EAST-Block einen erwartungskonforme Effekt lieferte, während die restlichen 3

Messblöcke ein ausgesprochen inkonsistentes Beziehungsmuster untereinander und zu anderen Einstellungskriterien aufwiesen, werden im EAST in Studie III lediglich 2 Messblöcke á 53 Trials dargeboten. Ein EAST bestand entsprechend aus fünf Blöcken (Abb. 6-4).

Block	Aufgabe	Trials	Material
1	Übung der Valenzbelegung	20	evaluative Stimuli
2	Übung der Farbklassifizierung	24	farbige Stimuli des Zielkonzepts
3	Simultane Übung beider Aufgaben aus 1 und 2	32	evaluative und Zielkonzeptstimuli
4	<b>erster Experimentalblock</b>	5 + 48	24 evaluative Stimuli
5	<b>zweiter Experimentalblock</b>	5 + 48	12 kongruente Zielkonzeptstimuli 12 inkongruente Zielkonzeptstimuli

Abbildung 6-4, EAST-Aufbau

Der erste EAST-Block besteht jeweils aus 20 evaluativen Trials, die aufgrund ihrer Valenz mit „angenehm“ oder „unangenehm“ beantwortet werden sollen. Im zweiten Block werden ausschließlich die farbigen Stimuli des Zielkonzepts präsentiert. Innerhalb der 24 Trials erscheint jeder Stimulus einmal in bläulicher und einmal in grünlicher Farbe. Der dritte Block setzt sich aus 32 gemischten Trials zusammen und dient der gemeinsamen Übung beider Aufgaben. Auf bunte Landschaftsdarstellungen soll nach der Wertigkeit mit „angenehm“ bzw. „unangenehm“ reagiert werden, wogegen bläuliche und grünliche Personendarstellungen allein aufgrund der Farbe mit „angenehm“ bzw. „unangenehm“ zu beantworten sind. An die drei Übungsblöcke schließen sich die beiden Experimentalblöcke an. Um das für die Messung kritische Gelingen der extrinsischen Valenzbelegung der Antworttasten zu unterstützen, werden die beiden Experimentalblöcke von jeweils fünf evaluativen Trials die aufgrund ihrer Valenz beantwortet werden sollen, eingeleitet. Auf diese ersten reinen evaluativen Trials folgen 48 gemischte Trials, in denen die evaluativen Stimuli (N=12) jeweils zweimal und die Zielkonzeptstimuli (N=12) jeweils einmal in jeder Farbe dargeboten werden.

Da Blau für alle Probanden gleichermaßen die extrinsisch angenehme und Grün die extrinsisch unangenehme Farbe darstellt, sind die vorurteils*kongruenten* Trials bläuliche Bilder Normalgewichtiger bzw. Heterosexueller und grünliche Bilder Übergewichtiger bzw. Homosexueller. Die vorurteils*inkongruenten* Trials stellen entsprechend bläuliche Bilder Übergewichtiger bzw. Homosexueller und grünliche Bilder Normalgewichtiger bzw. Heterosexueller dar. Aus der Differenz inkompatibler und kompatibler Trials der beiden Experimentalblöcke resultieren jeweils zwei EAST-Werte pro Einstellungsobjekt.

### 6.2.2.2 *Explizite Messung*

#### 6.2.2.2.1 *Heimset*

Da die explizite Vorurteilmessung *nach* der Erhebung der impliziten Maße erfolgt, wurde der Selbstbericht zweigeteilt, so dass der unverfänglichere Teil der Befragung gemeinsam mit der offenen Assoziativmessung im Intervall zwischen beiden Laborsitzungen zuhause bearbeitet werden kann. Die Probanden bekommen nach dem ersten Messzeitpunkt das so bezeichnete „Heimset“ mit der Aufforderung, es zum nächsten Messzeitpunkt ausgefüllt wieder mitzubringen. Der Aufbau des Heimsets kann Tabelle 6-1 entnommen werden.

Bis auf die SES-Skala (Stöber, 2001) sind alle Items auf einer fünfstufigen Likert-Skala zu beantworten. Die Berechnung der Skalenscores erfolgt über Itemmittelung. Gemäß dem Standardeinsatz für die SES-Skala, für die eine zweistufige Ja/Nein-Antwortskala vorgesehen ist, wird hier ein Summenscore gebildet. Ebenso wird mit dem Mehrfachwortwahl-Wortschatz-Intelligenztest (MWT) und dem Becks-Depressions-Inventar (BDI) verfahren.

Zur Selbstidentifikation hinsichtlich der eigenen Figur werden Selbsteinschätzung und Wunschvorstellung (Körperschema Real und Ideal) mit einer neunstufigen bildhaften Skala nach Rand & Resnick (1999) erfasst. Die nummerierten Bildanker zeigen schematisierte weibliche und männliche Figuren, die jeweils von untergewichtigem bis hin zu übergewichtigem Körperbau rangieren. Die Geschlechtsidentität misst ein elfstufiges Differential, auf dem die Probanden angeben sollen, wie stark sie sich als Mann (linker Pol=1) bzw. als Frau (rechter Pol=11) fühlen, der Mittelpunkt (6) ist mit „weder, noch“ markiert. Die Werte für männliche Probanden werden umkodiert, so dass hohe Werte gleichermaßen eine ausgeprägte Identifizierung mit dem jeweils eigenen Geschlecht anzeigen. Die sexuelle Orientierung der Probanden wird getrennt für Männer und Frauen ebenfalls über ein elfstufiges Differential erfasst. Die linken Pole (1) sind mit „lesbisch“ bzw. „schwul“ gekennzeichnet, die rechten Pole (11) einheitlich mit „heterosexuell“, der Mittelpunkt (6) ist mit „bisexuell“ benannt.

Der MWT besteht aus 37 Wort-Nichtwort-Reihen, aus denen das im deutschen Sprachgebrauch existierende Wort herausgesucht werden soll. Die maximal erreichbare Punktzahl im MWT beträgt 37, was einem IQ von 143 entspricht. Zum Zwecke korrelativer Untersuchungen wird keine Transformation der Rohwerte in IQ-Werte vorgenommen, da dies durch Zusammenfassen der Werte am oberen und unteren Rand der Verteilung einen ungerechtfertigten Informationsverlust bedeutet. Durch die unbeaufsichtigte Erhebung im Heimset, wodurch die Möglichkeit der Täuschung besteht, könnten für den MWT Validitätseinschränkungen in aktueller Untersuchung vorliegen.

Prinzipiell werden alle Items einheitlich in Konstruktichtung kodiert, so dass hohe Skalenwerte für alle Maße jeweils eine hohe Ausprägung entsprechend des erfassten Konstruktes widerspiegeln.

Messung	Abkürzung	N Items
1. <i>Situationseinschätzung</i>		8 x 7
2. <i>Offene Assoziativmessung</i> (1.: „Frühstück“, 2.: „Eis“, 3.: „Natur“, 4.: „Schuhwechsel“)		4
3. <i>Soziale-Erwünschtheits-Skala</i> (Stöber, 2001)	SES	16
4. <i>Trait-Anxiety-Inventory</i> (Laux, Glanzmann, Schaffner & Spielberger, 1981)	STAI	20
5. <i>Fragebogen zur Erfassung der Ekelempfindlichkeit</i> (Schienle, 2001)	Ekel	37
6. <i>Becks-Depressions-Inventar</i> (Hautzinger, Bailer, Worall & Keller, 1994)	BDI	22
7. <i>Ungewissheitstoleranzskala</i> (Dalbert, 1999)	UGTS	8
8. <i>Right-Wing-Authoritarianism</i> (Altemeyer, 1996, Kämpfe 2002 <sup>2</sup> )	RWA	12
9. <i>Social Dominance Orientation</i> (Sidanius & Pratto, 1999, Kämpfe 2002 <sup>2</sup> )	SDO	16
10. <i>Need for Cognition</i> (eigene Übersetzung nach Cacioppo & Petty, 1982)	NFC	17
11. <i>Balanced Inventory of Desirable Responding (Selbst- &amp; Fremdtäuschung)</i> (Musch, Brockhaus & Bröder, 2002)	BIDR	19
12. <i>Skala Konservatismus</i> (Cloetta, 1983)	Kons	16
13. <i>Selbstidentifikationsitems</i> [Körperschemata (Real, Ideal), Geschlechtsidentität, sexuelle Orientierung]	KS_Real/Ideal Sex-ID, Sex-O	4
14. <i>Demografische Variablen</i>		6
15. <i>Mehrfachwortwahl-Wortschatz-Intelligenztest</i> (Lehrl et al., 1991)	MWT	37

**Tabelle 6-1,** Aufbau des Heimsets

Einige der Messungen, die das Heimset umfasst, sind für hier anliegende Fragestellungen nicht vordergründig relevant. Das betrifft die Situationseinschätzung, den STAI und den BDI. Sie wurden an dieser Stelle deshalb lediglich der Vollständigkeit halber aufgeführt und finden im weiteren Verlauf keine Berücksichtigung.

#### 6.2.2.2.2 Laborset

Die explizite Vorurteilmessung bildet den Abschluss der Erhebung und wird unter Aufsicht im Labor bearbeitet. Den Aufbau des Laborsets beschreibt Tabelle 6-2.

Als Kriterien expliziter Einstellungen fungieren je Einstellungsobjekt eine Skala kognitiver Einstellungen, eine Skala affektiver Einstellungen sowie eine Gefühlsthermometermessung. Um wie oben dargestellt zwischen den *individuellen* Einstellungen der Probanden und den von ihnen auf ihre Umwelt *antizipierten* Einstellungen trennen zu können, werden für die expliziten Einstellungs- und Motivationsmessungen zwei verschiedene Antwortskalen verwendet. Neben dem zuerst abgeforderten Selbstbericht der individuellen Einstellungen sollte das jeweilige Item ein zweites Mal im Sinne einer projizier-

ten gesellschaftlichen Meinung beantwortet werden. Die nebeneinander positionierten Antwortskalen werden mit „ICH“ und „UNSERE GESELLSCHAFT“ überschrieben und wie folgt instruiert:

„Es kommt häufig vor, dass man sich in seinem Standpunkt zu manchen Themen mehr oder weniger stark von der Mehrzahl der Menschen unterscheidet. Dies wollen wir mit vorliegendem Fragebogen berücksichtigen und möchten Sie daher bitten, jede Aussage in zweifacher Hinsicht zu beurteilen: Zum Einen sind wir daran interessiert, wie Sie **persönlich** darüber denken. Zum Anderen möchten wir Sie bitte, die **allgemeine gesellschaftliche Meinung** zu den präsentierten Aussagen einzuschätzen. Für jede Beurteilungsperspektive – also Ihre eigene Sicht und die der Allgemeinheit – steht Ihnen eine Antwortspalte zur Verfügung.“

Für jede Fragebogenskala des Laborsets liegen also zwei verschiedene Scores vor: Der des expliziten Selbstberichts sowie ein expliziter Projektionswert, der zur Absorption sozialer Erwünschtheit dient, zur Überprüfung der Umweltassoziationsannahme für den IAT (Karpinski & Hilton, 2001) sowie zur Untersuchung der Bedeutung von relativer Übereinstimmung bzw. Divergenz zwischen individuellen und antizipierten Einstellungen. Zur Unterscheidung der Maße dieser beiden Dimensionen werden in der Datenauswertung die selbst berichteten Werte mit *i* („ich“) ausgewiesen, während die antizipierte Gesellschaftsmeinung mit „*a*“ („antizipiert“) gekennzeichnet ist.

Messung	Abkürzung	N Items
1. <i>Einstellungen zu übergewichtigen Menschen</i> (Degner, in prep)	ÜBK <i>O</i> <i>i/a</i>	19
2. <i>Kognitive Einstellungen zur Homosexualität</i> (Seise, Banse & Neyer, 2002 und Kämpfe, 2002 <sup>2</sup> )	HOK <i>O</i> <i>i/a</i>	18
3. <i>Affektive Einstellungen gegenüber übergewichtigen Menschen</i> (Autorin)	ÜBAF <i>i/a</i>	10
4. <i>Affektive Einstellungen zur Homosexualität</i> (Seise, Banse & Neyer, 2002)	HOAF <i>i/a</i>	9
5. <i>Gefühlsthermometer</i> (über-, normal-, untergewichtige-, homo- und heterosexuelle Menschen; z.B. Miller & Miller, 1977)	Th_Ü/N/U Th_Ho/He	5
6. <i>Modern Racism</i> (McConahay, 1986; überarbeitete und übersetzte Fassung)	MRS <i>i/a</i>	8
7. <i>Internal/ External Motivation to Respond Without Prejudice</i> (Plant & Devine, 1998, übersetzte und psychometrisch getestete Fassung)	IMS <i>i/a</i> EMS <i>i/a</i>	12
8. <i>Motivation zu vorurteilsfreiem Verhalten</i> (Banse & Gawronski, 2003)	MVV <i>i/a</i>	16
9. <i>Vorurteilstrollskala</i> (Gschwendner-Lukas, Hofmann & Schmitt, 2004)	VKS <i>i/a</i>	8
10. <i>Stereotypen-Schätzfragen</i> (Vorschaufehler)		20
11. <i>Anteilsschätzfragen zu relevanten Bevölkerungsgruppen in Deutschlands</i>		12
12. <i>Demografie</i>		6
13. <i>Kontrollitems für die Computerexperimente</i>		3
14. <i>Sexuelle Orientierung, übergewichtige-, homosexuelle Freunde, Körpergröße, -gewicht</i>	Sex-O, FreundÜ/H	5

**Tabelle 6-2**, Aufbau des Laborsets

### 6.2.2.3 Offene Assoziativmessung

Als Bildmotive für die offene Assoziativmessung werden die Bilder „Eis“ und „Schuhwechsel“ für Übergewicht sowie „Frühstück“ und „Natur“ verwendet. Analog zum Prozedere in Studie II werden die Versuchsteilnehmer gebeten, zu jedem Bild eine eigene Geschichte zu verfassen, die berücksichtigen soll, was die Illustration darstellt, was die dargestellten Personen denken und fühlen, wie es zu der abgebildeten Szene kam und wie sich die Situation weiter entwickeln könnte. Die offene Assoziativmessung ist eingebettet in das Heimset, wobei sich die Gestaltung an den in Studie II verwendeten Testheften orientiert.

Die vier pro Versuchsteilnehmer vorliegenden Geschichten werden von drei unabhängigen Ratern nach einem Schema kodiert, das ausgehend von dem allgemeingültigen Teil der Kodierschemata in Studie II um einige Variablen erweitert bzw. reduziert (siehe Tabelle 6-3 in Anhang A.6.2). Beibehalten wird die Kodierung der *Geschichtenlänge*, der *Valenz* und der *Stereotypenerkennung*. Zusätzlich berücksichtigt das Kodierschema den *Vorurteilsgehalt* und die *Kreativität* der Geschichten. Entsprechend der Unterscheidung zwischen individuellen und antizipierten Einstellungen im Selbstbericht, wird für den inhaltsorientiert kodierten Vorurteilsgehalt der Geschichten eine entsprechende Differenzierung zwischen eigenen Vorurteilen des Autors und Vorurteilen, die der Autor ausdrücklich der Gesellschaft zuschreibt, vorgenommen. Die getrennte Kodierung der eigenen Haltung des Autors und einer Rezipitation der gesellschaftlichen Meinungen seitens des Autors dient hier dazu, eine Vermischung der verschiedenen Perspektiven, aus denen eine Geschichte prinzipiell geschrieben werden kann, zu verhindern, um eine isolierte Erfassung der individuellen Einstellungen zu ermöglichen. Die Kreativität wird als interessante Kontrollvariable mit bedacht.

Im Gegensatz zu Studie II werden die *Stimmung*, die *Ernsthaftigkeit* und die *bildspezifischen Inhalte* hier nicht kodiert. Die Stimmung ist, wie sich bei der Kodierung der Geschichten in Studie II herausstellte, nur schwer von der Valenz zu unterscheiden. Die Ernsthaftigkeit erwies sich einerseits als wenig aussagekräftig und darüber hinaus zeigen die geringen Interrater-Reliabilitäten aus Studie II an, dass diese Variable nicht sehr treffsicher kodiert werden konnte. Die bildspezifisch variierenden Inhaltsvariablen fungierten bereits in Studie II als Kontrollvariablen für die Auswahl der Bilder, da die offene Assoziativmessung in Studie III als Einstellungsmessung eingesetzt wird, ist eine Vergleichbarkeit der pro Bild kodierten Variablen notwendig.



### 6.2.3 Durchführung

Die Datenerhebung ist dreigeteilt, sie erfolgt in zwei Labormessungen im Intervall von exakt sieben Tagen und Messungen, die von den Probanden in der Zwischenzeit zuhause bearbeitet werden. Von einer Erhebung aller vier impliziten Messungen (IAT- und EAST-Übergewicht und -Homosexualität), des Selbstberichtes und der offenen Assoziativmessung (4 Geschichten) in einer Sitzung wurde aus verschiedenen Gründen abgesehen. Einerseits sollten unerwünschte messmethodische Effekte ausgeschlossen werden, wie z.B. Lerneffekte oder Ermüdungserscheinungen. Andererseits bot der prozedurale Zusammenschluss der Datenerhebungen für hier verfolgte Fragestellungen und für eine weitere, inhaltlich unverwandte Untersuchung die Möglichkeit, finanzielle und prozedurale Synergien zu nutzen. Die Laborsitzungen finden gemäß der Computerausstattung jeweils in Gruppentestungen zu maximal 5 Teilnehmern statt. In der Zeit zwischen den Labormessungen fertigen die Probanden jeweils eine Geschichte zu den vier Bildmotiven der offenen Assoziativmessung an und bearbeiten einen umfassenden Selbstbericht („Heimset“), der jedoch noch nicht die eigentliche Vorurteilsmessung enthält. Das ausgefüllte Testmaterial wird zur zweiten Sitzung wieder eingesammelt. Die explizite Einstellungsmessung erfolgt als letzte Messung zum zweiten Labormesszeitpunkt („Laborset“). Nach Abschluss der zweiten Labormessung und bei Vorliegen aller Messdaten erfolgt die Entlohnung der Probanden mit 4 Versuchspersonenstunden oder € 25,-. Einen Überblick über das Untersuchungsdesign bietet Tabelle 6-4.

<b>T1</b>		<b>7 Tage Intervall</b> ←————→	<b>T2</b>		
<b>Ablauf Messungen</b>	<b>Dauer (min)</b>		<b>Ablauf Messungen</b>	<b>Dauer (min)</b>	
1.	<b>IAT „Über“</b>	<b>„Heimset“</b> Einstellungsprädiktoren offene Assoziativmessung	1.	Free Recall zum Ekel-Stroop	1
2.	STAI-State		2.	<b>IAT „Homo“</b>	10
3.	Rückschaufehler 1		3.	Rückschaufehler 2	10
4.	<b>EAST „Über“</b>		4.	Glaubwürdigkeit	3
5.	Ekel-Induktion		5.	STAI-State Ekel-State	3
6.	Ekel-State STAI-State		6.	<b>EAST „Homo“</b>	12
7.	Ekel-Stroop		7.	<b>„Laborset“</b> explizite Kriterien	35
<b>ca. 60</b>			<b>ca. 75</b>		

**Tabelle 6-4**, Überblick Untersuchungsdesign, fett gedruckt sind die für vorliegende Arbeit relevanten Messungen

#### 6.2.4 Auswertung

Die Dokumentation der Ergebnisse von Studie III ist in sechs Kapitel unterteilt. Im ersten Ergebniskapitel (6.3) werden die Basisanalysen der einzelnen Messungen berichtet, speziell Reliabilitäten, Deskriptive Statistiken sowie interne Zusammenhangsanalysen. Die fünf restliche Ergebniskapitel (6.4 bis 6.8) sind den Untersuchungen der fünf zugrunde liegenden Fragestellungen von Studie III gewidmet.

Für die impliziten Messverfahren werden reaktionszeit- und fehlerbasierte Maße betrachtet. Als Ausschlusskriterien für Ausreißer gelten gemäß den Konventionen ein individuelles Reaktionszeitmittel von mehr als 2000 ms und eine individuelle Fehlerrate von mehr als 3 Standardabweichungen über dem Gesamtmittelwert in der Stichprobe (z.B. Greenwald, Banaji & Schwartz, 1998; De Houwer, 2003a). Der Ausschluss von Probanden wird ggf. berichtet.

Zur Analyse der internen Konsistenz des IAT-Effekts werden die Split-Half-Reliabilitäten des Differenzmaßes basierend auf Übungs- und Experimentalblöcken, sowie auf erster und zweiter Hälfte (à 24 Trials) der Experimentalblöcke berichtet. Für die Reliabilitätsanalyse des EAST-Effekts werden in Anlehnung an De Houwer (2003) *Within-* und *Cross-Concept-Korrelationen*, sowie Split-Half-Reliabilitäten des Differenzmaßes basierend auf den beiden Experimentalblöcken dokumentiert. *Within-concept Korrelationen* bezeichnen die Zusammenhänge der nach Zielkonzeptstimuli getrennten Maßen zwischen erstem und zweitem Übungsblock. Hierfür werden die Korrelationen der für Stimuli der Zielkategorie (Übergewichts-, bzw. Homosexualität) und der jeweiligen Kontrastkategorie (Normalgewichts-, bzw. Heterosexualität) getrennt berechneten Effekte dokumentiert. *Cross-Concept Korrelationen* stellen die Zusammenhänge zwischen Ziel- und Kontrastkategorie jeweils innerhalb eines Blocks dar (vgl. De Houwer, 2003). Reliabilitäten werden sowohl für reaktionszeit-, als auch fehlerbasierte Effektmaße dokumentiert.

Für die deskriptive Auswertung der impliziten Maße werden zugunsten der Anschaulichkeit auch die auf den Rohwerten (in ms) beruhenden Effekte berichtet. Die Ergebnisdokumentation erfolgt für die jeweiligen Gesamtmaße sowie die stimuluspezifischen Effekte getrennt für Ziel- und Kontrastkategoriestimuli. Die Berechnung der stimuluspezifischen Effekte wird wie folgt vorgenommen: Für den Zielkategorie-Effekt des IAT werden die Reaktionszeiten bzw. Fehler der jeweils 12 Zielkategorientrials (Übergewicht bzw. Homosexualität) aus inkongruentem und kongruentem Block gemittelt und gemäß der Berechnung des IAT-Effekts subtrahiert. Analog erfolgt die Berechnung des IAT-Effekts für die Kontrastkategorie (Normalgewicht bzw. Heterosexualität). Der allgemeine IAT-Effekt beruht auf den Blockmittelwerten über jeweils *allen* 48 Trials und stellt das konventionelle IAT-Maß dar.

Für den Zielkategorie-Effekt des EAST werden die Reaktionszeiten bzw. Fehler der *relevanten* inkongruenten und kongruenten Übergewichts- bzw. Homosexualitätstrials getrennt je Block gemittelt und

subtrahiert. Entsprechend wird für den Kontrastkategorie-Effekt verfahren. Zur Berechnung des Gesamteffekts je EAST-Block gehen, im Gegensatz zum IAT-Maß, wofür die Blockmittelwerte *aller* präsentierten Trials verrechnet werden, ausschließlich die 24 relevanten Ziel- und Kontrastkategorie-Trials je Block ein (vgl. De Houwer, 2003a). Das bedeutet, IAT- und EAST-Effekt umfassen eine unterschiedliche Anzahl, wie auch unterschiedliche Arten von Stimuli. Während für den konventionellen IAT-Effekt alle Trials des kongruenten und inkongruenten Blocks verwendet werden, berechnet sich der EAST-Effekt ausschließlich aus den jeweiligen Zielkonzepttrials. Da keine Zwischensubjektvariationen im Prozedere vorgenommen wurden, reflektieren die gemäß IAT und EAST resultierenden Haupteffekte neben der erwünschten Konstruktvarianz, möglicherweise auch prozedurale Effekte (z.B. Blockreihenfolge). Daher muss von einer uneingeschränkten Interpretation der Stichprobenmittelwerte abgesehen werden. Nichtsdestotrotz werden für die beiden EASTs standardmäßig ANOVAs mit Messwiederholung berechnet, um Haupteffekte und Interaktionen auf Gruppenebene zu überprüfen (vgl. De Houwer, 2003a).

Analog zu Studie II wird mit den resultierenden Daten von IAT und EAST gleichermaßen verfahren. Im Unterschied zu Studie II werden die in mittlerweile in einigen Studien (z.B. De Houwer, Crombez, Koster & De Beul, 2004; Teige et al., 2004) etablierten Reaktionszeitschranken von 300 und 3000 ms nach De Houwer (2003a) übernommen. Gemäß der Prüfung unterschiedlicher Reaktionszeitschranken für Studie II wurden lediglich geringfügige Unterschiede in der Auswirkung beobachtet. Die resultierende Datenbehandlung umfasst:

- Eliminierung der Reaktionszeiten von Fehlertrials
- Rekodierung extremer Reaktionszeiten auf die Schranken 300 und 3000 ms
- Logarithmierung der Reaktionszeiten
- Ausschluss von Personen mit einer Fehlerquote, die mehr als 3 Standardabweichungen vom Gruppenmittelwert abweicht

## 6.3 Ergebnisse I: Basisanalysen

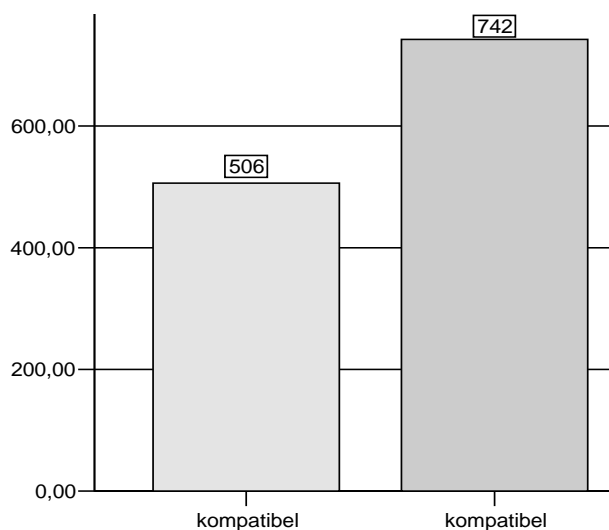
### 6.3.1 Implizite Messungen

#### 6.3.1.1 IAT Übergewicht

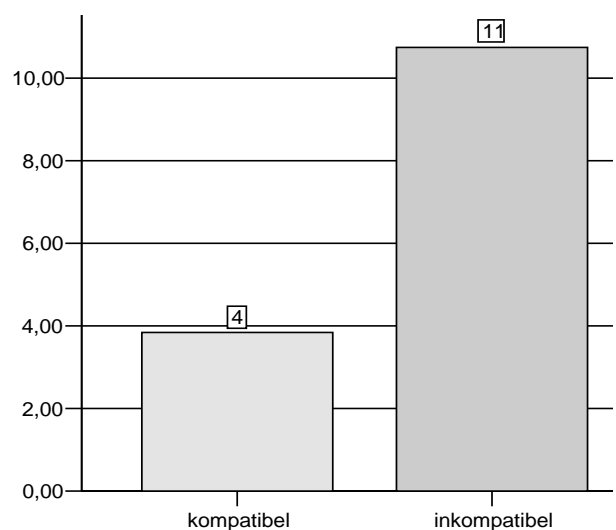
Fünf Versuchspersonen wurde aufgrund ihrer hohen Fehlerrate (22.22, 23.61, 25.00, 25.69 und 30.56%) von den Analysen ausgeschlossen. Nach Ausschluss dieser Fälle (die resultierende Stichprobe umfasst N=160 Personen) beträgt die Fehlerquote für alle Trials der Blöcke 3, 4, 6 und 7 im Mittel 6.61% (SD= 3.76). Die Differenz zwischen inkongruentem und kongruentem Block resultiert sowohl für Latenzen, als auch Fehlerraten in einem signifikanten Haupteffekt (vgl. Tab. 6-5; Abb. 6-5, 6-6). Der mittlere IAT-Effekt der Stichprobe beträgt bei Betrachtung der untransformierten Latenzen 236.32 ms (SD= 169.87) und für die Fehlerwerte 6.90% (SD= 7.13).

Block	logarithmierte Latenzen in ms			Fehlerraten in %		
	M (SD)	T	p	M (SD)	T	p
3	6.21 (.11)			4.56 (5.17)		
6	6.62 (.09)			14.09 (11.09)		
IAT-Effekt	.41 (.24)	21.30	.000	9.51 (11.63)	10.34	.000
4	6.20 (.09)			3.10 (3.18)		
7	6.40 (.24)			7.40 (7.40)		
IAT-Effekt	.20 (.21)	12.78	.000	4.30 (3.47)	10.31	.000

**Tabelle 6-5,** IAT Übergewicht, Haupteffekte der Reaktionszeiten und Fehlerraten für Übungs- (3, 6) und Experimentalblöcke (4, 7), Test der gepaarten Mittelwertsdifferenzen gegen Null, df= 159



**Abbildung 6-5,** IAT Übergewicht, Reaktionszeiten (Rohwerte in ms) für Mittelwert von Block 3 und 4 (kompatibel) sowie Block 6 und 7 (inkompatibel)



**Abbildung 6-6,** IAT Übergewicht, Fehlerraten (in %) Mittelwert von Block 3 und 4 (kompatibel) sowie Block 6 und 7 (inkompatibel)

Die IAT-Effekte aus Übungs- und Experimentalblock korrelieren signifikant positiv, und zwar zu .43 für Reaktionszeiten und .33 für Fehlerraten ( $p < .001$ , Tab. 6-6). Die Split-Half-Reliabilität basierend auf Übungs- und Experimentalblöcken beträgt entsprechend .60 für die Reaktionszeiten und .50 für die Fehler. Die allein auf den Experimentalblöcken basierenden IAT-Effekte korrelieren jeweils zu .18, woraus eine Split-Half-Reliabilität von .30 sowohl für Reaktionszeiten, als auch für Fehler resultiert. Reaktionszeit- und Fehler-IAT korrelieren zu .02 ( $p > .05$ ) in den Übungsblöcken, zu .22 ( $p < .01$ ) in den Experimentalblöcken und zu .27 ( $p < .001$ ) für den Gesamteffekt aus Übungs- und Experimentalblöcken.

Die Mittelwerte der einzelnen IAT-Effekte sind ohne Ausnahme signifikant verschieden von Null und reflektieren einen Reaktionsvorteil der vorurteilskongruenten Stimuluskombinationen (Übergewichtig x unangenehm beziehungsweise Normalgewichtig x angenehm) vor der inkongruenten Kombination. Da die Blockmittelwerte alle, also auch die evaluativen Trials umfassen, sind die Mittelwerte der stimuluspezifischen IAT-Effekte nicht identisch mit den dokumentierten Blockmittelwerten. Die Trennung der Effekte für beide Zielkonzeptkategorien dient lediglich der Veranschaulichung der Gewichtung und somit der besseren Interpretierbarkeit des Gesamteffekts des IAT. So können konsistent höhere Reaktionszeit- und Fehlereffekte für die Kontrastkategorie „Normalgewichtig“ im Vergleich zur Zielkategorie „Übergewichtig“ beobachtet werden. Für Stimuli, die als „normalgewichtig“ zu kategorisieren sind, bestehen also deutlichere Unterschiede zwischen kongruenter und inkongruenter Tastenbelegung als für Stimuli, die in die Kategorie „übergewichtig“ einzuordnen sind.

<b>IAT Übergewicht, N=160</b>		<b>Latenzen</b>	<b>Fehlerraten</b>
<b>Reliabilität</b>		<b><math>\alpha</math> (r)</b>	<b><math>\alpha</math> (r)</b>
Übungs- und Experimentalblöcke		.60 (.43)	.50 (.33)
Erste und zweite Hälfte der Experimentalblöcke		.30 (.18)	.30 (.18)
<b>Deskriptive Statistik (ms)</b>		<b>M (SD)</b>	<b>M (SD)</b>
Ü-Block	Übergewichtig	309.82 (263.74)	8.59 (20.42)
	Normalgewichtig	359.32 (321.93)	12.32 (22.94)
	<i>Blockmittel</i>	<i>326.24 (227.38)</i>	<i>9.51 (11.63)</i>
E-Block	Übergewichtig	124.28 (173.90)	4.45 (10.50)
	Normalgewichtig	199.09 (228.11)	8.59 (12.16)
	<i>Blockmittel</i>	<i>146.40 (176.18)</i>	<i>4.30 (5.27)</i>
Gesamt	Übergewichtig	219.15 (190.37)	6.52 (12.88)
	Normalgewichtig	359.32 (321.93)	10.46 (14.40)
	<i>Blockmittel</i>	<i>236.32 (169.87)</i>	<i>6.90 (7.13)</i>

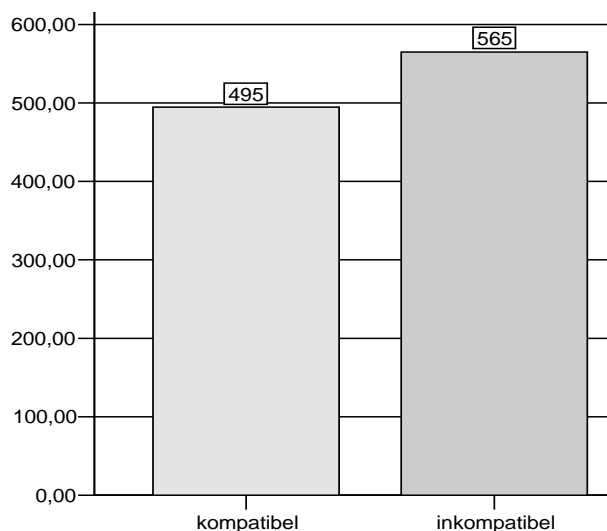
**Tabelle 6-6**, Reliabilitäten und stimuluspezifische Haupteffekte für IAT Übergewicht, Ü-Block= Übungsblöcke (3 und 6), E-Block= Experimentalblöcke (4 und 7), „Gesamt“: Mittelung von Übungs- und Experimentalblöcken

## 6.3.1.2 IAT Homosexualität

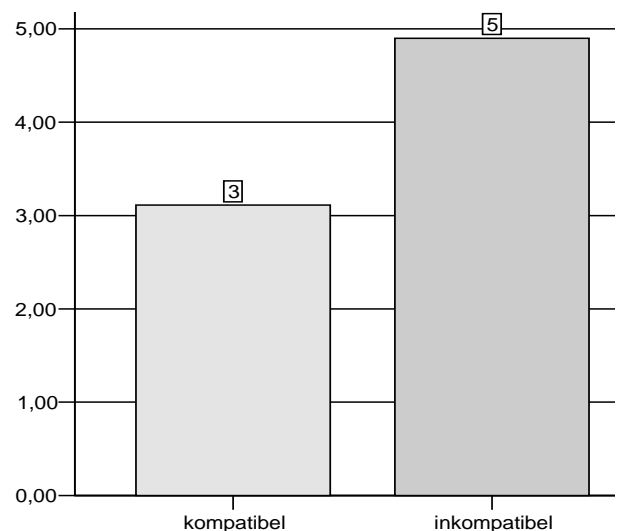
Da für drei Untersuchungsteilnehmer keine Daten der zweiten Messung vorliegen, und vier weitere Versuchspersonen aufgrund hoher Fehlerquoten eliminiert werden müssen, umfasst die Ausgangsstichprobe des IAT Homosexualität  $N=158$ . Die Fehlerquote für alle Trials der Blöcke 3, 4, 6 und 7 beträgt im Mittel 3.84 (SD= 2.56) und ist somit signifikant geringer als die des IAT Übergewicht (Mittel der Differenz= 2.77,  $T= 9.51$ ,  $df= 154$ ,  $p<.000$ ). Dieses Ergebnis könnte sowohl Lern-, als auch Materialeffekte widerspiegeln. Die Haupteffekte des IAT Homosexualität sind zwar nach wie vor in hypothesenkonformer Richtung statistisch bedeutsam (vgl. Tab. 6-7; Abb. 6-7, 6-8), jedoch auch ohne Ausnahme signifikant kleiner, als die IAT-Effekte für Übergewicht (vgl. Tab. 6-15). Der IAT-Effekt beträgt für die Reaktionszeitdaten 70.32 ms (SD= 76.86) und 1.79 % (SD= 3.61) für die Fehlerraten. zwar sind die IAT-Effekte nach wie vor statistisch bedeutsam, jedoch liegen sie betragsmäßig deutlich unter denen des IAT Übergewicht.

Block	logarithmierte Latenzen in ms				Fehlerraten in %		
	M (SD)	T	p	M (SD)	T	p	
3	6.18 (.10)			3.30 (3.88)			
6	6.33 (.18)			5.72 (5.03)			
IAT-Effekt	.15 (.14)	12.99	.000	2.43 (5.76)	5.29	.000	
4	6.18 (.08)			2.93 (3.01)			
7	6.25 (.12)			4.07 (3.54)			
IAT-Effekt	.07 (.09)	9.49	.000	1.15 (3.42)	4.22	.000	

**Tabelle 6-7**, IAT Homosexualität, Haupteffekte der Reaktionszeiten und Fehlerraten für Übungs- (3, 6) und Experimentaltalblöcke (4, 7), Test der gepaarten Mittelwertsdifferenzen gegen Null,  $df=157$



**Abbildung 6-7**, IAT Homosexualität, Reaktionszeiten (Rohwerte in ms) für Mittelwert von Block 3 und 4 (kompatibel) sowie Block 6 und 7 (inkompatibel)



**Abbildung 6-8**, IAT Homosexualität, Fehlerraten (in %) für Mittelwert von Block 3 und 4 (kompatibel) sowie Block 6 und 7 (inkompatibel)

Die Korrelationen der IAT-Effekte aus Übungs- und Experimentalblock sind wie im IAT Übergewicht signifikant positiv. Mit einer Korrelation von .61 ( $p < .001$ ) und der daraus resultierenden Split-Half-Reliabilität von .76 ist der Zusammenhang zwischen Übungs- und Experimentalblockeffekt für Reaktionszeiten jedoch deutlich höher (Tab. 6-8). Die Fehlerraten stehen dagegen in geringerem Zusammenhang als im IAT Übergewicht. Die Korrelation zwischen Übungs- und Experimentalblock beträgt .18 ( $p > .05$ ), die Split-Half-Reliabilität entsprechend .31. Die auf erster und zweiter Hälfte der Experimentalblöcke basierenden IAT-Effekte korrelieren mit .58 ebenfalls hoch für die Reaktionszeiten. Die Split-Half-Reliabilität beträgt .74. Die Korrelation der Fehler-IATs aus den Experimentalblockdaten ist dagegen nicht signifikant ( $r = .03$ ,  $p > .05$ ) und resultiert in einer Reliabilität von lediglich .05. Reaktionszeit- und Fehler-IAT korrelieren im IAT Homosexualität zu .35 ( $p < .001$ ) in den Übungsblöcken, zu .16 ( $p < .05$ ) in den Experimentalblöcken und zu .34 ( $p < .001$ ) für den Gesamteffekt aus Übungs- und Experimentalblöcken.

Innerhalb der Reaktionszeiten werden wiederum höhere Effekte für Trials der Kontrastkategorie „Heterosexuell“ beobachtet, für die Fehlereffekte sind diese Unterschiede dagegen nicht durchgängig.

<b>IAT Übergewicht, N=160</b>		<b>Latenzen</b>	<b>Fehlerraten</b>
<b>Reliabilität</b>		<b><math>\alpha</math> (r)</b>	<b><math>\alpha</math> (r)</b>
Übungs- und Experimentalblöcke		.76 (.61)	.31 (.18)
Erste und zweite Hälfte der Experimentalblöcke		.74 (.58)	.05 (.03)
<b>Deskriptive Statistik (ms)</b>		<b>M (SD)</b>	<b>M (SD)</b>
Ü-Block	Übergewichtig	97.10 (152.02)	3.90 (12.07)
	Normalgewichtig	117.38 (152.12)	3.06 (16.71)
	<i>Blockmittel</i>	<i>97.64 (105.82)</i>	<i>2.42 (5.76)</i>
E-Block	Übergewichtig	47.32 (87.57)	1.85 (7.96)
	Normalgewichtig	76.27 (94.52)	2.85 (9.58)
	<i>Blockmittel</i>	<i>43.00 (62.53)</i>	<i>1.15 (3.42)</i>
Gesamt	Übergewichtig	72.21 (99.47)	2.88 (7.61)
	Normalgewichtig	96.83 (110.24)	2.95 (10.30)
	<i>Blockmittel</i>	<i>70.32 (76.86)</i>	<i>1.79 (3.61)</i>

**Tabelle 6-8**, Reliabilitäten und stimuluspezifische Haupteffekte für IAT Homosexualität, Ü-Block= Übungsblöcke (3 und 6), E-Block= Experimentalblöcke (4 und 7), „Gesamt“: Mittelung von Übungs- und Experimentalblöcken

Gemäß der Basisanalyse werden für alle weiteren Analysen IAT-Effekte für Übergewicht und Homosexualität basierend auf Übungs- und Experimentalblöcken verwendet.

### 6.3.1.3 EAST Übergewicht

Aufgrund extremer Fehlerraten (20.92, 20.92, 21.01, 26.87 und 27.54%) werden fünf Probanden von den Analysen ausgeschlossen. Die Gesamtfehlerquote für die verbleibende Stichprobe von  $N=159$  beträgt demnach 5.48 ( $SD=3.34$ ).

Die Haupteffekte des EAST Übergewicht variieren deutlich zwischen den Blöcken (Tab. 6-9). Während die Differenz inkompatibler und kompatibler Trials des ersten Blocks in sowohl für Reaktionszeit-, als auch die Fehlerdaten negativen EAST-Effekten resultiert, wobei lediglich der Fehlereffekt signifikant verschieden von Null ist, bestehen für die Trials des zweiten Blocks signifikant positive EAST-Effekte. Da gemäß der Berechnung des EAST-Effekts positive Werte einen Reaktionsvorteil der Kombination „Übergewichtig“ und „unangenehm“ sowie „Normalgewichtig“ und „angenehm“ indizieren, besteht ein vorurteilskongruenter Effekt in der Gesamtstichprobe lediglich für die Trials des zweiten Blocks. Die durch Mittelung beider Blöcke erhaltenen Gesamteffekte sind weder für Reaktionszeiten, als für Fehler statistisch bedeutsam.

Block	logarithmierte Latenzen in ms				Fehlerraten in %		
	M (SD)	T	<i>p</i>	M (SD)	T	<i>p</i>	
1 kompatibel	6.33 (.15)			12.92 (11.72)			
1 inkompatibel	6.31 (.13)			6.48 (8.41)			
EAST 1-Effekt	-.01 (.10)	-1.57	.118	-6.44 (12.65)	-6.42	.000	
2 kompatibel	6.31 (.14)			7.50 (8.81)			
2 inkompatibel	6.34 (.16)			11.85 (10.50)			
EAST 2-Effekt	.03 (.10)	3.80	.000	4.35 (13.45)	4.08	.000	
EAST Gesamteffekt	.01 (.08)	1.33	.184	-1.04 (9.49)	-1.39	.167	

**Tabelle 6-9,** EAST Übergewicht, Haupteffekte der Reaktionszeiten und Fehlerraten für Block 1, 2 sowie den Gesamteffekt, Test der gepaarten Mittelwertsdifferenzen gegen Null,  $df=158$

Die Split-Half-Reliabilität basierend auf den Latenzen des ersten und zweiten Blocks beträgt .36. Die Korrelationen der nach Einstellungsobjekt getrennten EAST-Werten über beide Blöcke sind signifikant positiv, jedoch sehr gering. Die Korrelationen zwischen den Einstellungsobjekten sind ebenfalls gering und durchweg negativ. EAST-Werte basierend auf Fehlerraten zeigen kaum interne Zusammenhänge. Diese Werte replizieren die Befunde von De Houwer (2003a).

Die Korrelation zwischen Latenzen- und Fehler-Effekten beträgt .05 (n.s.) für Block 1, .16 ( $p<.05$ ) für Block 2 und .29 ( $p<.01$ ) für den Gesamtscore.



<b>EAST Übergewicht, N=159</b>		<b>Latenzen</b>	<b>Fehlerraten</b>
<b>Korrelationen und Reliabilität der EAST-Scores</b>			
Within-concept Korrelation	EAST_Über	.27 **	.04
Block 1 x Block 2	EAST_Normal	.26 **	.22 **
-----			
Cross-concept Korrelation	Block 1	-.17 *	.06
EAST_Über x EAST_Normal	Block 2	-.22 **	.11
-----			
Split-Half-Reliabilität Block 1 x 2	EAST_Über	.41	.08
	EAST_Normal	.43	.36
	gesamt	.36	.11
<b>Deskriptive Statistik (ms)</b>			
		<b>M (SD)</b>	<b>M (SD)</b>
Block 1	Übergewichtig	-59.01 (97.59)	-5.03 (15.77)
	Normalgewichtig	40.14 (142.33)	-8.34 (18.89)
	<i>gesamt</i>	<i>-11.43 (83.68)</i>	<i>-6.44 (12.65)</i>
-----			
Block 2	Übergewichtig	13.94 (104.68)	7.55 (17.96)
	Normalgewichtig	26.30 (115.09)	1.15 (18.19)
	<i>gesamt</i>	<i>19.00 (68.52)</i>	<i>4.35 (13.45)</i>
-----			
Gesamt	Übergewichtig	-31.30 (80.42)	.16 (11.15)
	Normalgewichtig	29.15 (94.48)	.59 (13.99)
	<i>gesamt</i>	<i>3.77 (57.29)</i>	<i>-1.04 (9.49)</i>

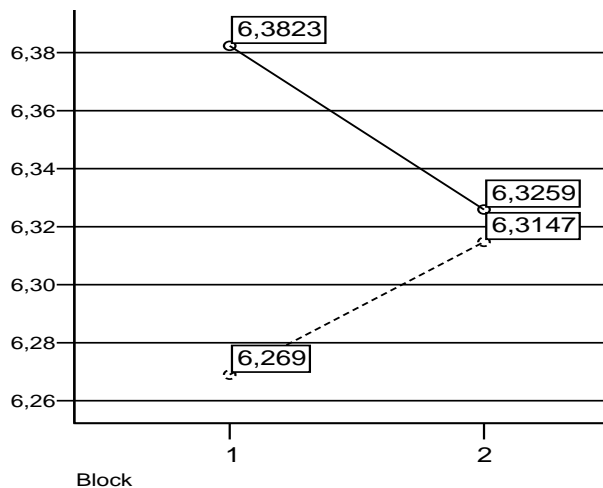
**Tabelle 6-10**, Reliabilitäten und stimuluspezifische Haupteffekte für EAST Übergewicht

Wie aus der unteren Hälfte von Tabelle 6-10 deutlich ersichtlich wird, bestehen große Unterschiede zwischen den EAST-Effekten in Abhängigkeit von der Stimuluskategorie bzw. -art und den Experimentalblöcken. Um diese Effekte zu untersuchen, wird eine für Latenzen und Fehlerwerte je eine 2 (Block) x 2 (Stimulusart: Übergewichtig vs. Normalgewichtig) x 2 (extrinsische Valenz: angenehm vs. unangenehm) ANOVA mit Messwiederholung berechnet (Deskriptive Statistiken siehe Tab. 6-11).

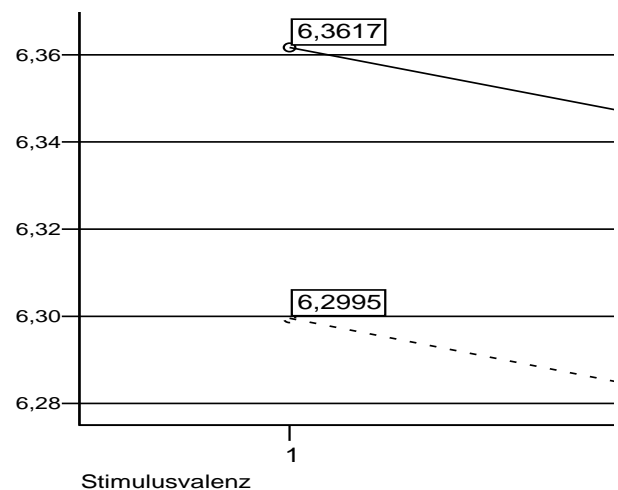
<b>Block</b>	<b>Stimulus - Response</b>	<b>Rohlatenzen</b>	<b>log. Latenzen</b>	<b>Fehlerraten in %</b>	
		<b>M (SD)</b>	<b>M (SD)</b>	<b>M (SD)</b>	
1	Über- gewichtig <sup>x</sup>	angenehm	532.62 (82.94)	6.27 (.15)	5.87 (9.77)
		unangenehm	620.88 (133.22)	6.41 (.20)	13.05 (17,53)
	Normal- gewichtig <sup>x</sup>	angenehm	538.04 (105.67)	7.25 (.18)	15.35 (16.41)
		unangenehm	581.136 (97.34)	6.35 (.16)	7.01 (11.46)
2	Über- gewichtig <sup>x</sup>	angenehm	572.78 (111.88)	6.33 (.18)	13.42 (13.97)
		unangenehm	558.05 (91.15)	6.31 (.15)	5.87 (10.12)
	Normal- gewichtig <sup>x</sup>	angenehm	549.63 (93.89)	6.30 (.16)	9.12 (12.81)
		unangenehm	575.08 (113.43)	6.34 (.18)	10.27 (14.01)

**Tabelle 6-11**, Deskriptive Statistik für ANOVA mit Messwiederholung, EAST Übergewicht

Die Analyse der Reaktionszeiten erbringt einen Haupteffekt der *Stimulusart* ( $F(1, 156)=7.10$   $p<.01$ ), resultierend aus kürzeren Reaktionszeiten für normalgewichtige- ( $M\cong 553$  ms) im Vergleich zu übergewichtigen Stimuli ( $M\cong 562$  ms). Der Haupteffekt *extrinsischer Valenz* ist ebenfalls signifikant ( $F(1, 156)=61.75$   $p<.001$ ) und zurückzuführen auf kürzere Reaktionszeiten bei extrinsisch positiver- ( $M\cong 540$  ms) im Vergleich zu extrinsisch negativer Valenz ( $M\cong 575$  ms, vgl. Tab. 6). Die längeren Latenzen für negative Stimuli deuten auf Stroop-Effekte hin, wie sie bereits De Houwer (2003a) beschreibt. Neben den Haupteffekten findet sich eine signifikante Interaktion zwischen *Block* und *extrinsischer Valenz* aufgrund deutlich längerer Reaktionszeiten bei extrinsisch negativer Valenz (Grün) im ersten Block ( $F(1, 156)=73.22$   $p<.001$ ; Abb. 6-9). Die kritische Interaktion zwischen Stimulusart und extrinsischer Valenz erreicht keine statistische Bedeutsamkeit ( $F(1, 156)=.00$   $p=.988$ , Abb. 6-10).

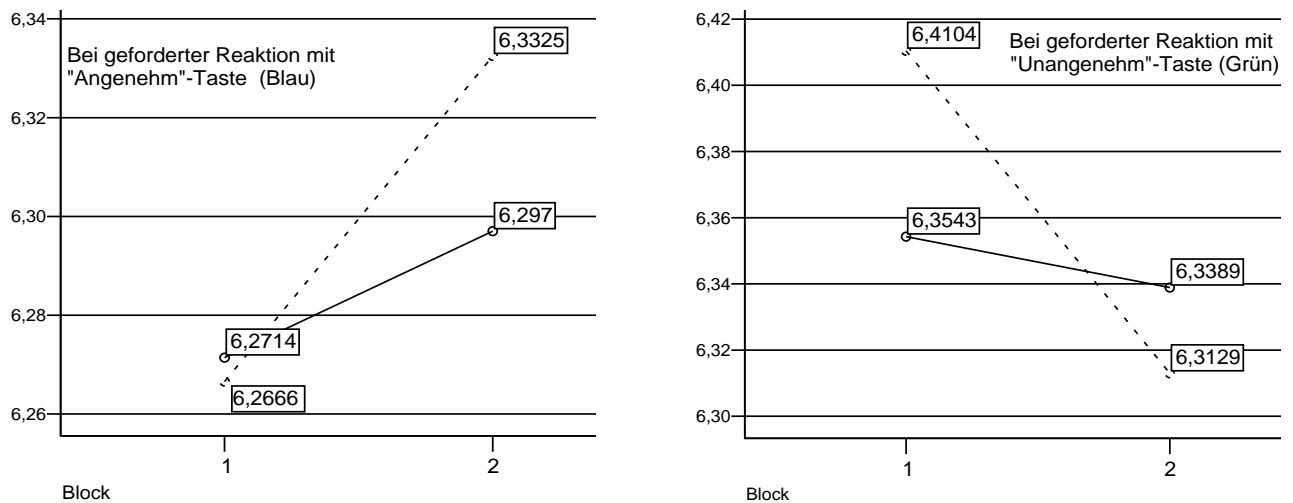


**Abbildung 6-9**, EAST Übergewicht, Reaktionszeiten (In ms), Interaktion zwischen *Block* und *extrinsischer Valenz*. gestrichelte Linie= extrinsisch angenehm (Blau)



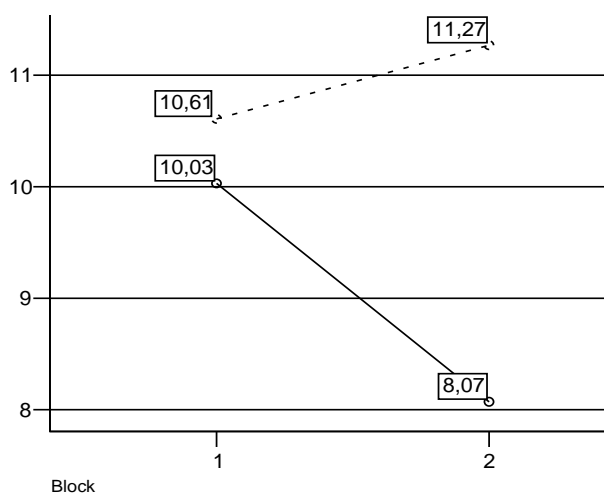
**Abbildung 6-10**, EAST Übergewicht, Reaktionszeiten (In ms), fehlende Interaktion zwischen *Stimulusart* und *extrinsischer Valenz*. Stimulusart 1= Über-, 2= Normalgewicht, gestrichelte Linie= extrinsisch angenehm (Blau)

Dagegen zeigt sich die Dreifach-Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz* signifikant ( $F(1, 156)=29.05$   $p<.001$ ). Sowohl für extrinsisch positive (blaue Stimuli), als auch negative Valenz (grüne Stimuli) treten vorurteilskongruente Effekte erst im zweiten Experimentalblock auf (extrinsisch positiv: höhere Reaktionszeiten für übergewichtige-, als für normalgewichtige Stimuli; extrinsisch negativ: höhere Reaktionszeiten für normalgewichtige-, als für übergewichtige Stimuli), während sich im ersten Experimentalblock umgekehrte Effekte zeigen (Abb. 6-11).

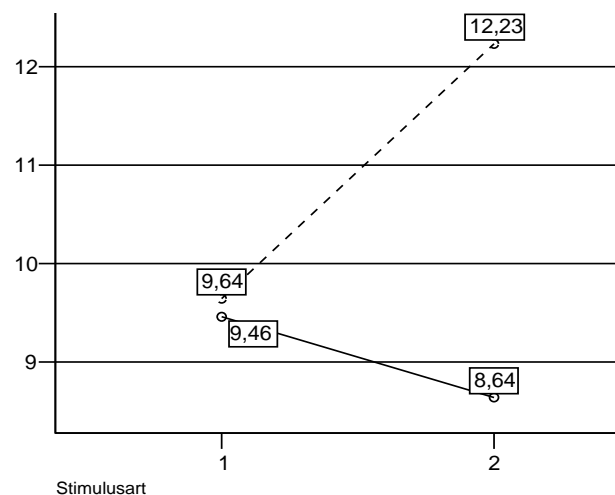


**Abbildung 6-11;** EAST Übergewicht, Reaktionszeiten (In ms), Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz*. gestrichelte Linie= Stimulusart Übergewichtig

Eine Analyse der Fehler-EASTs erbringt einen Haupteffekt *extrinsischer Valenz* ( $F(1, 158)=6.05$   $p<.05$ ), resultierend aus einer höheren Fehlerquote, wenn eine Reaktion mit der „Angenehm“-Taste erforderlich war (Blau,  $M=10.94\%$ ) im Vergleich zu extrinsisch negativer Valenz (Grün,  $M=9.05\%$ ). Die Interaktion zwischen *Block* und *extrinsischer Valenz* ist – analog zu den Reaktionszeiten – für die Fehlerraten signifikant ( $F(1, 158)=4.68$   $p<.05$ , Abb. 6-12). Die kritische Interaktion zwischen Stimulusart und *extrinsischer Valenz*, die in den Reaktionszeit-EASTs nicht nachgewiesen werden konnte, ist signifikant in den Fehlerdaten ( $F(1, 158)=4.80$   $p<.05$ , Abb. 6-13). Diese Interaktion ist jedoch insbesondere auf die unterschiedlichen Fehlerraten bei der Beantwortung normalgewichtiger Stimuli zurückzuführen. Während für übergewichtige Stimuli die Fehlerrate keine Funktion *extrinsischer Valenz* ist, zeigen sich deutlich höhere Fehlerraten, wenn normalgewichtige Stimuli aufgrund der Farbe, in der sie dargeboten sind, mit der Unangenehm-Taste beantwortet werden sollen (Grün).

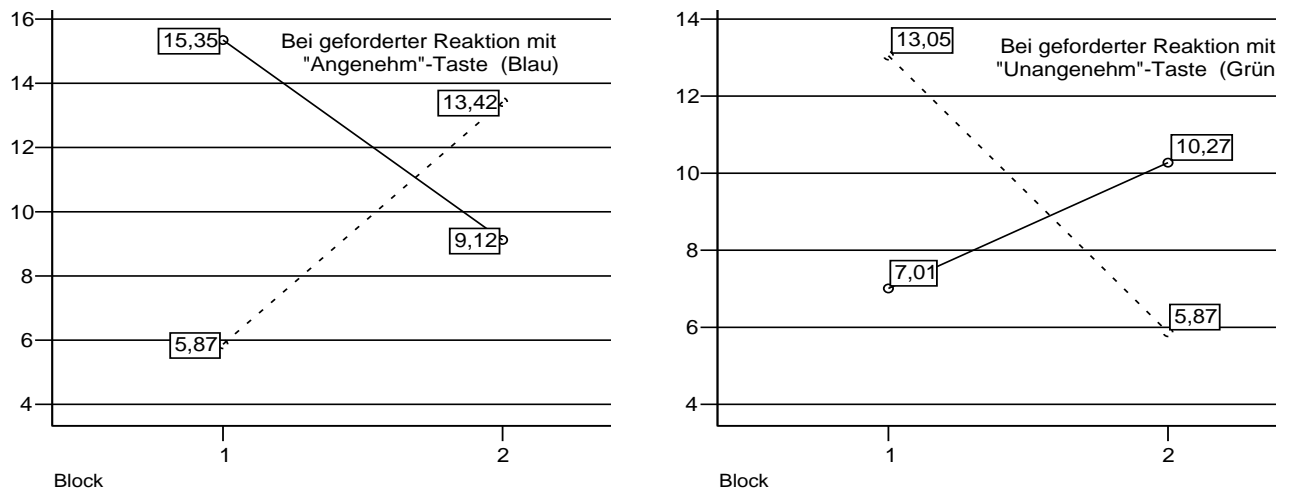


**Abbildung 6-12;** EAST Übergewicht, Fehler, Interaktion zwischen *Block* und *extrinsischer Valenz*. gestrichelte Linie= extrinsisch angenehm (Blau)



**Abbildung 6-13;** EAST Übergewicht, Fehler, Interaktion zwischen *Stimulusart* und *extrinsischer Valenz*. Stimulusart 1= Über-, 2= Normalgewichtig; gestrichelte Linie= extrinsisch angenehm (Blau)

Wie bereits für die Reaktionszeiten zu beobachten war, zeigt sich auch in den Fehlerwerten die Dreifach-Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz* signifikant ( $F(1, 158)=63.28$   $p<.001$ ). Wiederum ergeben sich vorurteilkongruente Effekte allein im zweiten Experimentalblock (Extrinsisch positive Valenz: höhere Fehlerraten für über- im Vergleich zu normalgewichtigen Trials; Extrinsisch negative Valenz: geringer Fehlerraten für übergewichtige- im Vergleich zu normalgewichtigen Trials; Abb. 6-14).



**Abbildung 6-14:** EAST Übergewicht, Fehler, Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz*. gestrichelte Linie= Stimulusart Übergewichtig

#### 6.3.1.4 EAST Homosexualität

Zwei Probanden werden aufgrund ihrer Fehlerraten (21.74%, 15.94%) ausgeschlossen, der resultierende Stichprobenumfang beträgt  $N=160$ . Die Fehlerquote für die Trials des Übungsblocks und der beiden Experimentalblöcke beträgt nunmehr 4.32 (SD 3.03). Analog zum IAT ist eine signifikante Verringerung der Fehlerquote zwischen erstem und zweitem EAST zu beobachten (Mittel der Differenz=1.31,  $T=6.76$ ,  $df=154$ ,  $p<.000$ ).

Für die Haupteffekte wird ein ähnliches Muster wie im EAST Übergewicht beobachtet. Die für den ersten Block berechneten EAST-Effekte sind negativer Natur, jedoch für die Reaktionszeiten nicht signifikant verschieden von Null. Die kompatiblen und inkompatiblen Trials des zweiten Blocks ergeben für Reaktionszeiten- und Fehlerdaten signifikant positive Differenzen. Die über beide Blöcke gemittelten Gesamteffekte sind nicht statistisch bedeutsam (Tab. 6-12)

Block	logarithmierte Latenzen in ms				Fehlerraten in %		
	M (SD)	T	p	M (SD)	T	p	
1 kompatibel	6.28 (.15)			9.95 (10.39)			
1 inkompatibel	6.27 (.13)			7.76 (8.92)			
EAST 1-Effekt	-.01 (.09)	-1.37	.173	-2.19 (12.35)	-2.24	.026	
2 kompatibel	6.28 (.15)			7.76 (9.72)			
2 inkompatibel	6.30 (.15)			11.62 (10.66)			
EAST 2-Effekt	.02 (.09)	2.49	.014	3.85 (12.81)	3.81	.000	
EAST Gesamteffekt	.00 (.06)	.74	.461	.83 (9.46)	1.11	.267	

**Tabelle 6-12**, EAST Homosexualität, Haupteffekte der Reaktionszeiten und Fehlerraten für Block 1, 2 sowie den Gesamteffekt, Test der gepaarten Mittelwertsdifferenzen gegen Null,  $df=159$

Die Reliabilität liegt deutlich unter der des EAST Übergewicht. Im Gegensatz zum EAST Übergewicht übersteigt die fehlerbasierte Split-Half-Reliabilität des ersten und zweiten Blocks mit .23 die latenzbasierte Reliabilität, die lediglich .09 beträgt. Sowohl die EAST-Werte für homosexuelle, als auch die für heterosexuelle Stimuli korrelieren positiv über beide Blöcke hinweg, die Gesamtwerte von Block 1 und 2 stehen dagegen in keiner signifikanten Beziehung zueinander. Übereinstimmend mit dem EAST Übergewicht sind die Korrelationen von Homo- und Heterosexualitäts-EAST innerhalb eines Blocks negativ. Auch hier zeigen die fehlerbasierten EAST-Werte wenig interne Zusammenhänge. Die Korrelation zwischen Latenzen- und Fehler-EAST-Werten beträgt .17 ( $p<.05$ ) für den ersten Block, .00 (n.s.) für den zweiten Block und .25 ( $p<.01$ ) für den Gesamtscore (Tab. 6-13, *Fortsetzung nächste Seite*).

EAST Homosexualität, N=160		Latenzen	Fehlerraten
<b>Korrelationen und Reliabilität der EAST-Scores</b>			
Within-concept Korrelation	EAST_Homo	.19 *	.14
Block 1 x Block 2	EAST_Hetero	.21 **	.18 *
Cross-concept Korrelation	Block 1	-.13	.01
EAST_Homo x EAST_Hetero	Block 2	-.19 *	.07
Split-Half-Reliabilität	EAST_Homo	.35	.31
Block 1 x 2	EAST_Hetero	.32	.32
	gesamt	.09	.23

**Tabelle 6-13**, Reliabilitäten und stimulusspezifische Haupteffekte für EAST Homosexualität, *Fortsetzung nächste Seite*

Deskriptive Statistik (ms)		M (SD)		M (SD)	
Block 1	Homosexuell	-48.83	(89.38)	-3.85	(18.24)
	Heterosexuell	120.40	(78.96)	-.52	(16.55)
	<i>gesamt</i>	<i>-10.26</i>	<i>(61.20)</i>	<i>-2.19</i>	<i>(12.35)</i>
Block 2	Homosexuell	-1.53	(91.70)	5.31	(17.97)
	Heterosexuell	20.27	(81.79)	2.40	(17.01)
	<i>gesamt</i>	<i>9.87</i>	<i>(54.38)</i>	<i>3.85</i>	<i>(12.81)</i>
Gesamt	Homosexuell	-27.36	(75.47)	.73	(13.65)
	Heterosexuell	59.89	(59.28)	.94	(12.92)
	<i>gesamt</i>	<i>-2.36</i>	<i>(42.55)</i>	<i>.83</i>	<i>(9.46)</i>

Fortsetzung Tab. 6-13

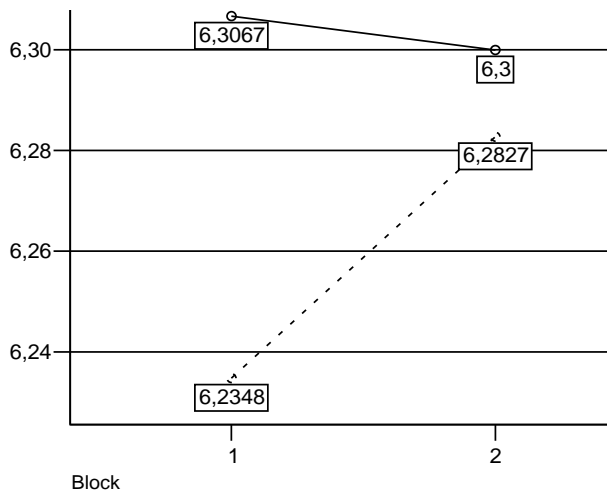
Auch der EAST-Homosexualität liefert unterschiedliche Effekte in Abhängigkeit von Stimulusart und Block. Sowohl für Reaktionszeiten, als auch für Fehler wird ein hypothesenkonformer EAST-Effekt allein im zweiten Messblock beobachtet. Zur systematischen Analyse dieser Maße wird für die auf Reaktionszeiten und Fehlern basierende Werte jeweils eine 2 (Block) x 2 (Stimulusart, Homosexuell vs. Heterosexuell) x 2 (extrinsische Valenz, angenehm vs. unangenehm) ANOVA mit Messwiederholung berechnet (Deskriptive Statistiken siehe Tab. 6-14).

Block	Stimulus - Response	Rohlatenzen	Latenzen	Fehlerraten	
		in ms	in ln (ms)	in %	
		M (SD)	M (SD)	M (SD)	
1	Homosexuell x	angenehm	522.79 (87.63)	6.25 (.16)	7.08 (11.61)
		unangenehm	567.47 (97.62)	6.33 (.16)	10.94 (14.48)
	Heterosexuell x	angenehm	510.98 (87.73)	6.22 (.16)	8.96 (12.80)
		unangenehm	541.85 (75.02)	6.29 (.13)	8.44 (11.22)
2	Homosexuell x	angenehm	555.26 (95.36)	6.31 (.16)	13.23 (14.49)
		unangenehm	556.57 (101.05)	6.31 (.17)	7.92 (12.25)
	Heterosexuell x	angenehm	528.92 (82.69)	6.26 (.15)	7.61 (11.65)
		unangenehm	548.06 (92.92)	6.29 (.16)	10.00 (13.63)

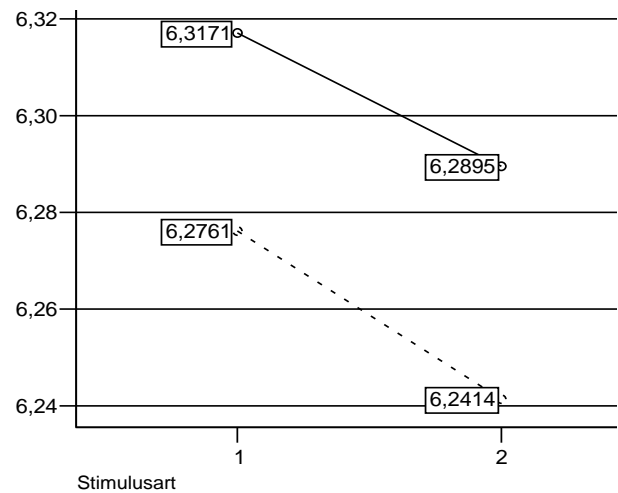
**Tabelle 6-14**, Deskriptive Statistik für ANOVA mit Messwiederholung, EAST Homosexualität

Die Analyse der Reaktionszeiten im EAST-Homosexualität liefert wiederum einen Haupteffekt der *Stimulusart* ( $F(1, 159)=39.73$   $p<.001$ ), resultierend aus kürzeren Latenzen für heterosexuelle ( $M \cong 526$  ms), im Vergleich zu homosexuellen Stimuli ( $M \cong 543$  ms). Analog wird auch hier ein Haupteffekt *extrinsischer Valenz* ( $F(1, 159)=46.79$   $p<.001$ ) beobachtet, aufgrund kürzerer Reaktionszeiten bei extrinsisch positiver Valenz ( $M \cong 527$  ms) im Vergleich zu extrinsisch negativer Valenz ( $M \cong 546$  ms).

Die Interaktion zwischen *Block* und *extrinsischer Valenz* ist signifikant und rührt von höheren Latenzen bei extrinsisch negativer Valenz im ersten Block ( $F(1, 159)=73.22$   $p<.001$ ; Abb. 6-15). Die Interaktion zwischen *Stimulusart* und *extrinsischer Valenz* ist nicht signifikant ( $F(1, 159)=.54$   $p=.47$ , Abb. 6-16).

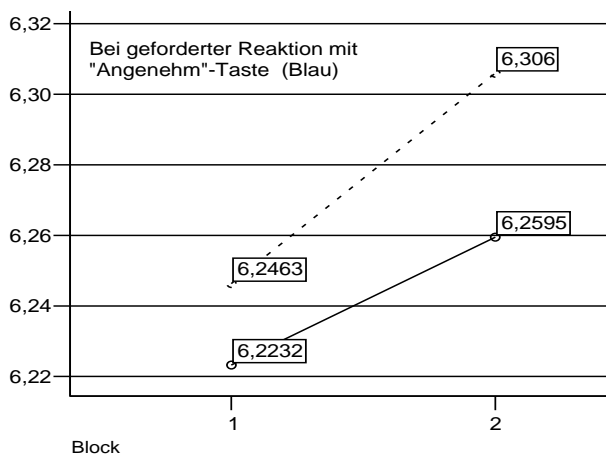


**Abbildung 6-15;** EAST Homosexualität, Latenzen (ln ms), Interaktion zwischen *Block* und *extrinsischer Valenz*. gestrichelte Linie= extrinsisch angenehm (Blau)

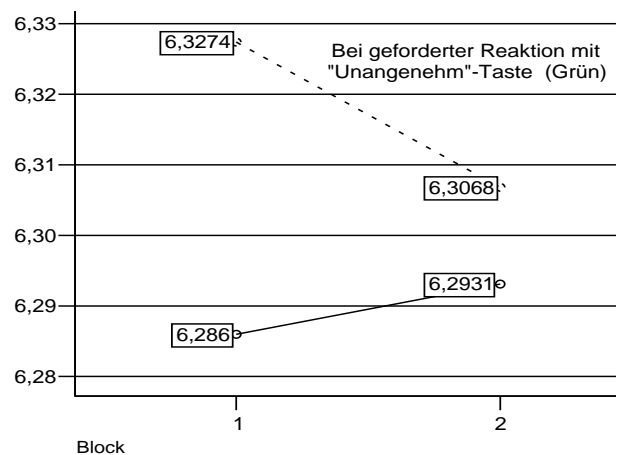


**Abbildung 6-16;** EAST Homosexualität, Latenzen (ln ms), fehlende Interaktion zwischen *Stimulusart* und *extrinsischer Valenz*. Stimulusart 1= Homo-, 2= Heterosexuell, gestrichelte Linie= extrinsisch angenehm (Blau)

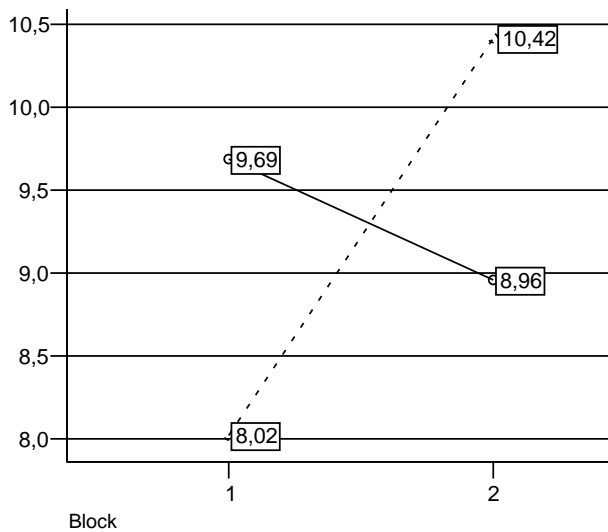
Die Dreifach-Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz* ist wiederum signifikant ( $F(1, 159)=7.33$   $p<.01$ , Abb. 6-17). Der hypothesenkonforme Einfluss extrinsischer Valenz auf die Beantwortung homosexueller Stimuli zeigt sich auch hier im zweiten Block: Während die Latenzen für heterosexuelle Stimuli ungeachtet der extrinsischen Valenz über die Blöcke hinweg (leicht) steigen, nimmt die Reaktionsgeschwindigkeit für homosexuelle Stimuli bei extrinsisch positiver Valenz ab (Reaktionsnachteil), bei negativer Valenz dagegen deutlich zu (Reaktionsvorteil). Diese Interaktion wird jedoch vom Haupteffekt der Stimulusart überlagert, so dass die allgemein höheren Latenzen homosexueller Stimuli auch bei extrinsisch negativer Valenz die der heterosexuellen Stimuli übersteigen.



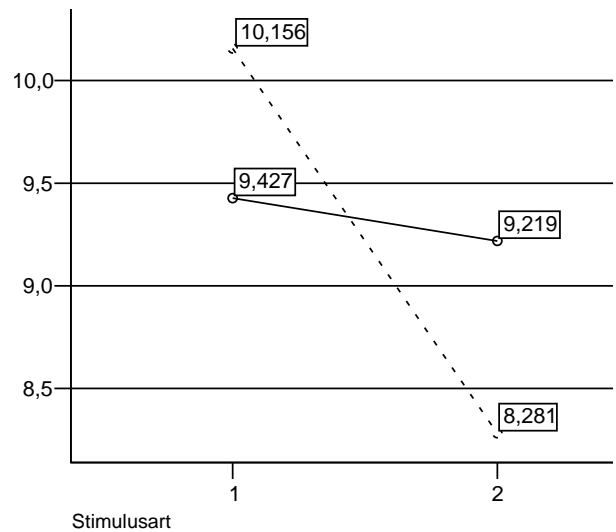
**Abbildung 6-17;** EAST Homosexualität, Reaktionszeiten (ln ms), Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz*. gestrichelte Linie= Stimulusart Homosexuell



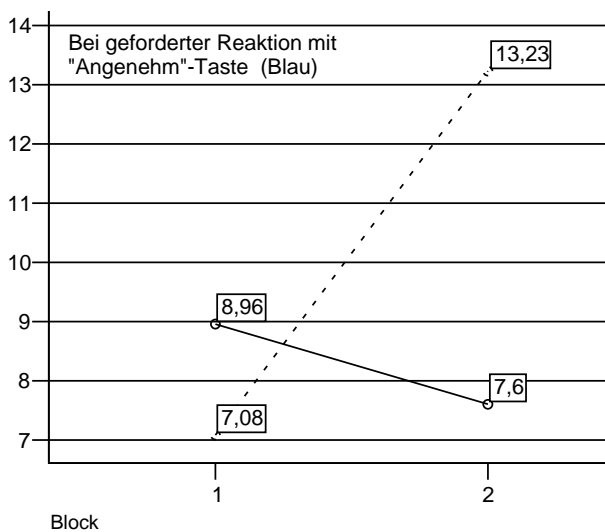
In den Fehlerwerten zeigt sich eine signifikante Interaktion zwischen *Block* und *extrinsischer Valenz* ( $F(1, 159)=6.58$   $p<.05$ ), die auf einen Fehleranstieg bei extrinsisch positiver Valenz und einen leichten Fehlerrückgang bei extrinsisch negativer Valenz von Block 1 zu Block 2 zurückzuführen ist (Abb. 6-18). Die kritische Interaktion zwischen *Stimulusart* und *extrinsischer Valenz* ist auch für die Fehlerwerte des EAST Homosexualität nicht statistisch bedeutsam ( $F(1, 159)=1.24$   $p=.267$ , Abb. 6-19). Die Dreifach-Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz* ist auch hier signifikant ( $F(1, 159)=21.23$   $p<.001$ ) und beruht auf vorurteilskongruenten Fehlerwerten für beide Stimulusarten und extrinsisch positive sowie negative Valenz im zweiten EAST-Block (Abb. 6-20).



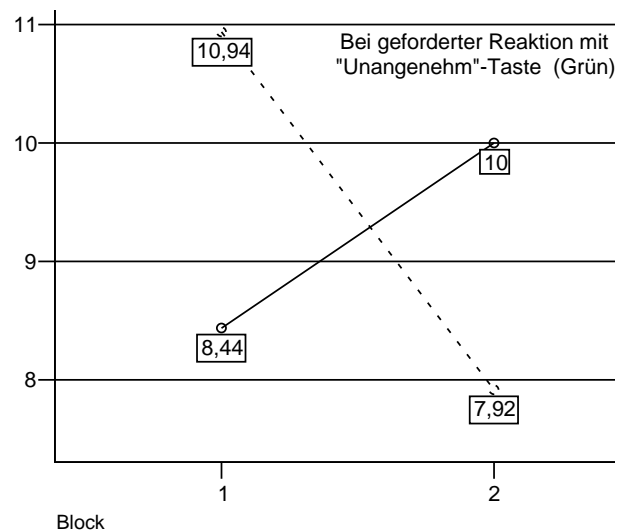
**Abbildung 6-18;** EAST Homosexualität, Fehler, Interaktion zwischen *Block* und *extrinsischer Valenz*. gestrichelte Linie= extrinsisch angenehm (Blau)



**Abbildung 6-19;** EAST Homosexualität, Fehler, fehlende Interaktion zwischen *Stimulusart* und *extrinsischer Valenz*. Stimulusart 1= Homo-, 2= Heterosexuell, gestrichelte Linie= extrinsisch angenehm (Blau)



**Abbildung 6-20;** EAST Homosexualität, Fehler, Interaktion zwischen *Block*, *Stimulusart* und *extrinsischer Valenz*. gestrichelte Linie= Stimulusart Homosexuell





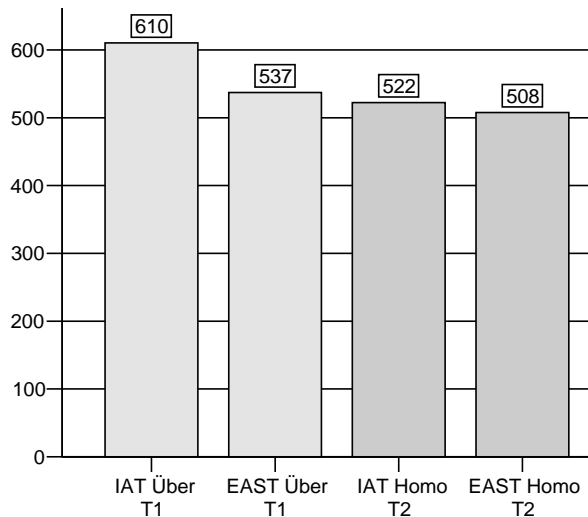
Den Basisanalysen beider EASTs zufolge, werden erwartungskonform gerichtete Effekte sowie die kritische Interaktion zwischen Stimulusart und extrinsischer Valenz mit bemerkenswerter Konsistenz lediglich im zweiten Experimentalblock beobachtet. Da die Effekte jeweils beider Blöcke im EAST Übergewicht und –Homosexualität in fast allen Fällen gegenläufig sind, ist von einer einfachen Mittelung abzusehen. Um die Bedeutung der reaktionszeit- und fehlerbasierten EAST-Effekte je Block besser einschätzen zu können, wird für die Zusammenhangsanalysen daher zunächst auf eine Aggregation verzichtet. Entsprechend der vorhergehenden Analysen sollten insbesondere Reaktionszeit- und Fehlereffekte des jeweils zweiten Experimentalblocks bedeutsam sein. Es müssen zusätzlich jedoch die unterschiedlichen Effekte der Stimulusart berücksichtigt werden, die eine Betrachtung der EAST-Effekte selbst auf Ebene der Experimentalblöcke noch als wenig zweckmäßig erscheinen lassen. Alternativ können stimuluspezifische EAST-Effekte betrachtet werden, wobei das dabei vergleichsweise noch stärker herabgesetzte Aggregationsniveau von Trials die Vergleichbarkeit der Effekte von IAT und EAST erschweren. Auffallend für beide EAST ist der beobachtete Reaktionszeitanstieg für übergewichtige oder homosexuelle Bildstimuli. Diese Ergebnisse können als Stroop-ähnliche Effekte der Aufmerksamkeitsbindung negativer oder bedrohlicher Reize interpretiert werden.

### 6.3.1.5 Mittelwertsvergleiche für IAT und EAST zu T<sub>1</sub> und T<sub>2</sub>

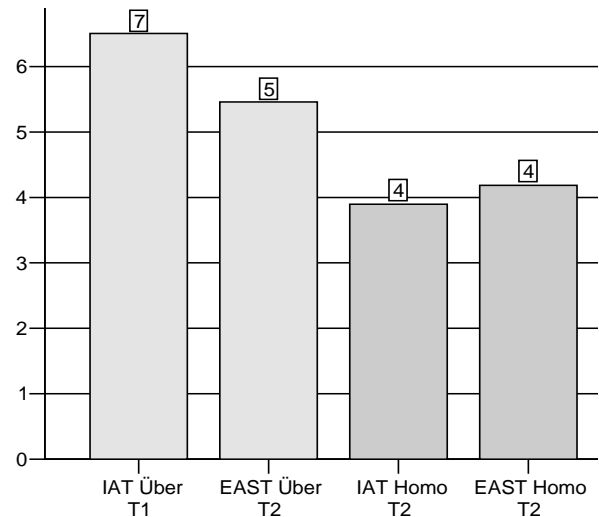
Die mittleren Reaktionszeiten der vier impliziten Maße sowie deren Streuung nehmen von der ersten zur letzten Messung deutlich ab. Auch die Gesamtfehlerrate sinkt, wobei der leichte Anstieg zwischen IAT- und EAST Homosexualität (zu T<sub>2</sub>) nicht signifikant ist (vgl. Tab. 6-15, Abb. 6-21, 6-22).

		<b>T<sub>1</sub>: Übergewicht</b>		<b>T<sub>2</sub>: Homosexualität</b>		<b>Vergleich T<sub>1</sub>-T<sub>2</sub></b>			
		<b>M (SD)</b>		<b>M (SD)</b>		<b>Δ</b>	<b>df</b>	<b>T</b>	<b>p</b>
<b>IAT</b>	Latenzen	613.53	(114.05)	524.29	(65.63)	89.23	154	12.97	.000
	Fehler	6.61	(3.75)	3.84	(2.56)	2.72	154	9.51	.000
<b>EAST</b>	Latenzen	537.71	(70.92)	507.16	(59.78)	29.36	154	7.62	.000
	Fehler	5.48	(3.34)	4.32	(3.03)	1.31	154	6.76	.000
<b>Vergleich</b>		<b>Δ</b>	<b>df</b>	<b>T</b>	<b>p</b>	<b>Δ</b>	<b>df</b>	<b>T</b>	<b>p</b>
<b>IAT - EAST</b>	Latenzen	72.88	155	10.63	.000	14.16	156	3.88	.000
	Fehler	1.00	155	3.16	.001	-.38	156	-1.72	.088

**Tabelle 6-15,** T-Tests für mittlere Reaktionszeiten und Fehlerraten der impliziten Maße



**Abbildung 6-21:** Mittlere Reaktionszeiten in Reihenfolge der Erhebung, hellgrau schattiert: erster Messzeitpunkt, dunkelgrau schattiert: zweiter Messzeitpunkt



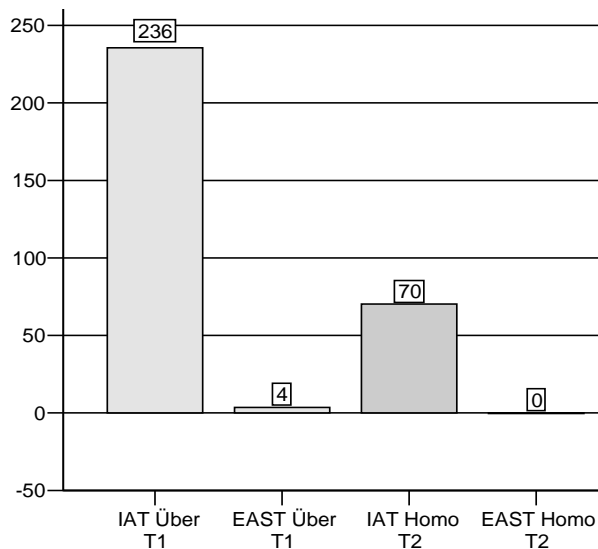
**Abbildung 6-22:** Mittlere Fehlerraten in Reihenfolge der Erhebung, hellgrau schattiert: erster Messzeitpunkt, dunkelgrau schattiert: zweiter Messzeitpunkt

Geringe Reaktionszeiten und Fehlerraten können prinzipiell Ausdruck von Lerneffekte sein, die aus der zunehmenden Erfahrung mit der auf Geschwindigkeit und Korrektheit fokussierenden Aufgabenstellung von IAT und EAST resultierend. Andererseits könnte es sich dabei auch um Effekte des stereotypenrelevanten Materials handeln, da sich IAT und EAST zu beiden Messzeitpunkten lediglich in diesen Eigenschaften unterscheiden.

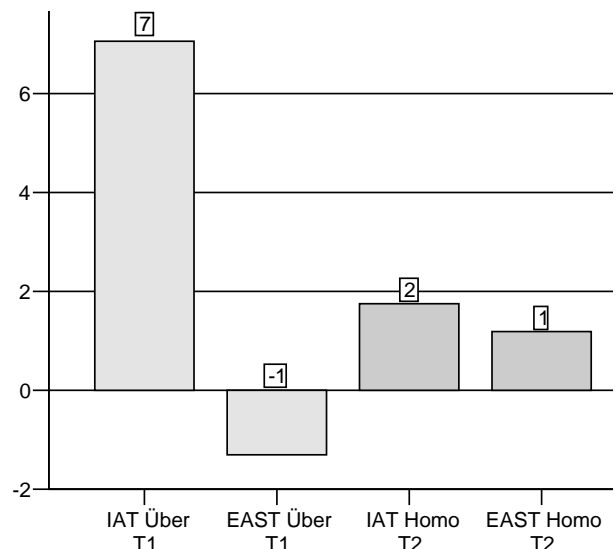
Im Folgenden werden die IAT- und EAST-Effekte miteinander verglichen. Vor der Auswertung der Stichprobenmittelwerte dieser Maße muss zum einen nochmals betont werden, dass das Fehlen jeglicher Zwischensubjektvariation keine gültige Interpretation der Haupteffekte erlaubt. Die betragsmäßig hohen IAT-Effekte könnten allein aufgrund der Blockreihenfolge (kongruent vor inkongruent), die ein Umlernen der Tastenbelegung erfordert, zustande kommen. Im EAST findet, zum Vergleich, kein Aufgabenwechsel statt. Zum anderen wurden für den direkten Effektvergleich die EAST-Blöcke für beide Einstellungsobjekte jeweils gemittelt, um einen Gesamtwert zu erhalten. Da zwischen den Effekten der Blöcke zum Teil große Unterschiede bestehen, besitzen Gesamtwerte eingeschränkte Aussagekraft. Sie dienen allein dem Vergleich, da auch der IAT-Effekt einen Gesamtwert über unterschiedliche Blöcke darstellt (Tab. 6-16; Abb. 6-23, 6-24).

		<b>T<sub>1</sub>: Übergewicht</b>		<b>T<sub>2</sub>: Homosexualität</b>		<b>Vergleich T<sub>1</sub>-T<sub>2</sub></b>			
		<b>M</b>	<b>(SD)</b>	<b>M</b>	<b>(SD)</b>	<b>Δ</b>	<b>df</b>	<b>T</b>	<b>p</b>
<b>IAT</b>	Latenzen	236.32	(169.87)	70.32	(76.86)	167.00	154	13.44	.000
	Fehler	6.90	(7.13)	1.79	(3.61)	5.12	154	8.43	.000
<b>EAST</b>	Latenzen	3.77	(57.29)	-2.36	(42.55)	4.80	154	.87	.388
	Fehler	-1.04	(9.49)	.83	(9.46)	-2.39	154	-2.32	.022
<b>Vergleich</b>		<b>Δ</b>	<b>df</b>	<b>T</b>	<b>p</b>	<b>Δ</b>	<b>df</b>	<b>T</b>	<b>p</b>
<b>IAT - EAST</b>	Latenzen	230.95	155	17.09	.000	70.62	156	10.38	.000
	Fehler	8.17	155	9.31	.000	.89	156	1.06	.298

**Tabelle 6-16,** T-Tests für IAT- und EAST-Gesamteffekte



**Abbildung 6-23,** Reaktionszeiteffekte von IAT und EAST



**Abbildung 6-24,** Fehlereffekte von IAT und EAST

Nicht nur die Mittelwerte von Reaktionszeiten und Fehlerraten beider Verfahren, sondern auch die interessierenden Differenzmaße zwischen kompatiblen und inkompatiblen Trials unterscheiden sich in ihrer Effektgröße in drei von vier Fällen signifikant voneinander. Bis auf die vergleichbaren Fehlereffekte von IAT und EAST zu T<sub>2</sub> werden betragsmäßig höhere Effekte des IAT beobachtet. Mit Ausnahme der fehlerbasierten EAST-Effekte weisen die Maße zudem einen deutlichen Abfall vom ersten zum zweiten Messzeitpunkt auf. Dies könnte inhaltlich begründbare Unterschiede zwischen beiden Einstellungsobjekten und/ oder Effekte der Messwiederholung reflektieren. Hierin bestehen jedoch deutliche Differenzen zwischen IAT und EAST.

Zwischen beiden, also reaktionszeit- und fehlerbasierten, IAT-Effekten wird eine deutliche Abnahme von T<sub>1</sub> zu T<sub>2</sub> hin beobachtet, nichtsdestotrotz ist der IAT Homosexualität (T<sub>2</sub>) jedoch nach wie vor hypothesenkonform gerichtet. Zwischen EAST Übergewicht und -Homosexualität besteht nur für die Fehlereffekte eine signifikante Verringerung von T<sub>1</sub> zu T<sub>2</sub>, die zudem mit einem Vorzeichenwechsel

verbunden ist. Der negative Mittelwert des EAST Übergewicht zu  $T_1$  deutet auf einen vorurteilsinkongruenten Effekt, also eine vergleichsweise positivere Bewertung übergewichtiger bzw. eine negativere Bewertung normalgewichtiger Personen in der Stichprobe hin. Auch für die Reaktionszeiten wird ein Vorzeichenwechsel von  $T_1$  zu  $T_2$  beobachtet, dieser ist jedoch dem der Fehlereffekte entgegengesetzt. Der Haupteffekt der Reaktionszeiten im EAST Homosexualität ( $T_2$ ) weist eine vorurteilsinkongruente Richtung auf, jedoch ist hier der Unterschied zwischen  $T_1$  und  $T_2$  nicht statistisch bedeutsam. Allein hinsichtlich der fehlerbasierten Effekte zu  $T_2$  liegt kein signifikanter Unterschied zwischen IAT und EAST vor. Die Beträge aller Effekte, mit Ausnahme des Fehler-EASTs, tendieren von  $T_1$  zu  $T_2$  zu einer signifikanten Abnahme. Daraus resultieren kleinere Effekte mit geringerer Streuung. Dies kann negative Auswirkungen auf ihre Zusammenhänge zu anderen Maßen haben.

Den Vergleich von IAT und EAST für Übergewicht ( $T_1$ ) und Homosexualität ( $T_2$ ) resümierend, bleiben insbesondere die großen Differenzen zwischen den Effektbeträgen festzuhalten. Für hier verfolgte Fragestellung interessieren jedoch weniger die mittleren Effektbeträge, sondern vielmehr die individuelle Merkmalsvariation und -Kovariation. Die Streuungen der einzelnen Effekte sind zwar ebenfalls unterschiedlich groß, lassen jedoch für jeden Effekt auf eine substantielle Variation zwischen den Probanden schließen.

### 6.3.1.6 Zusammenhänge der impliziten Einstellungsmaße

Eine Analyse der Beziehungen zwischen den vier impliziten Einstellungsmaßen sollte im Sinne von Konstruktvalidität jeweils positive Korrelationen zwischen den reaktionszeit- und fehlerbasierten Maßen je Einstellungsobjekt erbringen (Tab. 6-17).

Von den insgesamt 66 Korrelationen zwischen den impliziten Maßen sind 14 signifikant positiv (21.2% signifikante Statistiken). Betragsmäßig hohe Korrelationen bestehen zwischen den Effekten des IAT-Übergewicht und -Homosexualität (IATÜ, IATH:  $r=.35$ , IATH, IATerH:  $r=.34$ , IATerÜ, IATerH:  $r=.25$ ,  $p<.001$ ). Innerhalb der EAST-Effekte korrelieren besonders die Maße für Einstellungen zu Übergewicht miteinander. Die Reaktionszeiten des ersten und zweiten Blocks des EAST Übergewicht stehen in positiver Beziehung ( $r=.22$ ,  $p<.01$ ), außerdem korrelieren die Reaktionszeiteffekte des zweiten Blocks mit den Fehlereffekten beider Blöcke (EASTÜ2, EASTerÜ1:  $r=.27$ ,  $p<.001$ ; EASTÜ2, EASTerÜ2:  $r=.16$ ,  $p<.05$ ). Für den EAST Homosexualität werden signifikante Korrelationen zwischen den Reaktionszeiteffekten des ersten Blocks und den Fehlereffekten von Block 1 und 2 beobachtet (EASTH1, EASTerH1:  $r=.17$ ; EASTH1, EASTerH2:  $r=.16$ ,  $p<.05$ ). Zusätzlich besteht eine positive Beziehung zwischen den Einstellungsobjekten: Die Reaktionszeiten des jeweils ersten Blocks korrelieren zu  $.24$  ( $p<.01$ ). Insgesamt handelt es sich innerhalb des EAST um relativ geringe Korrelationsbeträge.

Zwischen IAT und EAST bestehen nur innerhalb des Einstellungsobjekts Übergewicht bedeutsame Zusammenhänge. Beide IAT-Effekte korrelieren mit den Reaktionszeiten des zweiten EAST-Blocks (IATÜ, EASTÜ2:  $r=.17$ , IATerÜ, EASTÜ2:  $r=.19$ ,  $p<.05$ ). Außerdem stehen der Fehler-IAT und der Fehler-EAST des ersten Blocks in Beziehung. Eine weitere signifikante Korrelation wird zwischen dem Reaktionszeiteffekt des zweiten Blocks im EAST Übergewicht und dem Reaktionszeiteffekt des IAT Homosexualität beobachtet ( $r=.18$ ,  $p<.05$ ).

Es werden keine Geschlechtseffekte beobachtet, dagegen besteht ein positiver Zusammenhang zwischen dem Alter der Probanden und ihren EAST-Übergewichtseffekten (1. Block, Reaktionszeiten). In negativer Beziehung steht das Alter dagegen zum Fehler-IAT für Homosexualität und zum EAST-Homosexualität (2. Block, Reaktionszeiten).

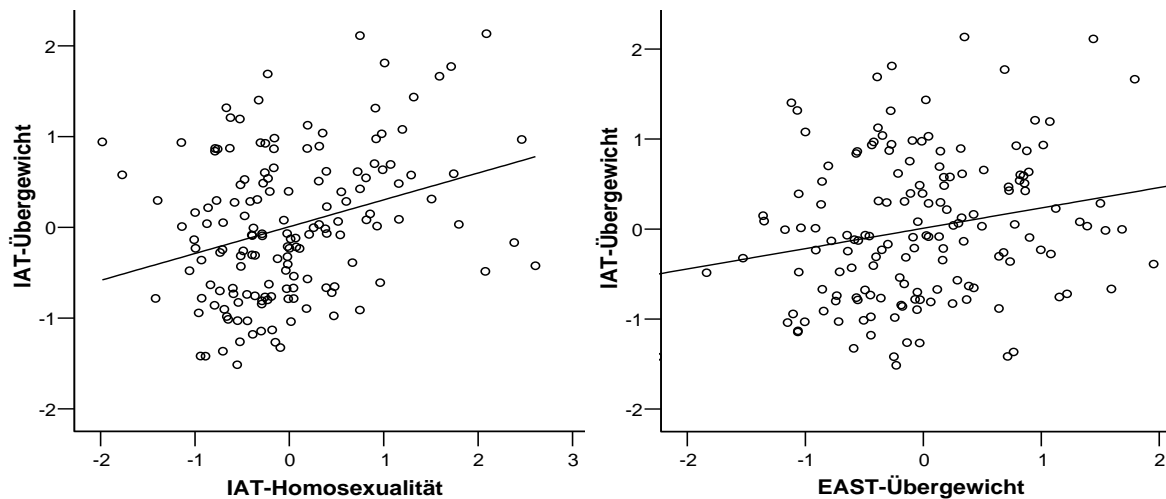
		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	
1.	<b>IATÜ</b>												
2.	<b>IATerÜ</b>	<b>.27</b>											
3.	<b>EASTÜ1</b>	.10	.05										
4.	<b>EASTÜ2</b>	.17	.19	<b>.22</b>									
5.	<b>EASTERÜ1</b>	.10	.17	.05	<b>.27</b>								
6.	<b>EASTERÜ2</b>	.02	.07	.04	.16	.06							
7.	<b>IATH</b>	<b>.35</b>	<b>.25</b>	-.07	.18	.06	.01						
8.	<b>IATerH</b>	.06	.14	.12	.08	.00	.02	<b>.34</b>					
9.	<b>EASTH1</b>	-.16	.06	<b>.24</b>	-.08	.01	-.09	.03	-.01				
10.	<b>EASTH2</b>	.01	.01	-.03	-.04	.12	-.12	-.00	.04	.05			
11.	<b>EASTERH1</b>	-.15	.03	-.01	.0	-.03	.06	-.03	.14	.17	.14		
12.	<b>EASTERH2</b>	-.05	.07	.01	-.09	-.11	.04	-.11	.05	.16	.00	.13	
		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
<b>Alter</b>		.14	-.01	-.02	<b>.22</b>	.15	-.01	.16	.10	-.16	-.07	-.18	-.04
<b>Geschlecht</b>		.11	-.10	-.08	-.14	-.07	-.09	.04	-.15	-.13	-.03	-.03	-.03

**Tabelle 6-17**, Pearson-Korrelationen zwischen den Kennwerten des IAT und EAST (oben); Alters- und Geschlechtseffekte (unten); Ü= Übergewicht, H= Homosexualität; IATÜ, EASTH: latenzbasierte Maße; IATerÜ, EASTERÜ: fehlerbasierte Maße; EASTÜ1/2: Block 1/ Block 2

		1.	2.	3.
1.	<b>IATÜ</b>			
2.	<b>EASTÜ</b>	<b>.22</b>		
3.	<b>IATH</b>	<b>.31</b>	.10	
4.	<b>EASTH</b>	-.05	-.02	.03

**Tabelle 6-18**, Latenz-Fehler-Effekte

Da sowohl Reaktionszeit- als auch Fehlerraten die relevanten Informationen enthalten können, werden diese Effekte für IAT und EAST jeweils z-standardisiert und innerhalb eines Einstellungsobjektes und einer Messmethode gemittelt. Zwischen den IAT-Effekten für Übergewicht und Homosexualität besteht nach wie vor ein signifikant positiver Zusammenhang ( $r=.31$ ,  $p<.001$ ), ebenso wie zwischen IAT- und EAST-Übergewicht ( $r=.22$ ,  $p<.01$ , Tab. 6-18, Abb. 6-25).



**Abbildung 6-25.** Streudiagramme für IAT- Übergewicht und IAT-Homosexualität (links) und IAT-Übergewicht und EAST-Übergewicht (rechts)

### 6.3.2 Selbstberichtsmessung

Der Stichprobenumfang für die Fragebogendaten ist notwendigerweise an den Versuchspersonenverlust bzw. -ausschluss zwischen beiden Messzeitpunkten gebunden und umfasst N=161 Probanden.

#### 6.3.2.1 Vorurteils- und Motivationsskalen

Die Reliabilitäten (Cronbach's Alpha) der Einstellungsskalen sind befriedigend bis gut (Tab. 6-19). Modern Racism (MRS) weist dagegen eine geringe interne Konsistenz (.60) auf. Mit Ausnahme der Vorurteilskontrollskala (VKS) sind die Reliabilitäten der Motivationsskalen ebenfalls akzeptabel. Die Mittelwerte der individuellen Vorurteilsmaße liegen mit Ausnahme von HOKO*i* nur leicht unter dem theoretischen Skalenmittelwert von 3 (Range von 1.85 für HOKO*i* bis 2.93 für ÜBAF*i*). Dies ist für die explizite Messung von Vorurteilen in Studentenstichproben bemerkenswert, da hier aufgrund einer liberaleren Grundhaltung häufig eher unterdurchschnittliche Werte erhalten werden (Six, 1996). Ausgesprochen hoch im Vergleich dazu schätzen die Probanden die Prävalenz von Vorurteilen innerhalb unserer Gesellschaft ein (antizipierte Dispositionen). Hier sinken die Mittelwerte der Messungen nicht unter das theoretische Skalenmittel und bewegen sich zwischen 3.32 (HOKO*a*) und 4.03 (HOAF*a*). Entsprechend unterscheiden sich die gesellschaftlich antizipierten Einstellungen und Motive für jede Skala signifikant von der angegebenen eigenen Haltung der Probanden. Die Richtung dieser Mittelwertsunterschiede suggeriert in allem Fällen eine deutlich positivere *eigene* Einstellung gegenüber übergewichtigen und homosexuellen Menschen (sowohl kognitiv, als auch affektiv) und eine geringere Ausprägung von Modern Racism. Die Differenzen (antizipierte – individuelle Disposition) bewegen sich betragsmäßig im Mittel zwischen .60 (ÜBAF) und 1.47 (HOKO), also zwischen mehr als einem halben und eineinhalb Skalenwerten. Bei einer fünfstufigen Antwortskala handelt es sich dabei also

um beträchtlich hohe Unterschiede. Im Hinblick auf die vorurteilsbezogene Motivation zeigt sich, wie zu erwarten, eine höhere eigene intrinsische Motivation und eine höhere antizipierte extrinsische Motivation, während die Skala Motivation zu vorurteilsfreiem Verhalten sowie die Vorurteilskontrollskala auf eine höhere individuelle Ausprägung im Vergleich zur Meinung innerhalb der Gesellschaft schließen lassen. Hier sind jedoch geringere Differenzen zu beobachten.

Die durchweg niedrigen Beträge der Korrelationen zwischen individueller und antizipierter Perspektive weisen zudem auf keine rein additive Skalenverschiebung hin. Hierbei handelt es sich offensichtlich um jeweils zwar nicht unabhängige, jedoch stark dissoziierte Konstrukte. Die Korrelationen zwischen eigener und antizipierter Haltung rangieren zwischen  $-.03$  (n.s., MVV) und  $.28$  ( $p < .001$ , EMS), betragen im Mittel  $.18$  und sind signifikant für die drei Vorurteilsskalen ÜBKO, ÜBAF, HOAF und folgende drei Motivationsmessungen IMS, EMS und VKS.

Explizites Selbstberichtmaß	Deskriptivstatistik						Korrelation		T-Test		
	M	SD	Min	Max	Range	$\alpha$	r	p	$\Delta$	T	p
1. ÜBKO $i$	2.65	.55	1.53	4.05	2.53	.85	.23	.003	-1.06	-21.31	.000
2. ÜBKO $a$	3.71	.48	2.53	4.79	2.26	.85					
3. HOKO $i$	1.85	.52	1.00	3.82	2.82	.84	.14	.073	-1.47	-26.38	.000
4. HOKO $a$	3.32	.56	1.65	4.88	3.24	.87					
5. ÜBAF $i$	2.93	.46	1.60	4.40	2.80	.77	.17	.032	-.60	-13.76	.000
6. ÜBAF $a$	3.53	.41	2.00	4.70	2.70	.78					
7. HOAF $i$	2.74	.71	1.00	4.67	3.67	.89	.22	.005	-1.29	-21.81	.000
8. HOAF $a$	4.03	.44	2.44	5.00	2.56	.82					
9. MRS $i$	2.48	.52	1.25	4.13	2.88	.60	.14	.072	-.90	-17.63	.000
10. MRS $a$	3.38	.46	1.83	4.50	2.67	.58					
11. IMS $i$	3.87	.72	2.17	5.00	2.83	.80	.21	.008	.72	10.97	.000
12. IMS $a$	3.15	.60	1.83	5.00	3.17	.74					
13. EMS $i$	2.54	.78	1.00	4.67	3.67	.76	.28	.000	-.82	-11.84	.000
14. EMS $a$	3.36	.69	1.17	5.00	3.83	.74					
15. MVV $i$	3.76	.50	1.94	5.00	3.06	.82	-.03	.721	.48	8.66	.000
16. MVV $a$	3.27	.49	1.94	4.69	2.75	.83					
17. VKS $i$	3.40	.51	2.00	5.00	3.00	.48	.21	.007	.31	6.15	.000
18. VKS $a$	3.09	.50	1.57	4.71	3.14	.57					

**Tabelle 6-19.** Deskriptive Statistiken und Reliabilitäten der Einstellungs- und Motivationsmessungen, Korrelationen und T-Tests für individuelle Disposition ( $i$ ) und korrespondierende antizipierte Haltung ( $a$ ); ÜBKO=kognitive Einstellungen zu übergewichtigen Menschen, ÜBAF=affektive Einstellungen zu übergewichtigen Menschen, HOKO=kognitive Einstellungen zu Homosexualität, HOAF=affektive Einstellungen zu Homosexualität, MRS= Modern Racism, IMS/EMS=Internal/External Motivation to Respond Without Prejudice, MVV=Motivation zu Vorurteilsfreiem Verhalten, VKS=Vorurteilskontrollskala

Für keines der Maße liegen auffällige Range-Restriktionen vor. Die Spannweiten der expliziten Vorurteilsmaße reichen von 57% (ÜBK0a) bis 92% (HOAF<sub>i</sub>) und betragen im Mittel 2.96 für die individuellen Werte und 2.69 für die Projektionswerte. Insgesamt werden geringere Streuungen, also homogenere Einschätzungen, für die gesellschaftlich antizipierten Werte beobachtet.

### 6.3.2.2 Gefühlsthermometer

Die fünf Thermometer-Maße werden durch millimetergenaues Ausmessen des jeweiligen Füllbereichs gewonnen. Die Thermometer-Skala rangiert von 0° („unangenehm“, entspricht 0 mm) bis 100° („angenehm“, entspricht 75 mm) bilden. Hohe Werte repräsentieren angenehmes Empfinden gegenüber entsprechenden Personen. Die interessierenden Maße hierbei sind insbesondere die Relationen zwischen den Einzelwerten, also die kritischen Empfindungsdifferenzen zwischen Übergewichtigen und Normalgewichtigen, Untergewichtigen und Normalgewichtigen, Übergewichtigen und Untergewichtigen, sowie Homosexuellen und Heterosexuellen. Wie der unteren Hälfte von Tabelle 6-20 zu entnehmen ist, erreicht jede dieser Differenzen statistische Signifikanz in vorurteilskonformer Richtung. Normalgewichtige werden angenehmer empfunden im jeweiligen Vergleich zu übergewichtigen und untergewichtigen Personen, wobei übergewichtige Menschen im Mittel sogar noch angenehmer empfunden werden als untergewichtige. Obwohl diese Differenz mit ca. 5 mm vergleichsweise gering erscheint, erreicht sie Signifikanzniveau. Die Differenz in der gefühlsmäßigen Bewertung homosexueller und heterosexueller Menschen ist ebenfalls signifikant, heterosexuelle Menschen werden im Mittel als angenehmer empfunden.

<b>Gefühlsthermometer Absolutwerte</b>	<b>M (SD)</b>	<b>Min</b>	<b>Max</b>
ThÜb	44.36 (15.80)	0	75
ThNo	60.74 (12.74)	33	75
ThUn	39.22 (17.81)	0	75
ThHo	53.92 (15.46)	0	75
ThHe	61.67 (12.78)	37	75

<b>Thermometervergleich</b>	<b>M (SD)</b>	<b>T</b>	<b>p</b>
Th_NÜ	16.38 (17.06)	12.23	.000
Th_NU	21.53 (17.94)	15.27	.000
Th_UÜ	-5.14 (18.67)	-3.51	.001
Th_HH	7.75 (13.98)	7.06	.000

**Tabelle 6-20**, Thermometermaße, ThÜb= übergewichtige-, ThNo = normalgewichtige-, ThUn= untergewichtige-, ThHo= homosexuelle-, ThHe= heterosexuelle Personen;  
Th\_NÜ= Differenz ThNo-ThÜb,  
Th\_NU= Differenz ThNo-ThUn,  
Th\_UÜ= Differenz ThUn-ThÜb,  
Th\_HH= Differenz ThHe-ThHo



Sämtliche mögliche Korrelationen zwischen den einzelnen Thermometermaßen sind signifikant positiv. Der Zusammenhang zwischen den Maßen für Normal- und Übergewichtige beträgt .30 ( $p < .001$ ), für Normal- und Untergewichtigen .35 ( $p < .001$ ), für Unter- und Übergewichtige .39 ( $p < .001$ ) sowie .52 ( $p < .001$ ) für homo- und heterosexuelle Personen. Die Empfindung gegenüber den beiden Minderheiten übergewichtiger und homosexueller Menschen korreliert zu .56 ( $p < .001$ ); erwartungsgemäß steht die Bewertung der sozialen Majoritäten Normalgewichtiger und Heterosexueller in noch höherem Zusammenhang ( $r = .80$ ,  $p < .001$ ).

### 6.3.2.3 Prädiktormäße I

Mit Ausnahme von zwei Skalen weisen die Prädiktorskalen eine zufrieden stellende bis gute Reliabilität auf. Mit einem Cronbach's Alpha von .51 ist die Skala Konservatismus keine verlässliche Messung, zudem wird hier eine relativ geringe Varianz ( $SD = .36$ ) beobachtet. Auch Soziale Erwünschtheit erreicht mit einer Reliabilität von .64 keine befriedigende interne Konsistenz. Die höchsten Reliabilitäten ( $\alpha = .90$ ) erzielen die Ekelempfindlichkeitsskala sowie die Trait-Skala des STAI. Die Mittelwerte für Autoritarismus und Soziale Dominanzorientierung, zwei zentrale Konstrukte im nomologischen Netzwerk sozialer Einstellungen, liegen mit mindestens einem halben Skalenpunkt weit unter dem theoretischen Mittelwert der Skala von 3. Die höchsten Stichprobenmittelwerte werden dagegen für die Skalen zur Erfassung von Need for Cognition, Konservatismus und Ekelempfindlichkeit beobachtet (siehe Tab. 6-21).

<b>Prädiktormmaß</b>	<b>M (SD)</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b><math>\alpha</math></b>
<b>SES</b>	8.03 (3.08)	.00	16.00	16.00	.64
<b>STAI</b>	2.06 (.44)	1.15	3.30	2.15	.90
<b>EKEL</b>	3.24 (.56)	1.95	4.49	2.54	.90
<b>BDI</b>	9.42 (6.56)	.00	40.00	40.00	.86
<b>UGTS</b>	2.85 (.60)	1.38	4.25	2.88	.70
<b>RWA</b>	2.37 (.55)	1.08	3.58	2.50	.73
<b>SDO</b>	2.50 (.55)	1.00	4.00	3.00	.81
<b>NFC</b>	3.58 (.50)	2.12	5.00	2.88	.81
<b>KONS</b>	3.34 (.36)	1.94	4.38	2.44	.51
<b>BIDR_ST</b>	2.87 (.51)	1.20	4.10	2.90	.70
<b>BIDR_FT</b>	2.55 (.66)	1.00	5.00	4.00	.70

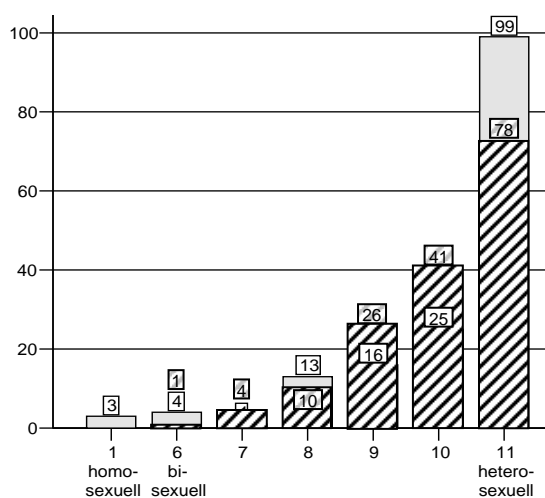
**Tabelle 6-21**, Deskriptive Statistiken der Prädiktormäße I, Fragebogenskalen, SES= Soziale Erwünschtheit, STAI= Trait-Anxiety, EKEL= Ekelempfindlichkeit, BDI= Depressionsneigung, UGTS= Ungewissheitstoleranz, RWA= Right-Wing Authoritarianism, SDO= Soziale Dominanzorientierung, NFC= Need for Cognition, KONS = Konservatismus, BIDR\_ST= Erwünschtheitsverhalten Selbsttäuschung, BIDR\_FT= Erwünschtheitsverhalten Fremdtäuschung

## 6.3.2.4 Prädiktormasse II

Das reale Körperbild der Probanden (KS-Real) liegt mit 4.19 im Mittel unterhalb des theoretischen Skalenmittels. Eigenes und ideales Körperbild weichen zudem signifikant voneinander ab ( $\Delta=.74$ ,  $T=9.61$ ,  $df=160$ ,  $p<.000$ ). Das Körperideal der Probanden (KS-Ideal) liegt im Mittel etwa einen Skalenpunkt unter dem ihres jeweiligen Realbildes. Die Korrelation der Körperschemata beträgt .54 ( $p<.000$ ). Der mittlere BMI in der Untersuchungsstichprobe, liegt mit 21.66 im unteren Normalgewichtsbereich. Die Streubreite des BMI schließt sowohl untergewichtige, als auch adipöse Probanden ein. Die Korrelation zwischen BMI und Körperschema (Real) beträgt .68 ( $p<.000$ ), BMI und körperliches Idealbild korrelieren zu .47 ( $p<.000$ ). Probanden, die gemäß BMI korpulenter sind, wählen also auch ein höheres Körperschema zur subjektiven Selbstbeschreibung (vgl. Tab. 6-22).

Prädiktormaß	M (SD)	Min	Max	Range
<b>KS-Real</b>	4.19 (1.12)	2	9	7
<b>KS-Ideal</b>	3.45 (.86)	2	9	7
<b>KS_diff</b>	-.74 (.98)	-3	2	5
<b>Body-Mass-Index (BMI)</b>	21.66 (2.73)	15.04	32.87	17.84
<b>Geschlechtsidentität</b>	10.10 (1.11)	6	11	5
<b>Sexuelle Orientierung</b>	10.10 (1.74)	1	11	10
<b>MWT-IQ</b>	30.23 (3.44)	22	37	15

**Tabelle 6-22**, Deskriptive Statistiken der Prädiktormasse II, KS-Real=eigenes Körperschema, KS-Ideal=Körperidealbild, KS\_diff=Differenz (KS-Ideal – KS-Real), MWT-IQ=Mehrfachwortwahltest Punktwert



**Abbildung 6-26**; Graue Balken= Sex. Orientierung, Gestreifte Balken= Geschlechtsidentität

Die Identifizierung der Probanden mit ihrem eigenen Geschlecht fällt mit einem Mittelwert von 10.10 durchweg hoch aus (Abb. 6-26). Ebenso stark ist eine heterosexuelle Orientierung, wobei hier eine deutlich höhere Streuung und entsprechend auch Spannweite beobachtet wird, resultierend aus vier Probanden, die mit einem Skalenwert von 6 angeben, bisexuell orientiert zu sein und drei Probanden, die mit dem Wert 1 eine eindeutig homosexuelle Orientierung berichten. Bemerkenswerterweise gibt es keine Variation zwischen bisexueller und homosexueller Orientierung, während jeder Skalenpunkt

während jeder Skalenpunkt zwischen bisexueller und heterosexueller Orientierung mindestens einmal gewählt worden ist (Abb. 6-26). Die Korrelation zwischen sexueller Orientierung und Geschlechtsidentität beträgt .42 ( $p<.001$ ).

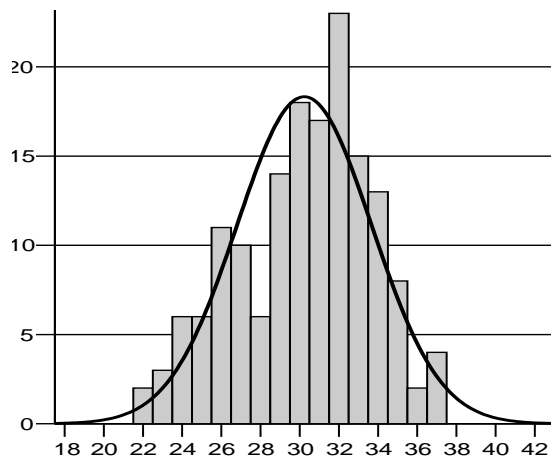


Abbildung 6-27; Punktwerte im MWT

Gemäß Mehrfachwortwahltest beträgt der mittlere Intelligenzquotient der Untersuchungsteilnehmer etwa 111 (Punktwertmittel: 30.23, SD: 3.44). Erwartungsgemäß zeichnet sich die studentische Stichprobe also durch eine im Mittel hohe Intelligenz aus (vgl. Lehrl, Merz, Burkard & Fischer, 1991). Der Kolmogorov-Smirnov-Test zeigt keine signifikante Abweichung von der Normalverteilung an ( $K-S-Z=1.35$ ,  $p=.05$ , vgl. Abb. 6-27). Es liegen zudem keine MWT-Unterschiede für männliche und weibliche Probanden vor ( $\Delta=.48$ ,  $T=.76$ ,  $df=156$ ,  $p=.45$ ), jedoch

besteht ein signifikant positiver Zusammenhang des MWT mit dem Alter ( $r=.30$ ,  $p<.001$ ).

Der Mittelwert der Anzahl übergewichtiger Freunde beträgt 2.75 ( $SD=.98$ , Median= 3  $\cong$  „3 bis 5“). Der Mittelwert der Menge homosexueller Freunde ist mit 2.24 ( $SD= 1.09$ , Median= 2  $\cong$  „1 bis 2“) geringfügig kleiner. Kein Proband gab an, jeweils mehr als 10 übergewichtige oder homosexuelle Freunde zu haben (Abb. 6-28 und 6-29).

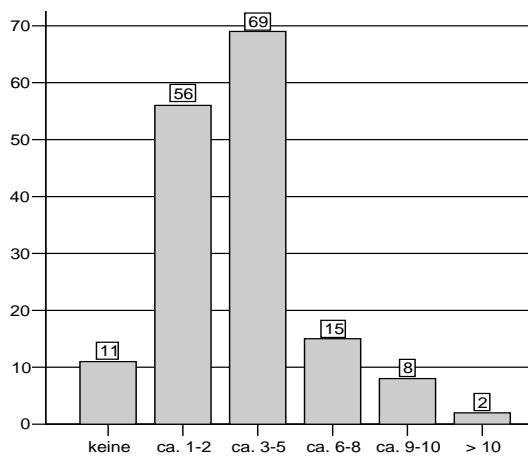


Abbildung 6-28, Anzahl übergewichtiger Freunde und Bekannte

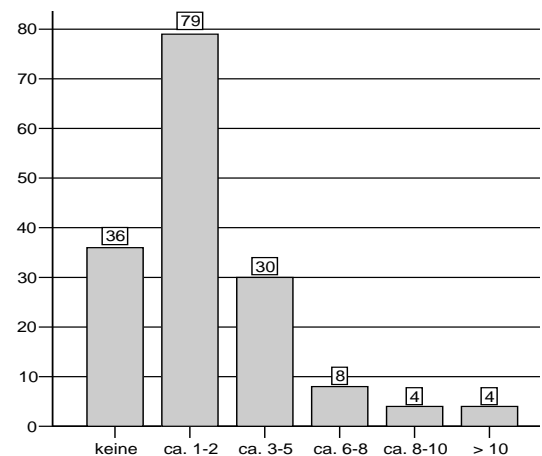


Abbildung 6-29, Anzahl homosexueller Freunde und Bekannte

### 6.3.2.5 Interne Zusammenhänge der Fragebogenmaße

Vorab ist an dieser Stelle festzuhalten: Alle Fragebogenmaße teilen die methodische Varianz der Papier-Bleistift-Messung, wobei jedoch die methodische Ähnlichkeit sowohl unter den reinen *Selbstbericht*maßen (individuelle Einstellungen, Prädiktoren), als auch unter den *antizipierten* Einstellungs- und Motivmaßen jeweils höher sein sollte, als *zwischen* beiden Antwortskalen. Dies sollte allgemein höhere Korrelationen *innerhalb* der individueller sowie *antizipierten* Dispositionen bewirken und eine Minderung der Zusammenhänge *zwischen* diesen Skalen (vgl. Campbell & Fiske, 1959).

### 6.3.2.5.1 *Motivationsmaße*

Die verschiedenen Maße für die Motivation zur Vorurteilsfreiheit weisen relativ hohe positive Zusammenhänge auf, wobei mit Ausnahme von „External Motivation to Respond Without Prejudice“ (EMS) deutlich höhere Korrelationen zwischen den Maßen innerhalb der individuellen (IMS<sub>i</sub>, MVV<sub>i</sub>:  $r=.52$ , IMS<sub>i</sub>, VKS<sub>i</sub>:  $r=.41$ , MVV<sub>i</sub>, VKS<sub>i</sub>:  $r=.50$ ,  $p<.001$ ) und der antizipierten Messebene beobachtet werden, (IMS<sub>a</sub>, MVV<sub>a</sub>:  $r=.58$ , IMS<sub>a</sub>, VKS<sub>a</sub>:  $r=.43$ , MVV<sub>a</sub>, VKS<sub>a</sub>:  $r=.58$ ,  $p<.001$ ), als für die individuellen und antizipierten Versionen einer Motivationskala (IMS<sub>i</sub>, IMS<sub>a</sub>:  $r=.21$ ,  $p<.01$ ; MVV<sub>i</sub>, MVV<sub>a</sub>:  $r=-.03$ ,  $p>.05$ ; VKS<sub>i</sub>, VKS<sub>a</sub>:  $r=.21$ ,  $p<.01$ ). Der Zusammenhang zwischen EMS<sub>i</sub> und EMS<sub>a</sub> ist mit  $r=.28$  ( $p<.001$ ) zwar vergleichbar hoch, jedoch sind die jeweiligen Korrelationen zu den anderen Vorurteilsmaßen deutlich geringer bzw. nicht signifikant. Die jeweiligen Differenzen für eine Vorurteilskala korrelieren ebenfalls sehr hoch miteinander, was jedoch wiederum nicht für EMS gilt (IMS, MVV:  $r=.69$ , IMS, VKS:  $r=.50$ , MVV, VKS:  $r=.52$ ,  $p<.001$ ; Tab. 6-23).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. <b>IMS<sub>i</sub></b>											
2. <b>IMS<sub>a</sub></b>	<b>.21</b>										
3. <b>IMS</b>	<b>-.71</b>	<b>.54</b>									
4. <b>EMS<sub>i</sub></b>	-.09	.07	.13								
5. <b>EMS<sub>a</sub></b>	.18	.17	-.04	<b>.28</b>							
6. <b>EMS</b>	<b>.22</b>	.07	-.14	<b>-.67</b>	<b>.52</b>						
7. <b>MVV<sub>i</sub></b>	<b>.52</b>	-.06	<b>-.48</b>	.15	-.00	-.14					
8. <b>MVV<sub>a</sub></b>	-.11	<b>.58</b>	<b>.51</b>	.13	<b>.22</b>	.06	-.03				
9. <b>MVV</b>	<b>-.44</b>	<b>.44</b>	<b>.69</b>	-.02	.16	.14	<b>-.72</b>	<b>.71</b>			
10. <b>VKS<sub>i</sub></b>	<b>.41</b>	.11	<b>-.26</b>	.09	.07	-.03	<b>.50</b>	.06	<b>-.30</b>		
11. <b>VKS<sub>a</sub></b>	-.06	<b>.43</b>	<b>.36</b>	.03	.18	.11	.06	<b>.58</b>	<b>.36</b>	<b>.21</b>	
12. <b>VKS</b>	<b>-.36</b>	<b>.26</b>	<b>.50</b>	-.05	.09	.11	<b>-.33</b>	<b>.42</b>	<b>.52</b>	<b>-.63</b>	<b>.63</b>

**Tabelle 6-23**, Interne Zusammenhänge der Motivationsmaße, IMS/EMS/MVV/VKS=Differenzenmaße aus antizipierten (a) und individuellen (i) Einstellungen; Abkürzungen siehe Tab. 6-19

Eine orthogonal rotierte Hauptkomponentenanalyse der individuellen und antizipierten Motivationsmaße erbringt gemäß Kaiser-Kriterium (Eigenwerte > 1) und Screeplot eine Dreifaktorenlösung, mit der 66.71% der Gesamtvarianz der 8 Variablen erklärt werden. Der erste Faktor (Eigenwert: 2.30, erklärte Varianz: 28.69%) kann eindeutig als antizipierte IMS-, MVV- und VKS-Disposition identifiziert werden. Den zweiten Faktor (Eigenwert: 1.88, erklärte Varianz: 23.46%) bilden entsprechend die individuell erfassten Dispositionen von IMS, MVV und VKS. Der dritte Faktor (Eigenwert: 1.17, erklärte Varianz: 14.56%) fasst ausschließlich EMS, und zwar sowohl das, als auch das antizipierte Motivationsmaß (vgl. Tab. 6-24). Externale Motivation ist demzufolge nach ein eigenständiges, von individueller und antizipierter Vorurteilsmotivation unabhängiges Konstrukt. Für weitere Analysen zur vorur-

teilsbezogenen Motivation werden fortan die drei orthogonalen Faktoren betrachtet Faktor 1 „Antizipierte Vorurteilmotivation“ (AVMF), Faktor 2 „Individuelle Vorurteilmotivation“ (IVMF) sowie Faktor 3 „External orientierte Vorurteilmotivation“ (EMSF).

	<b>Komponente</b>		
	1=AVMF	2=IVMF	3=EMSF
<b>IMS<sub>i</sub></b>	.00	.81	-.03
<b>IMS<sub>a</sub></b>	.80	.10	.04
<b>EMS<sub>i</sub></b>	-.03	.02	.86
<b>EMS<sub>a</sub></b>	.21	.08	.71
<b>MVV<sub>i</sub></b>	-.08	.82	.10
<b>MVV<sub>a</sub></b>	.86	-.09	.17
<b>VKS<sub>i</sub></b>	.16	.76	.05
<b>VKS<sub>a</sub></b>	.81	.06	.04

**Tabelle 6-24**, Rotierte Faktorenmatrix der Hauptkomponentenanalyse der 4 (Skalen) x 2 (individuell vs. antizipiert) Motivationsmaße

#### 6.3.2.5.2 Prädiktionsmaße

Die Prädiktormäße lassen sich in Einstellungskonstrukte (RWA, SDO, Konservatismus), Persönlichkeitsnahe Konstrukte (NFC, UGTS, MWT-IQ), motivationale Variablen (SES, BIDR) und in habituelle- bzw. Selbstidentifikationsmaße (KS-, BMI, Sex-O, Geschlechts-ID) unterteilen. Die drei eher klinischen Maße zur Traitangst (STAI), Ekelempfindlichkeit (EKEL) und Depressionsneigung (BDI) werden an dieser Stelle lediglich der Vollständigkeit halber betrachtet, nichtsdestotrotz liefert ihre Berücksichtigung in der Untersuchung der Beziehungsstrukturen der Einstellungsprädiktoren interessante Informationen über dieselben. Es ist jedoch zu beachten, dass hier keine Betrachtung unter klinischen Aspekten erfolgen kann. Die klinischen Maßen werden analog zu den eigentlichen Prädiktorskalen als kontinuierliche Merkmalsverteilungen behandelt. Tabelle 6-25 enthält die Zusammenhänge der Prädiktormäße.

Von den 190 Korrelationen erreichen 57 statistische Bedeutsamkeit, was einem Anteil signifikanter Statistiken von 30% entspricht. Autoritarismus und Soziale Dominanzorientierung stehen mit  $r=.37$  ( $p<.001$ ) in vergleichsweise hohem positiven Zusammenhang zueinander (z.B. Altemeyer, 1998, Sidanius & Pratto, 1999) und korrelieren außerdem positiv mit Unsicherheitstoleranz (RWA:  $r=.17$ ,  $p<.05$ ; SDO:  $r=.22$ ,  $p<.01$ ) und Konservatismus (RWA:  $r=.24$ ,  $p<.01$ ; SDO:  $r=.20$ ,  $p<.05$ ). Gleichsam stehen RWA und SDO jedoch in negativem Zusammenhang zu Need for Cognition (RWA:  $r=-.32$ ,  $p<.001$ ; SDO:  $r=-.21$ ,  $p<.01$ ). Autoritärere Probanden beschreiben sich außerdem als leicht schlanker (KS-Real:  $r=-.17$ ,  $p<.05$ ), obwohl kein Zusammenhang zum BMI besteht, zeigen eine deutlich stärkere Identifikation mit dem eigenen Geschlecht ( $R=.29$ ,  $p<.001$ ) und geben signifikant weniger homosexu-

elle Freunde an ( $r=-.21, p<.01$ ). Die negativen Korrelationen zwischen SDO und den Indikatoren sozialer Erwünschtheit, insbesondere der zum Zwecke der Fremdtäuschung motivierten Erwünschtheit (BIDR\_FT:  $r=-.40$ , SES:  $r=-.26, p<.001$ ) veranschaulichen eindrücklich die Nähe des Konstruktes zu Persönlichkeitseigenschaften wie Verträglichkeit (Zitat). Sozial dominantere Probanden haben geringere Punktwerte im MWT ( $r=.17, p<.05$ ), was auch Ausdruck ihrer mangelnden sozialen Erwünschtheitsmotivation sein kann, da die Art der Durchführung des MWT eine Schönfärbung der eigenen Resultate leicht ermöglichte (vgl. Abschnitt 6.2.2.2.1). Jedoch besteht seinerseits kein substantieller Zusammenhang zwischen den MWT-Werten und den verschiedenen Maßen sozialer Erwünschtheit.

Konservativere Probanden neigen zu weniger Selbsttäuschung ( $r=-.25, p<.01$ ), geben ein schlankeres Körperidealbild an ( $r=-.19, p<.05$ ) und weisen analog eine höhere Real-Ideal-Differenz der Figur auf ( $r=-.19, p<.05$ ). Auch aus einer Berücksichtigung der geringen Reliabilität der Skala mittels Minde-rungskorrektur wären keine zusätzlichen bedeutsamen Zusammenhänge für Konservatismus zu erwarten.

Die Erwünschtheitsmaße stehen in erwartungsgemäß positiver Beziehung zueinander. Je sozial erwünschter eine Person angibt zu sein, desto höher ist ihre Selbsteinschätzung bezüglich NFC ( $r=.19, p<.05$ ) und desto schlanker fällt die Wahl der realen Körperform aus ( $r=-.18, p<.05$ ). Die Fremdtäuschung korreliert ebenfalls positiv mit NFC ( $r=.25, p<.01$ ), während die Selbsttäuschung eine ebensolche Beziehung zur figurbezogenen Real-Ideal-Differenz aufweist ( $r=.20, p<.05$ ).

Die Anzahl übergewichtiger Freunde steht zu keinem weiteren Prädiktormass, als der Anzahl homosexueller Freunde, in Beziehung ( $r=.32, p<.05$ ). Letzteres Item korreliert dagegen nicht nur negativ mit RWA (wie bereits erwähnt) und UGTS ( $r=-.16, p<.05$ ), sondern auch positiv mit NFC ( $r=.16, p<.05$ ), dem körperlichen Ideal ( $r=.18, p<.05$ ), dem BMI ( $r=.17, p<.05$ ) und zu einem hohen Ausmaß mit der sexuellen Orientierung ( $r=.52, p<.001$ ).

Als relevant für die beiden Einstellungsobjekte aus den drei klinischen Maßen herausgegriffen steht die Ekelempfindlichkeit steht in negativem Zusammenhang zum BMI ( $r=-.22, p<.01$ ), zur körperlichen Idealkonstitution ( $r=-.26, p<.01$ ), der figürlichen Real-Ideal-Differenz ( $r=-.19, p<.05$ ) und in positivem Zusammenhang zur Geschlechtsidentität ( $r=.20, p<.05$ ). Während die negativen Beziehungen möglicherweise als Konsequenz eines Ekels vor körperlichem Fett interpretiert werden können, bleibt es schwer nachzuvollziehen, weshalb eine stärkere Identifikation mit dem eigenen Geschlecht und eine höhere Ekelempfindlichkeit miteinander einhergehen. Ein Effekt des Alters liegt lediglich für NFC ( $r=-.22, p<.01$ ) und BDI ( $r=.22, p<.01$ ) vor, ältere Probanden berichten ein geringeres Ausmaß an Need for Cognition und eine höhere Depressionsneigung. Geschlechtsunterschiede werden allein für die Fremdtäuschung und die Ekelempfindlichkeit beobachtet, in beiden Fällen liegen höhere Ausprägungen bei Frauen im Vergleich zu Männern vor.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.
1. <b>RWA</b>																			
2. <b>SDO</b>	<b>.37</b>																		
3. <b>NFC</b>	<b>-.32</b>	<b>-.21</b>																	
4. <b>UGTS</b>	.17	<b>.22</b>	<b>-.25</b>																
5. <b>KONS</b>	<b>.24</b>	.20	-.13	-.02															
6. <b>SES</b>	.08	<b>-.26</b>	.19	-.07	.03														
7. <b>BIDR_ST</b>	-.01	-.10	.13	-.18	<b>-.25</b>	<b>.37</b>													
8. <b>BIDR_FT</b>	-.08	<b>-.40</b>	<b>.25</b>	-.06	-.12	<b>.60</b>	<b>.20</b>												
9. <b>KS-Real</b>	-.17	-.11	.00	.01	.04	-.18	-.10	-.05											
10. <b>KS-Ideal</b>	-.14	-.09	.05	-.03	-.19	-.15	.11	.03	<b>.54</b>										
11. <b>KS_diff</b>	.07	.05	.04	-.03	-.19	.07	.20	.08	<b>-.67</b>	<b>.26</b>									
12. <b>BMI</b>	-.11	.06	.01	.00	.01	-.14	-.04	-.07	<b>.68</b>	<b>.47</b>	<b>-.36</b>								
13. <b>Sex-O</b>	.13	.13	-.06	.01	.11	.11	.00	-.05	-.06	-.12	-.02	-.09							
14. <b>G-ID</b>	<b>.29</b>	.07	<b>-.22</b>	.04	.08	.07	.18	-.06	.02	-.06	-.07	-.07	<b>.42</b>						
15. <b>FreundÜ</b>	-.04	.04	.11	.01	-.01	-.09	-.01	.01	.05	.14	.07	.10	-.12	-.12					
16. <b>FreundH</b>	<b>-.21</b>	-.15	.16	-.16	-.09	-.14	.06	-.00	.13	.18	.00	.17	<b>.52</b>	-.15	<b>.32</b>				
17. <b>MWT-IQ</b>	-.12	-.17	.07	.12	-.10	.03	-.03	.13	-.01	.10	.09	-.04	-.07	-.07	-.05	.06			
18. <b>STAI</b>	-.07	.08	-.18	.22	.11	<b>-.37</b>	<b>-.59</b>	-.17	.15	-.11	<b>-.25</b>	.08	-.07	-.20	.01	-.01	.10		
19. <b>EKEL</b>	.09	-.01	-.02	.01	.06	.01	-.06	.10	-.01	<b>-.26</b>	-.19	<b>-.22</b>	.14	.20	-.15	-.13	-.06	.11	
20. <b>BDI</b>	.12	.07	<b>-.28</b>	.10	.09	-.17	<b>-.39</b>	-.12	.12	-.06	-.17	.07	<b>-.23</b>	-.11	.02	.13	.14	<b>.60</b>	-.01

**Tabelle 6-25.** Interne Zusammenhänge der Prädiktormasse; RWA= Right-Wing Authoritarianism, SDO= Soziale Dominanzorientierung, NFC= Need for Cognition, UGTS= Ungewissheitstoleranz, KONS = Konservatismus, SES= Soziale Erwünschtheit, BIDR\_ST= Erwünschtheitsverhalten Selbsttäuschung, BIDR\_FT= Erwünschtheitsverhalten Fremdtäuschung, KS-Real=eigenes Körperschema, KS-Ideal= Körperidealbild, BMI= Body-Mass-Index, Sex-O= Sexuelle Orientierung, G-ID= Geschlechtsidentität, FreundÜ/H= Anzahl übergewichtiger/homosexueller Freunde, MWT-IQ= Punktwert im Mehrfachwortwahltest, Geschlecht: 1=Frauen, 2=Männer STAI= Trait-Anxiety, EKEL= Ekelempfindlichkeit, BDI= Depressionsneigung,

Eine Hauptkomponente der Einstellungsprädiktoren RWA, SDO, NFC, KONS, UGTS, SES, BIDR\_ST und BIDR\_FT legt laut Kaiser-Kriterium eine Vierfaktorenlösung mit einer Gesamtvarianzaufklärung von 72.87% nahe (Tab. 6-26). Der erste Faktor (Eigenwert: 2.29, erklärte Varianz: 28.66%) fast SES und BIDR\_FT und ist somit als extern orientierte soziale Erwünschtheit zu interpretieren. Der zweiten Faktor (Eigenwert: 1.46, erklärte Varianz: 18.29%) wird schwerpunktmäßig von RWA und SDO gebildet. Die hohen Ladungen von NFC (negativ) und UGTS (positiv) auf den dritten Faktor (Eigenwert: 1.05, erklärte Varianz: 13.18%) legen eine Interpretation im Sinne kognitive Geschlossenheit nahe. Der vierte Faktor (Eigenwert: 1.02, erklärte Varianz: 12.75%) – maßgeblich geprägt durch negative Ladungen von Konservatismus und positive Ladungen von Selbsttäuschung – erscheint schwerer zu interpretieren (Tab. 23). Für die weiteren Analysen mit den Einstellungsprädiktoren werden demgemäß vier orthogonale Faktoren betrachte: Faktor 1 „Soziale Erwünschtheit“ (SEF), Faktor 2 „RWA und SDO“ (RSF), Faktor 3 „kognitive Geschlossenheit“ (KGF) und Faktor 4 „Liberalismus und Selbsttäuschung“ (LSF).

	<b>Komponente</b>			
	1=SEF	2=RSF	3=KGF	4=LSF
<b>RWA</b>	.12	.80	.20	-.09
<b>SDO</b>	-.40	.68	.10	-.02
<b>NFC</b>	.13	-.38	-.56	.08
<b>KONS</b>	.09	.40	-.17	-.80
<b>UGTS</b>	.02	.07	.87	.01
<b>SES</b>	.88	.11	-.11	.10
<b>BIDR_ST</b>	.33	.24	-.34	.74
<b>BIDR_FT</b>	.85	-.25	.03	.02

**Tabelle 6-26**, Rotierte Faktorenmatrix der Hauptkomponentenanalyse der 8 Prädiktormasse, Abkürzungen siehe Tab. 5-25

### 6.3.2.5.3 Zusammenhänge zwischen Motivations- und Prädiktorfaktoren

Da die Motivations- und Prädiktionsfaktoren jeweils durch orthogonale Rotation erhalten wurden und aus diesem Grund statistisch unabhängig voneinander sind, werden lediglich die Zusammenhänge zwischen beiden Faktorengruppen dargestellt (Tab. 6-27). Die für die Gesellschaft antizipierte Vorurteilmotivation (AVMF) korrespondiert mit keinem der Prädiktionsfaktoren, die ausnahmslos auf individueller Ebene erfasst wurden. Die individuelle Motivation, keine Vorurteile vorzuweisen (IVMF), korreliert dagegen positiv mit dem Einstellungsfaktor „Soziale Erwünschtheit“ (IVMF, SEF:  $r=.34$ ,  $p<.001$ ) und negativ mit „Autoritarismus und Soziale Dominanzorientierung“ (IVMF, RSF:  $r=-.28$ ,  $p<.001$ ). „External orientierte Vorurteilmotivation“ (EMSF) unterhält eine positive Beziehung zu RSF (EMSF, RSF:  $r=.34$ ,  $p<.001$ ) und korreliert zudem negativ mit „Liberalismus und Selbsttäuschung“ (EMSF, LSF:  $r=-.22$ ,  $p<.001$ ).

		<b>Motivationsfaktoren</b>		
		AVMF	IVMF	EMSF
<b>Prädiktorfaktoren</b>	<b>SEF</b>	.06	<b>.34</b>	-.12
	<b>RSF</b>	-.00	<b>-.28</b>	<b>.34</b>
	<b>KGF</b>	-.06	-.05	-.15
	<b>LSF</b>	-.03	-.13	<b>-.22</b>

**Tabelle 6-27**, Zusammenhänge der Motivations- und Prädiktionsfaktoren, siehe Tab. 6-24 und 6-26



### 6.3.2.5.4 Kriteriumsskalen der Fragebogenmessung

Aufgrund des Umfangs verfügbarer Maße erfolgt die Analyse der Zusammenhangsstruktur expliziter Einstellungen zunächst getrennt für beide Einstellungsobjekte

#### a) Einstellungsobjekt Übergewichtige Menschen

Die relevanten *Einstellungsmaße* zu übergewichtigen Menschen sind die kognitiven (ÜBKO*i/a*) und affektiven eigenen und antizipierten Einstellungen (ÜBAF*i/a*), sowie die jeweilige Differenz zwischen eigener und antizipierter Haltung (ÜBKO/ÜBAF) und die Thermometermessung für normal-, über- und untergewichtige Menschen (ThNo, ThÜb, ThUn) sowie deren Differenzen (Th\_NÜ, Th\_NU, Th\_U-Ü). Weitere bedeutungsvolle *Dispositionen* für das Einstellungsobjekt stellen die Körperschematamaße (Real- und KS-Ideal), deren Differenz (KS\_diff), der BMI und zusätzlich die Anzahl übergewichtiger Freunde (Freunde Ü) dar (Tab. 6-28).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	
1. <b>ÜBKO<i>i</i></b>																	
2. <b>ÜBKO<i>a</i></b>	<b>.23</b>																
3. <b>ÜBKO</b>	<b>-.68</b>	<b>.55</b>															
4. <b>ÜBAF<i>i</i></b>	<b>.46</b>	-.08	<b>-.46</b>														
5. <b>ÜBAF<i>a</i></b>	-.04	<b>.46</b>	<b>.38</b>	.17													
6. <b>ÜBAF</b>	<b>-.40</b>	<b>.41</b>	<b>.65</b>	<b>-.69</b>	<b>.59</b>												
7. <b>ThÜb</b>	<b>-.40</b>	<b>.20</b>	<b>.50</b>	<b>-.50</b>	.10	<b>.48</b>											
8. <b>ThNo</b>	.14	.19	.03	-.04	.08	.09	<b>.30</b>										
9. <b>ThUn</b>	-.08	-.04	.04	-.07	-.02	.04	<b>.39</b>	<b>.35</b>									
10. <b>Th_NÜ</b>	<b>.48</b>	-.05	<b>-.44</b>	<b>.43</b>	-.04	<b>-.38</b>	<b>-.70</b>	<b>.47</b>	-.10								
11. <b>Th_NU</b>	.17	.17	-.02	.05	.07	.02	-.17	<b>.37</b>	<b>-.75</b>	<b>.43</b>							
12. <b>Th_UÜ</b>	<b>.27</b>	<b>-.21</b>	<b>-.39</b>	<b>.35</b>	-.11	<b>-.36</b>	<b>-.48</b>	.08	<b>.63</b>	<b>.50</b>	<b>-.57</b>						
13. <b>KS-Real</b>	<b>-.21</b>	-.05	.14	-.19	.00	.16	.12	-.04	-.09	-.14	.06	-.19					
14. <b>KS-Ideal</b>	<b>-.21</b>	-.18	.04	-.16	-.13	.03	.01	<b>-.23</b>	-.18	-.18	.01	-.17	<b>.54</b>				
15. <b>KS_diff</b>	.06	-.10	-.13	.08	-.11	-.15	-.14	-.16	-.05	.01	-.06	.07	<b>-.67</b>	<b>.26</b>			
16. <b>BMI</b>	-.20	-.17	.04	-.08	-.04	.04	-.04	-.11	-.11	-.05	.04	-.08	<b>.68</b>	<b>.47</b>	<b>-.36</b>		
17. <b>FreundÜ</b>	-.13	-.15	.00	-.16	-.00	.13	.06	-.19	.01	-.20	-.15	-.04	.05	.14	.07	.10	
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
<b>Alter</b>	.11	.02	-.08	.14	-.10	-.19	-.13	-.15	-.16	.00	.05	-.04	-.00	.16	.12	.13	.02
<b>Geschlecht</b>	.11	<b>-.28</b>	<b>-.30</b>	<b>.24</b>	-.17	<b>-.32</b>	-.19	-.17	-.05	.05	-.07	.11	-.09	<b>.40</b>	<b>.40</b>	<b>.25</b>	<b>.21</b>

**Tabelle 6-28.** Interne Zusammenhänge der Maße für das Einstellungsobjekt **Übergewicht** (oberer Teil), Alter- und Geschlechtseffekte (unterer Teil), ÜBKO= kognitive Einstellungen zu übergewichtigen Menschen, ÜBAF= affektive Einstellungen zu übergewichtigen Menschen, ThÜb/No/Un= Gefühlsthermometermessung, Th\_NÜ/NU/UÜ= Thermometerdifferenzen, KS-Real= eigenes Körperschema, KS-Ideal= Körperidealbild, BMI= Body-Mass-Index, FreundÜ= Anzahl übergewichtiger Freunde, Geschlecht: 1= Frauen, 2= Männer

Kognitive und affektive Einstellungen zu Übergewichtigen stehen innerhalb der individuellen, wie der gesellschaftlich antizipierten Messebene in gleichmäßigem positivem Zusammenhang ( $\ddot{U}BK\ddot{O}i$ ,  $\ddot{U}BAFi$   $r=.46$ ,  $p<.01$ ;  $\ddot{U}BK\ddot{O}a$ ,  $\ddot{U}BAFa$   $r=.46$ ,  $p<.001$ ). Die Korrelation zwischen individuellen und antizipierten Maßen einer Skala sind von deutlich geringerem Ausmaß (vgl. Tab. 12) und zwischen beiden Skalen bestehen keine signifikanten Beziehungen über die Messebenen hinweg ( $\ddot{U}BK\ddot{O}i$ ,  $\ddot{U}BAFa$ ,  $r=-.04$ ,  $p>.05$ ;  $\ddot{U}BK\ddot{O}a$  und  $\ddot{U}BAFi$ ,  $r=-.08$ ,  $p>.05$ ). Die Differenzen zwischen individuellen und antizipierten Einstellungen je Skala korrelieren dagegen hoch ( $\ddot{U}BK\ddot{O}$ ,  $\ddot{U}BAF$ :  $r=.65$ ,  $p<.001$ ).

Individuelle kognitive und affektive Einstellungen korrelieren signifikant negativ mit den Absolutwerten des Thermometermaß Übergewichtige (kognitiv:  $r=-.40$ , affektiv:  $r=-.50$ ;  $p<.01$ ), jedoch weder mit den Maßen für normal-, noch untergewichtige Personen. Die Bewertung Normalgewichtiger im Gefühlsthermometer korrespondiert lediglich schwach mit den antizipierten kognitiven Einstellungen ( $r=.19$ ,  $p<.05$ ). Die Werte der Differenzen zwischen individuellen und antizipierten kognitiven und affektiven Einstellungen stehen ebenfalls ausschließlich zum Thermometermaß für Übergewichtige in statistisch bedeutsamer Beziehung (kognitiv:  $r=.50$ , affektiv:  $r=.48$ ;  $p<.001$ ). Für die Thermometerdifferenzen Normal-Übergewichtige (Th\_NÜ) und Unter-Übergewichtige (Th\_UÜ) zeigen sich erwartungskonforme und einheitliche Zusammenhänge zu den Einstellungsskalen. Je größer die jeweiligen Empfindungsdifferenzen, desto negativer kognitive und affektive Einstellungen zu übergewichtigen Menschen, was insbesondere für die individuellen Einstellungen gilt. Entsprechende Korrelation mit antizipierten Einstellungen sind nicht durchgehend signifikant, weisen jedoch einheitlich negative Vorzeichen auf. Die beiden genannten Thermometerdifferenzen korrelieren signifikant negativ mit den Einstellungsdifferenzen eigener und gesellschaftlicher Haltung (NÜ, kognitiv:  $r=-.44$ ; NÜ, affektiv:  $r=-.38$ ; UÜ, kognitiv:  $r=-.39$ , UÜ, affektiv  $r=-.36$ ;  $p<.001$ ). Je höher also die Empfindungsdifferenzen, desto geringer jeweils die Unterschiede zwischen individuellen und antizipierten Einstellungen. Für die Thermometerdifferenz Normal-Untergewichtige werden signifikante, jedoch betragsmäßig geringe Zusammenhänge lediglich zu beiden kognitiven Einstellungsmaßen ( $\ddot{U}BK\ddot{O}i$  und  $\ddot{U}BK\ddot{O}a$ , jeweils  $r=.17$ ,  $p<.05$ ) beobachtet.

Die körperliche Konstitution (KS-Real) der Probanden steht in systematisch negativem Zusammenhang zu deren individuellen Einstellungen (kognitiv:  $r=-.21$ , affektiv:  $r=-.19$ ;  $p<.05$ ). Je schlanker das Körperschema ist, das die Untersuchungsteilnehmer zur Darstellung ihrer Person wählen, desto negativer sind die Einstellungen übergewichtigen Personen gegenüber. Gleiches gilt für das Idealbild des eigenen Körpers (kognitiv:  $r=-.21$ , affektiv:  $r=-.16$ ,  $p<.05$ ). Die Differenz zwischen realem und idealem Körperschema steht lediglich zum Gefühlsthermometer für Normalgewichtige in Beziehung. Je größer die Abweichung zwischen realem und idealem Körperschemata, je schlanker die Probanden also ihr Idealbild angaben, desto negativer sind ihre Empfindung gegenüber normalgewichtigen Personen ( $r=-.16$ ,  $p<.05$ ).

Der BMI korreliert ausschließlich mit individuellen und antizipierten kognitiven Einstellungen (kognitiv:  $r=-.20$ , affektiv:  $r=-.17$ ,  $p<.05$ ) und ebenso wie die Körperschemata sind die Zusammenhänge negativer Natur. Allgemein ist also ein stabiler negativer Zusammenhang zwischen eigenem Gewicht bzw. Körperumfang und den Einstellungen gegenüber übergewichtigen Menschen festzuhalten

Auch die Anzahl übergewichtiger Freunde steht in negativem Zusammenhang zu individuellen affektiven Einstellungen ( $r=-.16$ ,  $p<.05$ ). Die ebenfalls negative Korrelation zum Gefühlsthermometer Übergewichtige ( $r=-.19$ ,  $p<.05$ ) suggeriert ein negativeres Empfinden, je mehr übergewichtige Freunde angegeben werden. Signifikant negativ ist außerdem die Beziehung zwischen der Anzahl übergewichtiger Freunde und der Thermometerdifferenz Normal-Übergewichtige ( $r=-.20$ ,  $p<.05$ ).

Alterseffekte spielen eine eher untergeordnete Rolle für die Einstellungsmaße zu übergewichtigen Menschen. Ältere Probanden zeigen geringere Differenzen zwischen individuellen und antizipierten affektiven Einstellungen ( $r=-.19$ ,  $p<.05$ ) und bewerten untergewichtige Menschen schlechter im Gefühlsthermometer ( $r=-.16$ ,  $p<.05$ ), als jüngere Probanden. Diese Zusammenhänge sind zwar signifikant, jedoch nicht sehr stark im Effekt. Das Geschlecht der Probanden ist dagegen ein wichtiger Prädiktor für die Einstellungen gegenüber Übergewicht. Die individuellen affektiven Einstellungen von Frauen zu Übergewicht sind negativer, als die von Männern ( $r=.24$ ,  $p<.01$ ). Ebenso fällt die Bewertung von Normal- und Übergewichtigen im Gefühlsthermometer leicht negativer unter Frauen aus (ThÜb:  $r=-.19$ , ThNo:  $r=-.17$ ,  $p<.05$ ). Dagegen antizipieren weibliche Probanden signifikant positivere Einstellungen der Gesellschaft gegenüber übergewichtigen Menschen (kognitiv:  $r=-.28$ ,  $p<.001$ ; affektiv:  $r=-.17$ ,  $p<.05$ ) und weisen weniger große Diskrepanzen zwischen individuellen und antizipierten Einstellungen auf (kognitiv:  $r=-.30$ , affektiv:  $r=-.32$ ,  $p<.001$ ).

Die weiblichen Untersuchungsteilnehmer haben einen niedrigeren BMI, als die männlichen ( $r=.21$ ,  $p<.01$ ;  $\Delta=-1.56$ ,  $T=-3.18$ ,  $df= 155$ ,  $p<.01$ ), unterscheiden sich jedoch nicht in der Selbsteinschätzung ihrer körperlichen Konstitution (KS-Real:  $\Delta=.22$ ,  $T=1.12$ ,  $df= 158$ ,  $p=.263$ ). Stattdessen geben Frauen ein signifikant schlankeres Idealkörperbild als Männer an (KS-Ideal:  $r=.40$ ,  $p<.001$ ;  $\Delta=-.68$ ,  $T=-6.06$ ,  $df= 82.20$ ,  $p<.001$ ) und weisen eine deutlich höhere Differenz zwischen realem und idealem Körperbild auf, da ihr Figurideal fast einen Skalenpunkt mehr von ihrem Realbild abweicht, als es für Männer der Fall ist ( $p=.25$ ,  $p<.01$ ;  $\Delta=-.89$ ,  $T=-5.42$ ,  $df= 158$ ,  $p<.001$ ).

Außerdem ist dem Vergleich der Angaben zum Freundes- und Bekanntenkreis zu entnehmen, dass Männer mehr übergewichtige Personen kennen bzw. zu ihren Freunden zählen als Frauen ( $p=.21$ ,  $p<.01$ ;  $\Delta=-.48$ ,  $T=-2.69$ ,  $df= 158$ ,  $p<.01$ ).

## b) Einstellungsobjekt Homosexuelle Menschen

Die relevanten Einstellungsmaße zu homosexuellen Menschen sind die kognitiven (HOKOi/a) und affektiven eigenen und antizipierten Einstellungen (HOAFi/a), die daraus resultierenden Differenzen (HOKO/HOAF), die Thermometermessung für homo- und heterosexuelle Personen (ThHo, ThHe) und die entsprechend hieraus resultierende Differenz (Hetero-Homosexuell: Th\_HH). Weitere wichtige Angaben in diesem Rahmen sind die sexuelle Orientierung, die Geschlechtsidentität und die Anzahl bi- oder homosexueller Freunde (Tab. 6-29).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	
1. <b>HOKOi</b>												
2. <b>HOKOa</b>	.14											
3. <b>HOKO</b>	<b>-.62</b>	<b>.69</b>										
4. <b>HOAFi</b>	<b>.64</b>	.04	<b>-.44</b>									
5. <b>HOAFa</b>	.19	<b>.59</b>	<b>.33</b>	<b>.22</b>								
6. <b>HOAF</b>	<b>-.50</b>	<b>.31</b>	<b>.61</b>	<b>-.82</b>	<b>.38</b>							
7. <b>ThHo</b>	<b>-.55</b>	.01	<b>.41</b>	<b>-.58</b>	-.07	<b>.51</b>						
8. <b>ThHe</b>	.01	.06	.04	-.13	.03	.14	<b>.52</b>					
9. <b>Th_HH</b>	<b>.61</b>	.04	<b>-.42</b>	<b>.52</b>	.10	<b>-.43</b>	<b>-.63</b>	<b>.34</b>				
10. <b>Sex-O</b>	<b>.31</b>	-.04	<b>-.26</b>	<b>.35</b>	.05	<b>-.30</b>	<b>-.20</b>	-.02	<b>.21</b>			
11. <b>G-ID</b>	.05	-.01	-.05	.05	-.05	-.07	.08	.20	.10	<b>.42</b>		
12. <b>FreundH</b>	<b>-.32</b>	-.03	<b>.21</b>	<b>-.37</b>	.01	<b>.35</b>	.18	-.03	<b>-.22</b>	<b>-.52</b>	-.15	
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	
<b>Alter</b>	.10	.11	.01	.19	.03	-.17	<b>-.25</b>	<b>-.21</b>	.08	-.05	-.01	.08
<b>Geschlecht</b>	<b>.38</b>	-.05	<b>-.32</b>	<b>.38</b>	-.02	<b>-.38</b>	<b>-.39</b>	-.16	<b>.28</b>	-.05	-.12	-.02

**Tabelle 6-29,** Interne Zusammenhänge der Maße für das Einstellungsobjekt **Homosexualität** (oberer Teil), Alter- und Geschlechtseffekte (unterer Teil); HOKO= kognitive Einstellungen zu homosexuellen Menschen, HOAF= affektive Einstellungen zu homosexuellen Menschen, ThHo/He= Gefühlsthermometer, Th\_HH= Thermometerdifferenz; Sex-O= sexuelle Orientierung, G-ID= Geschlechtsidentität, FreundH= Anzahl bi- oder homosexueller Freunde, Geschlecht: 1= Frauen, 2= Männer

Hohe positive Korrelationen sind zwischen kognitiven und affektiven Einstellungen innerhalb der individuellen ( $r=.64$ ,  $p<.001$ ) und antizipierten Messebene ( $r=.59$ ,  $p<.001$ ) zu beobachten, während von den restlichen Korrelationen zwischen den Messebenen lediglich zwei signifikant und dabei betragsmäßig deutlich geringer sind (HOAFi, HOAFa:  $r=.22$ ,  $p<.01$ ; HOKOi, HOAFa:  $r=.19$ ,  $p<.05$ ). Die für kognitive und affektive Einstellungen getrennt berechneten Differenzmaße individueller und antizipierter Haltung gegenüber homosexuellen Menschen korrelieren zu  $.61$  ( $p<.001$ ).

Wie bereits für die Einstellungen zu übergewichtigen Menschen beobachtet, besteht ein negativer Zusammenhang zwischen dem Gefühlsthermometermaß der Zielgruppe, also homosexuelle Personen, und den individuellen Einstellungen zu homosexuellen Menschen (kognitiv:  $r=-.55$ , affektiv:  $r=-.58$ ,

$p < .001$ ). Der Differenzwert zwischen individuellen und antizipierten Einstellungen korreliert wiederum positiv mit dem Gefühlsthermometer Homosexuelle. Die über das Thermometermaß erfasste Empfindung gegenüber heterosexuellen Personen, also der sozialen Majorität, zeigt keinerlei Beziehungen zu den Einstellungsmaßen. Die Differenz zwischen beiden Thermometermessungen korrelieren positiv mit den individuellen Einstellungen (HOKO:  $r = .61$ , HOAF:  $r = .52$ ,  $p < .001$ ) und negativ mit den Einstellungsunterschieden individueller und antizipierter Haltung (HOKO:  $r = -.42$ , HOAF:  $r = -.43$ ,  $p < .001$ ).

Die sexuelle Orientierung korreliert erwartungsgemäß positiv mit den individuellen Einstellungen (kognitiv:  $r = .31$ , affektiv:  $r = .35$ ,  $p < .001$ ) und negativ mit den Differenzwerten individueller und antizipierter Einstellungen (kognitiv:  $r = -.26$ , affektiv:  $r = -.30$ ,  $p < .001$ ). Je deutlicher die Probanden eine heterosexuelle Orientierung angeben, desto negativer sind ihre Einstellungen zur Homosexualität und desto geringer ist die Diskrepanz zwischen ihrer individuellen Meinung und der gesellschaftlich antizipierten. Des Weiteren ist ein hypothesengerechter negativer Zusammenhang zwischen der sexuellen Orientierung und dem Gefühlsthermometer Homosexuelle zu beobachten, je stärker die Probanden heterosexuell orientiert sind, desto negativer bewerten sie homosexuelle Menschen. Die Geschlechtsidentität steht ebenso folgerichtig in positiver Beziehung zur Bewertung heterosexueller Personen im Gefühlsthermometer ( $r = .20$ ,  $p < .05$ ).

Die Anzahl bi- oder homosexueller Freunde und Bekannte ist ein stabiler Einstellungsprädiktor. Die Variable (Freunde H) korreliert signifikant negativ mit den individuellen Einstellungen zu homosexuellen Menschen (kognitiv:  $r = -.32$ , affektiv:  $r = -.37$ ,  $p < .001$ ), außerdem besteht ein positiver Zusammenhang zu den beiden Differenzen individueller und antizipierter Einstellungen (kognitiv:  $r = .21$ ,  $p < .01$ ; affektiv:  $r = .35$ ,  $p < .001$ ). Des Weiteren korreliert die Anzahl bi- oder homosexueller Freunde positiv mit der Empfindung gegenüber homosexuellen Menschen ( $r = .18$ ,  $p < .05$ ) und negativ mit der Differenz der Thermometermaße ( $r = -.22$ ), die Probanden mit mehr Freunden anderer als heterosexueller Orientierung machen weniger Unterschiede in der gefühlsmäßigen Bewertung homo- und heterosexueller Menschen. Die ferner beobachtete hohe negative Korrelation zwischen eigener sexueller Orientierung und der Anzahl bi- oder homosexueller Freunde ( $r = -.52$ ,  $p < .001$ ) bestätigt den Indikatorcharakter dieses Items.

Das Alter der Probanden spielt insbesondere für die Bewertung homo- und heterosexueller Menschen im Gefühlsthermometer eine Rolle. Je älter die Probanden sind, desto negativer empfinden sie sowohl homo- als auch heterosexuelle Personen (ThHo:  $r = -.25$ , ThHe:  $r = -.21$ ,  $p < .01$ ). Die individuellen affektiven Einstellungen zu Homosexualität älterer Probanden sind zudem negativer als die jüngerer ( $r = .19$ ,  $p < .05$ ) und die Diskrepanz zwischen individuellen und antizipierten affektiven Einstellungen ist geringer ( $r = -.17$ ,  $p < .05$ ).

Das Geschlecht der Probanden hat erwartungskonforme Bedeutung für die individuellen Einstellungen zu Homosexualität und die jeweilige Differenz individueller und antizipierter Einstellungen. Männer geben signifikant negativere individuelle Einstellungen zu homosexuellen Menschen an als Frauen (kognitiv:  $r=.38$ , affektiv:  $r=.38$ ,  $p<.001$ ) und demzufolge nach ist die Differenz zwischen eigener und gesellschaftlicher Meinung, die allgemein negativer antizipiert wird (vgl. Tab. 6-19), bei Männern deutlich geringer (kognitiv:  $r=.32$ , affektiv:  $r=.38$ ,  $p<.001$ ). Im Gefühlsthermometer werden sowohl homo- als auch heterosexuelle Menschen von Männern negativer bewertet (ThHo:  $r=-.39$ ,  $p<.001$ ; ThHe:  $r=-.16$ ,  $p<.01$ ). Da homosexuelle Menschen von Männern jedoch als sehr viel negativer empfunden werden, als heterosexuelle, steht auch die Differenz beider Thermometermaße in Zusammenhang zum Geschlecht ( $r=.28$ ,  $p<.001$ ).

Keine Geschlechtsdifferenzen lassen sich für die sexuelle Orientierung, die Geschlechtsidentität und die Anzahl homosexueller Freunde und Bekannter feststellen.

### c) Zusammenhänge zwischen beiden Einstellungsobjekten

Innerhalb der individuellen und der antizipierten Messebene bestehen positive Zusammenhänge zwischen den kognitiven und affektiven Einstellungsmaßen übergewichtiger und homosexueller Menschen (z.B. ÜBK*O*<sub>*i*</sub>, HOKO*i*:  $r=.29$ , ÜBK*O*<sub>*a*</sub>, HOKO*a*:  $r=.42$ , ÜBAF*i*, HOAF*i*:  $r=.37$ ,  $p<.001$ ). Keine der Korrelationen zwischen individueller und antizipierter Messebene erreicht die Signifikanzschwelle. Die jeweiligen Differenzen der beiden Messebenen stehen über beide Einstellungsobjekte hinweg in positivem Zusammenhang (ÜBK*O*, HOKO:  $r=.37$ , ÜBAF, HOAF:  $.48$ ,  $p<.001$ ), ebenso die Differenzen der Gefühlsthermometer, wobei sich erwartungsgemäß der stärkste Zusammenhang zwischen der Differenz Heterosexuell-Homosexuell und Normalgewicht-Übergewicht beobachten lässt (Th\_NÜ, Th\_HH:  $r=-.47$ ,  $p<.001$ ; Tab. 6-30).

	ÜB- KO <i>i</i>	ÜB- KOg	ÜBK O	ÜB- AF <i>i</i>	ÜB- AFa	ÜBAF	ThÜb	ThNo	ThUn	Th_ NÜ	Th_ NU	Th_ UÜ
<b>HOKO<i>i</i></b>	<b>.29</b>	.02	<b>-.24</b>	<b>.23</b>	.07	-.14	<b>-.26</b>	.04	-.03	<b>.27</b>	.06	.19
<b>HOKO<i>a</i></b>	.08	<b>.42</b>	<b>.25</b>	.02	<b>.36</b>	<b>.25</b>	.02	.06	.01	.02	.04	-.01
<b>HOKO</b>	-.15	<b>.32</b>	<b>.37</b>	-.15	<b>.23</b>	<b>.30</b>	<b>.21</b>	.02	.03	-.18	-.01	-.15
<b>HOAF<i>i</i></b>	<b>.21</b>	-.10	<b>-.25</b>	<b>.47</b>	.04	<b>-.35</b>	<b>-.26</b>	-.06	.00	.20	-.05	<b>.22</b>
<b>HOAFa</b>	-.02	<b>.34</b>	<b>.28</b>	.05	<b>.39</b>	<b>.24</b>	-.02	-.02	-.10	.00	.04	-.03
<b>HOAF</b>	<b>-.21</b>	<b>.29</b>	<b>.40</b>	<b>-.42</b>	.19	<b>.48</b>	<b>.24</b>	.05	-.03	-.18	.07	<b>-.23</b>
<b>ThHo</b>	<b>-.23</b>	.13	<b>.30</b>	<b>-.30</b>	.12	<b>.34</b>	<b>.56</b>	<b>.44</b>	<b>.31</b>	-.19	.00	-.18
<b>ThHe</b>	-.02	.12	.11	-.11	.13	.18	<b>.34</b>	<b>.80</b>	<b>.27</b>	<b>.28</b>	<b>.29</b>	-.02
<b>Th_HH</b>	<b>.24</b>	-.03	<b>-.23</b>	<b>.24</b>	-.02	<b>-.21</b>	<b>-.31</b>	<b>.24</b>	-.09	<b>.47</b>	<b>.27</b>	.17

**Tabelle 6-30**, Zusammenhänge zwischen den Maße für beide Einstellungsobjekte; Abkürzungen siehe Tab. 6-28 und 6-29

Werden die 8 Einstellungsmaße (kognitiv und affektive individuelle und antizipierte Einstellungen gegenüber Übergewicht und Homosexualität) in einer VARIMAX-rotierten Hauptkomponentenanalyse auf die inhärenten Zusammenhangsdimensionen untersucht, zeigt sich weder eine vorrangige Bedeutung des Einstellungsobjekts, noch eine Dominanz der Messebene individuell vs. antizipiert, was jeweils in einer Zweifaktorenlösung resultieren müsste. Die Analyse erbringt eine Dreifaktorenlösung, mit der insgesamt 69.52% der Gesamtvarianz erklärt werden. Die Faktoren sind unschwer zu interpretieren: Der erste Faktor fasst allgemeine antizipierte Einstellungen (Eigenwert: 2.50, erklärte Varianz: 31.27%). Der zweite Faktor stellt die individuellen Einstellungen zu Homosexualität dar (Eigenwert: 1.97, erklärte Varianz: 24.66%) und der dritte Faktor die individuellen Einstellungen zu Übergewicht (Eigenwert: 1.09, erklärte Varianz: 13.60%; Tab. 6-31).

	<b>Komponente</b>		
	1	2	3
<b>ÜBKO<sub>i</sub></b>	.06	.10	.90
<b>ÜBKO<sub>a</sub></b>	.75	-.28	.29
<b>HOKO<sub>i</sub></b>	.11	.80	.17
<b>HOKO<sub>a</sub></b>	.79	.11	-.04
<b>ÜBAF<sub>i</sub></b>	-.02	.43	.67
<b>ÜBAF<sub>a</sub></b>	.71	.00	.08
<b>HOAF<sub>i</sub></b>	.01	.90	.18
<b>HOAF<sub>a</sub></b>	.76	.33	-.20

**Tabelle 6-31,** Rotierte Faktorenmatrix der individuellen und antizipierten Einstellungsmaße, Abkürzungen siehe Tab. 6-28 und 6-29

An dieser Stelle wird nochmals die außerordentlich stabile Unterscheidbarkeit zwischen individuellen und gesellschaftlich antizipierten Einstellungen deutlich. Während das Einstellungsobjekt für individuelle Einstellungen von Bedeutung ist, spielt für die antizipierten Einstellungen keine Rolle, ob Einstellungen zu Übergewicht oder Homosexualität betrachtet werden – die Gesellschaft wird stringent und undifferenziert voreingenommener gegenüber beiden sozialen Stereotypen eingeschätzt.

In diesem Zusammenhang ist die Unterscheidung zwischen *kognitiven* und *affektiven* Einstellungen relativ unbedeutend – innerhalb eines Einstellungsobjektes könnten die beiden Maße unter Umständen jeweils zu einem Einstellungsindikator zusammengefasst werden. Da hierdurch jedoch – insbesondere für das Einstellungsobjekt Übergewicht – Informationsverluste einkalkuliert werden müssen, bleibt zu analysieren, inwiefern die Unterscheidung zwischen kognitiven und affektiven Einstellungen eine differenzielle Bedeutung hinsichtlich der impliziten Einstellungsmaße hat (vgl. Studie I).

### 6.3.3 Offene Assoziativmessung

#### 6.3.3.1 Deskriptive Statistiken der Geschichtenvariablen

Die Interraterreliabilität (Intra-Klassen-Korrelation, ICC, Zwei-Weg, zufällig, Shrout & Fleiss, 1979) für die 20 zu kodierenden Variablen rangiert zwischen .53 (VOR<sub>a</sub> Ü4) und .98 (Stereo\_H1) und beträgt im Mittel .81. Nach Fleiss (1981) handelt es sich demnach um gute (ICC>.50) bis exzellente (ICC>.75) Interrater-Reliabilitäten.

Zunächst erfolgt die deskriptive Auswertung der Geschichtenvariablen (Tab. 6-32). Hohe Werte von „Valenz“ reflektieren eine positive Wertigkeit (Range 1-5). Hohe Werte für „VOR<sub>i</sub>“ und „VOR<sub>a</sub>“ deuten auf jeweils stärkeren Vorurteilsgehalt der Geschichten hin (Range 1-5). Die Variable „Stereo“ spiegelt die Erkennung bzw. Thematisierung des betreffenden Einstellungsobjekts wider, hohe Werte zeigen eine eindeutige Stereotypenerkennung an (Range 1-3). Eine ausgeprägt kreative Geschichte wird mit hohen Werten der variablen „Kreativ“ kodiert.

Einst.- Objekt	Bild- name	Valenz	VOR <sub>i</sub>	VOR <sub>a</sub>	Kreativ	Stereo	Anzahl
		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
H1	„Früh“	2.56 (.81)	1.11 (.39)	1.03 (.20)	3.91 (.90)	1.41 (.76)	139.30 (70.37)
Ü1	„Eis“	2.41 (.93)	2.27 (1.11)	1.61 (.88)	3.85 (.88)	2.47 (.80)	136.19 (62.94)
H2	„Natur“	3.04 (1.1)	1.20 (.48)	1.32 (.63)	3.87 (.91)	2.40 (.84)	130.04 (70.71)
Ü2	„Schuh“	2.67 (.96)	2.08 (1.75)	1.52 (.65)	3.78 (.85)	2.43 (.81)	119.40 (54.53)
Ü <sub>gesamt</sub>		2.54 (.77)	2.18 (1.15)	1.57 (.61)	3.81 (.77)	2.45 (.60)	127.82 (54.11)
H <sub>gesamt</sub>		2.80 (.72)	1.15 (.36)	1.18 (.35)	3.89 (.80)	1.90 (.56)	134.85 (65.75)
		**	**	**		**	**

**Tabelle 6-32.** Deskriptive Statistiken für die Kodiervariablen, Valenz= Wertigkeit, VOR<sub>i</sub>= individuelle Vorurteile, VOR<sub>a</sub>= antizipierte Vorurteile, Kreativ= Kreativität, Stereo= Stereotypenerkennung, Anzahl= Wortanzahl; Ü/ H= Bilder zum Einstellungsobjekt Übergewicht/ Homosexualität; Ü/ H<sub>gesamt</sub> = Bilder je Einstellungsobjekt gemittelt, \*\*= auf  $p<.01$ -Niveau signifikante Mittelwertsunterschiede zwischen Ü- und H<sub>gesamt</sub>

Zunächst fällt auf, dass die Wortanzahl von der ersten zur dritten zu schreibenden Geschichte leicht- und zwischen dritter und letzter Geschichte stärker abnimmt. Zwischen dritter und vierter Geschichtenlänge besteht ein signifikanter Mittelwertsunterschied ( $\Delta=10.87$ ,  $T=2.01$ ,  $df= 158$ ,  $p<.05$ ). Bis auf die Kreativität, bestehen signifikante Unterschiede zwischen allen je Einstellungsobjekt gemittelten Variablen. Die Richtung der Differenzen weisen für die Geschichten zum Einstellungsobjekt Übergewicht eine geringere Länge (Anzahl:  $\Delta=7.04$ ,  $T=2.62$ ,  $df= 160$ ,  $p<.01$ ), eine negativere Wertigkeit (Valenz:  $\Delta=.27$ ,  $T=3.65$ ,  $df= 160$ ,  $p<.001$ ), stärkere individuelle und antizipierte Vorurteilen (VOR<sub>i</sub>:  $\Delta=-$



.98,  $T=-15.22$ ,  $df=160$ ,  $p<.001$ ; VORa:  $\Delta=-.38$ ,  $T=-7.73$ ,  $df=160$ ,  $p<.001$ ) sowie eine bessere Stereotypenerkennung (Stereo:  $\Delta=-.55$ ,  $T=-9.01$ ,  $df=160$ ,  $p<.001$ ) im Vergleich zum Einstellungsobjekt Homosexualität aus. Eine 6 (Variablen) x 4 (Bilder) ANOVA mit Messwiederholung bestätigt die Unterschiede zwischen den 4 Bildern mit Ausnahme der Kreativität ( $F=26.05$ ,  $df=18, 141$ ,  $p<.001$ ).

### 6.3.3.2 Zusammenhänge der Geschichtenmaße

Um die Zusammenhänge zwischen den Kodiervariablen zu untersuchen, werden die 276 möglichen Korrelationen berechnet, von denen 89 Werte signifikant sind (Tab. 6-33). Systematisch sehr hohe Korrelationen innerhalb der Variablen einer Bildmessung deuten auf mangelnde diskriminante Validität der kodierten Merkmale hin, während hohe Korrelationen zwischen den Kodiervariablen der vier Bilder die erwünschte konvergente Validität der Kodierung bestätigen.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	
1. Ü1A																								
2. Ü1V	-.11																							
3. Ü1Vi	-.07	.06																						
4. Ü1Va	.09	.04	<b>.26</b>																					
5. Ü1K	<b>.54</b>	-.13	.02	-.04																				
6. Ü1S	.06	.15	<b>.55</b>	<b>.29</b>	-.01																			
7. Ü2A	<b>.53</b>	-.06	-.06	.09	<b>.44</b>	-.07																		
8. Ü2V	<b>-.29</b>	.09	-.12	-.14	-.20	-.01	<b>-.26</b>																	
9. Ü2Vi	-.05	-.00	<b>.40</b>	-.06	.00	.06	-.06	<b>-.22</b>																
10. Ü2Va	.05	-.09	-.04	<b>.25</b>	.01	-.13	.06	-.13	.05															
11. Ü2K	<b>.43</b>	-.01	-.08	.07	<b>.55</b>	-.04	<b>.50</b>	-.14	-.20	-.07														
12. Ü2S	<b>-.26</b>	-.06	.03	.00	-.14	-.05	-.17	<b>.25</b>	<b>.26</b>	<b>.32</b>	<b>-.46</b>													
13. H1A	<b>.75</b>	-.06	-.16	.00	<b>.46</b>	-.05	<b>.73</b>	-.16	-.09	.09	<b>.50</b>	-.17												
14. H1V	-.05	<b>.25</b>	-.15	-.08	.05	-.14	.09	.03	-.11	-.15	.14	-.08	.02											
15. H1Vi	.12	-.05	.15	.07	-.02	.15	-.03	-.05	.21	.04	-.14	.06	.07	<b>-.49</b>										
16. H1Va	.16	-.06	-.09	.07	.07	-.01	.08	-.00	-.08	.13	.09	-.11	.22	-.13	.16									
17. H1K	<b>.50</b>	-.09	-.16	.01	<b>.62</b>	-.08	<b>.49</b>	<b>-.27</b>	-.12	.06	<b>.63</b>	<b>-.28</b>	<b>.60</b>	.18	<b>-.23</b>	-.07								
18. H1S	.10	.03	-.06	.05	-.03	-.01	.07	.02	.03	-.05	-.02	.03	.15	-.14	<b>.63</b>	<b>.42</b>	<b>-.21</b>							
19. H2A	<b>.74</b>	-.05	-.08	.11	<b>.45</b>	.02	<b>.54</b>	-.20	-.08	.06	<b>.44</b>	-.11	<b>.70</b>	.01	.10	<b>.22</b>	<b>.45</b>	.10						
20. H2V	-.00	.10	-.11	-.20	.01	-.17	.01	<b>.21</b>	-.00	-.13	-.02	.01	.11	<b>.33</b>	-.07	.11	-.01	.07	.05					
21. H2Vi	.10	-.12	<b>.26</b>	.09	.10	.18	.01	-.11	.19	.01	-.01	.09	-.03	<b>-.30</b>	<b>.32</b>	-.06	-.02	.11	.08	<b>-.38</b>				
22. H2Va	.13	.06	-.07	.19	.06	.02	.05	-.11	-.11	.17	.12	.02	.10	-.11	.18	<b>.26</b>	.08	<b>.22</b>	<b>.22</b>	<b>-.29</b>	<b>.26</b>			
23. H2K	<b>.43</b>	-.10	-.15	-.01	<b>.51</b>	-.06	<b>.32</b>	-.08	-.12	-.02	<b>.62</b>	-.20	<b>.45</b>	.18	-.09	.06	<b>.59</b>	-.05	<b>.57</b>	.13	-.12	.07		
24. H2S	.10	.03	.15	.11	.02	.19	.00	-.07	.01	.02	.00	.05	.01	<b>-.30</b>	<b>.26</b>	.08	-.06	.13	.15	-.19	<b>.59</b>	<b>.47</b>	-.14	

**Tabelle 6-33.** Interne Zusammenhänge der Geschichtenvariablen; Ü1/2= 1./2. Übergewichtsgeschichte, H1/2= 1./2. Homosexualitätsgeschichte; A=Wortanzahl, V=Valenz, Vi=individuelle Vorurteile, Va=antizipierte Vorurteile, S=Stereotypenerkennung, K= Kreativität

Innerhalb der Variablen eines Bildes bestehen systematisch hohe Zusammenhänge zwischen der Wortanzahl und der Kreativität der Geschichte (Ü1:  $r=.54$ , Ü2:  $r=.50$ , H1:  $r=.60$ , H2:  $r=.57$ ,  $p<.001$ ) sowie zwischen den beiden Vorurteilsvariablen (VOR*i*, VOR*a*) und der Stereotyperkennung (Ü1: VOR*i*:  $r=.55$ , VOR*a*:  $r=.29$ ; Ü2: VOR*i*:  $r=.26$ , VOR*a*:  $r=.26$ ; H1: VOR*i*:  $r=.63$ , VOR*a*:  $r=.42$ ; H2: VOR*i*:  $r=.59$ , VOR*a*:  $r=.47$ ;  $p<.01$ ). Umfangreichere Geschichten werden demnach als kreativer bewertet und ein stärkeres Thematisieren individueller oder antizipierter Vorurteile geht mit der Einschätzung deutlicherer Stereotyperkennung einher. Für die Einschätzung der Kreativität bedarf es sicherlich genügen Textmaterial, ähnlich wie für die Beurteilung, inwiefern der jeweilige Stereotyp erkannt wurde, stereotypische Geschichtenthemen einbezogen werden. Hypothesenkonform negative Korrelationen zwischen Valenz und individuellen Vorurteilen bestehen für alle Bilder außer Ü1 (Ü2:  $r=-.22$ ,  $p<.01$ ; H1:  $r=-.49$ , H2:  $r=-.38$ ,  $p<.001$ ).

Die beiden Korrelationsmatrizen der jeweiligen 2 Bilder pro Einstellungsobjekt (dunkelgrau schattiert) veranschaulichen vergleichbare Beziehungsmuster, wie die vier Matrizen der Korrelationen zwischen den Einstellungsobjekten (hellgrau schattiert), wobei die Korrelationsbeträge erstgenannter höher sind. Außerhalb dieser Korrelationsmatrizen bestehen beständig hohe Zusammenhänge zwischen der objektiven Länge und der Kreativität aller vier Geschichten. Diese hohen Korrelationen können als Ausdruck individueller Schreibstile der Probanden interpretiert werden.

Innerhalb der Einstellungsobjekte werden die beiden Geschichtenmessungen gemittelt. Tabelle 6-34 dokumentiert der Vollständigkeit und Übersichtlichkeit halber die resultierenden Zusammenhänge der Kodiervariablen für beiden Einstellungsobjekte sowie Alters- und Geschlechtseffekte.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	
1. <b>ÜbA</b>												
2. <b>ÜbV</b>	.07											
3. <b>ÜbVi</b>	.07	<b>-.46</b>										
4. <b>ÜbVa</b>	<b>.26</b>	-.14	.19									
5. <b>ÜbK</b>	<b>.63</b>	.16	-.17	.04								
6. <b>ÜbS</b>	.14	<b>-.22</b>	<b>.66</b>	<b>.48</b>	-.17							
7. <b>HoA</b>	<b>.86</b>	.02	.06	.15	<b>.55</b>	.10						
8. <b>HoV</b>	-.18	<b>.23</b>	-.12	-.05	<b>-.22</b>	.00	<b>-.29</b>					
9. <b>HoVi</b>	-.13	-.13	<b>.29</b>	-.13	-.18	.05	-.09	-.15				
10. <b>HoVa</b>	.09	-.20	.05	<b>.21</b>	.03	.01	.08	-.16	.04			
11. <b>HoK</b>	<b>.57</b>	.07	-.03	.12	<b>.76</b>	-.01	<b>.62</b>	-.19	-.10	-.03		
12. <b>HoS</b>	-.13	-.17	.20	-.05	<b>-.25</b>	.12	-.20	.17	<b>.37</b>	<b>.20</b>	<b>-.29</b>	
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
<b>Alter</b>	-.15	.00	.09	-.11	-.16	-.01	-.12	.08	<b>.25</b>	-.16	-.06	.02
<b>Geschlecht</b>	-.17	-.02	.02	-.05	-.04	-.12	-.16	-.06	<b>.29</b>	.03	-.06	.05

**Tabelle 6-34**, Zusammenhänge der gemittelten Geschichtenvariablen, Üb: Übergewicht, Ho: Homosexualität (oben), Alters- und Geschlechtseffekte (unten); Abkürzungen siehe Tab. 6-33

Das für die Einzelmessungen beobachtete Beziehungsmuster, zeigt sich auch nach Mittelung der Ko-diervariablen. Die Konvergenz der Variablen zwischen den Einstellungsobjekten (grau schattiert) deutet einerseits auf einen stabilen Schreibstil hin (z.B. für Anzahl:  $r=.86$  und Kreativ:  $r=.76$ ,  $p<.001$ ), kann speziell für die einstellungsrelevanten Variablen (Valenz,  $VOR_i$ ,  $VOR_a$ ) aber auch als grundlegende Tendenz im Umgang mit sozialen Minderheiten betrachtet werden. Allein die Stereotyperkennung in den Bildern zu Übergewicht und Homosexualität steht in keinem signifikanten Zusammenhang zueinander. Dies ist sicher auch der deutlich schwierigeren Identifizierbarkeit homosexueller vor Übergewichtigen Personen im Bildmaterial geschuldet (vgl. Tab. 28), was prinzipiell mehr Möglichkeiten der Alternativinterpretation für die Probanden eröffnet und insbesondere von vorurteilsfrei motivierten Probanden genutzt werden dürfte.

Für eine Reihe der Geschichtenvariablen liegen Alters- und Geschlechtseffekte vor. So werden die Übergewichtsgeschichten ältere Probanden weniger kreativ eingeschätzt (ÜbK:  $r=-.16$ ,  $p<.05$ ), während ihre Homosexualitätsgeschichten von einem stärkeren individuellen und einem geringeren antizipierten Vorurteilsgehalt geprägt sind (HoV*i*:  $r=.25$ ,  $p<.01$ ; HoV*a*:  $r=-.16$ ,  $p<.05$ ). Das Geschlecht steht in Beziehung zur Wortanzahl, Frauen verfassen signifikant längere Geschichten (ÜbA:  $r=-.17$ , HoA:  $r=-.16$ ,  $p<.05$ ), und zu den individuellen Vorurteilen in den Homosexualitätsgeschichten, worin Männer deutlich höhere Werte erzielen (HoV*i*:  $r=.30$ ,  $p<.001$ ).

#### 6.3.4 Zusammenhangsanalyse der Einstellungsmaße

Um eine übersichtliche Ergebnisdokumentation zu gewährleisten, werden zunächst die Zusammenhänge zwischen direkter und indirekter Einstellungsmessung, also zwischen Fragebogen- und impliziten Maßen gemäß IAT und EAST, dargestellt. Im Anschluss daran erfolgt die Dokumentation der Zusammenhangsanalyse für die offene Assoziativmessung mit den Maßen aus Fragebogen, IAT und EAST, um schließlich die simultane Auswertung der Einstellungsmaße mittels latenter Multitrait-Multimethod-Analyse vorzunehmen. Aufgrund der Unregelmäßigkeiten in den Effekten EAST-Blöcke werden an dieser Stelle die ungemittelten Effekte betrachtet.

##### 6.3.4.1 Zusammenhänge expliziter und impliziter Einstellungen

Von den insgesamt 252 Korrelationen erreichen 25 Werte statistische Bedeutsamkeit, wovon 5 Korrelationen innerhalb des Einstellungsobjekts Übergewicht und 10 innerhalb des Einstellungsobjekts Homosexualität bestehen (Tab. 6-35). Das entspricht einem leicht überzufälligen Anteil von knapp 10% signifikanten Ergebnissen.<sup>1</sup>

<sup>1</sup> Die Zusammenhänge für die Gesamteffekte (z.B. IATÜ *gesamt*) sind hier nicht einkalkuliert, da die Gesamteffekte durch Mittelung der reaktionszeit- und fehlerbasierten Effekte erhalten werden.

	Übergewicht								Homosexualität							
	IATÜ	IATerÜ	EASTÜ1	EASTÜ2	EASTerÜ1	EASTerÜ2	IATÜ gesamt	EASTÜ gesamt	IATH	IATerH	EASTH1	EASTH2	EASTerH1	EASTerH2	IATH gesamt	EASTH gesamt
<b>ÜBKO<sub>i</sub></b>	.07	-.05	-.07	.15	-.01	.07	.01	.06	.06	.04	<b>-.21</b>	-.10	-.16	.02	.06	<b>-.19</b>
<b>ÜBKO<sub>a</sub></b>	.03	.07	.03	.17	-.06	.06	.06	.08	-.03	.04	-.13	-.05	-.01	-.03	.01	-.10
<b>ÜBKO</b>	-.03	.09	.08	.01	-.04	-.02	.04	.01	-.07	-.00	.08	.05	.13	-.04	-.04	.09
<b>ÜBAF<sub>i</sub></b>	.10	-.09	-.04	.05	.08	-.01	.01	.04	.11	-.01	<b>-.17</b>	-.05	-.14	-.05	.06	<b>-.17</b>
<b>ÜBAF<sub>a</sub></b>	.08	.04	-.02	.05	.02	.03	.08	.03	.12	.05	-.12	.09	-.05	-.06	.10	-.07
<b>ÜBAF</b>	-.02	.10	.02	-.00	-.05	.02	.05	-.00	-.00	.04	.05	.10	.08	-.01	.02	.10
<b>ThÜb</b>	-.01	<b>.24</b>	-.00	.09	.03	.13	.15	.11	-.10	.00	.09	-.00	.02	.00	-.06	.05
<b>ThNo</b>	.05	.17	.05	.16	.01	.14	.14	.15	-.00	-.05	-.03	-.02	-.07	-.09	-.03	-.09
<b>ThUn</b>	-.07	.17	.14	.07	-.00	.14	.07	.15	-.06	.02	.05	-.07	-.03	-.02	-.02	-.03
<b>Th_NÜ</b>	.04	-.10	.04	.04	-.02	-.02	-.04	.02	.09	-.04	-.11	-.01	-.07	-.07	.03	-.11
<b>Th_NU</b>	.11	-.05	-.10	.04	.01	-.04	.03	-.04	.06	-.06	-.07	.05	-.02	-.04	.00	-.04
<b>Th_UÜ</b>	-.06	-.04	.13	-.01	-.03	.02	-.07	.05	.03	.02	-.03	-.06	-.05	-.02	.03	-.07
<b>HOKO<sub>i</sub></b>	.08	.04	-.02	.01	-.09	-.04	.08	-.06	<b>.30</b>	-.08	.01	.08	.01	.12	.13	.10
<b>HOKO<sub>a</sub></b>	.05	.00	.06	<b>.24</b>	.17	.05	.03	<b>.22</b>	.07	.01	.02	.07	.10	-.11	.05	.03
<b>HOKO</b>	-.02	-.03	.06	.19	.20	.07	-.03	<b>.21</b>	-.16	.06	.00	-.00	.07	-.18	-.06	-.05
<b>HOAF<sub>i</sub></b>	-.01	.05	.01	.07	.04	-.06	.03	.02	<b>.34</b>	.04	-.00	.03	-.04	.02	<b>.23</b>	.01
<b>HOAF<sub>a</sub></b>	.10	-.02	.14	.16	.13	.01	.05	.18	.18	-.05	.04	.11	<b>.21</b>	-.02	.08	.15
<b>HOAF</b>	.07	-.06	.07	.03	.04	.06	.01	.09	<b>-.22</b>	-.06	.03	.03	.16	-.03	-.17	.08
<b>ThHo</b>	.00	.05	.06	.03	.03	.18	.03	.13	<b>-.25</b>	-.01	.02	.03	.04	-.01	-.16	.03
<b>ThHe</b>	.05	.20	-.05	.08	.06	.12	.16	.09	.01	-.11	-.01	.00	.04	-.06	-.06	-.01
<b>Th_HH</b>	.04	.13	-.11	.04	.02	-.08	.11	-.06	<b>.29</b>	-.09	-.03	-.03	-.00	-.04	.12	-.05

**Tabelle 6-35.** Zusammenhänge expliziter und impliziter Einstellungsmaße; IATÜ, EASTÜ: latenzbasierte Effekte für Übergewicht, IATerÜ, EASTerÜ: fehlerbasierte Effekte für Übergewicht, IATH, EASTH: latenzbasierte Effekte für Homosexualität, IATerH, EASTerH: fehlerbasierte Effekte für Homosexualität; IATÜ/H *gesamt* bzw. EASTÜ/H *gesamt* Mittelwert z-standardisierter Latenz- und Fehlereffekte; grau schattiert: Homotrait-Heteromethod-Matrizen, weitere Abkürzungen siehe Tab. 6-28 und 6-29

Die Korrelationen zwischen expliziten und impliziten Einstellungsindizes innerhalb der Einstellungsobjekte entsprechen laut Multitrait-Multimethod-Analyse der Homotrait-Heteromethod-Matrix; sie sollten bei konvergenter Validität der gemessenen Konstrukte größer sein, als die Korrelationen der Maße *zwischen* den Einstellungsobjekten (Heterotrait-Heteromethod-Matrix). Wie Tabelle 6-35 entnommen werden kann, bestehen 7 signifikante Korrelationen zwischen den impliziten Maßen für das Einstellungsobjekt Übergewicht und den Fragebogenmaßen zu Homosexualität sowie drei weitere Korrelation in der zweiten Heterotrait-Heteromethod-Matrix, die jedoch ein negatives Vorzeichen aufweisen.

Konsistente und richtig gerichtete Korrelationen bestehen zwischen dem reaktionszeitbasierten IAT-Homosexualität und insbesondere den individuellen Einstellungen zu Homosexualität gemäß Fragebogen. Je höher die Diskrepanz zwischen inkongruenter und kongruenter Tastenbelegung im IAT,

desto negativer waren die Einstellungen gegenüber homosexuellen Menschen im Fragebogen (HOKO:  $r=.30$ ,  $p<.05$ , HOAFi:  $r=-.22$ ,  $p<.01$ ), desto geringer unterschieden sich individuelle von antizipierten Vorurteilen (HOKO:  $r=-.16$ , HOAF:  $r=.34$ ,  $p<.001$ ), desto unangenehmer wurde die Minderheit im Gefühlsthermometer beurteilt ( $r=-.25$ ,  $p<.001$ ) und desto größer war die Differenz der Gefühlsthermometer heterosexuelle vs. homosexuelle Personen ( $r=.29$ ,  $p<.001$ ).

Für das Einstellungsobjekt Übergewicht bestehen keine signifikanten Zusammenhänge zum reaktionszeitbasierten IAT-Effekt. Der korrespondierende Fehler-IAT korreliert ausschließlich mit allen drei Gefühlsthermometern, am höchsten jedoch mit der Empfindung gegenüber übergewichtigen Personen (ThÜb:  $r=.24$ ,  $p<.001$ , ThNo:  $r=.17$ , ThUn:  $r=.17$ ,  $p<.05$ ).

Für die EAST-Effekte werden lediglich sporadisch Zusammenhänge innerhalb des Einstellungsobjekts Übergewicht beobachtet. Der zweite reaktionszeitbasierte EAST-Block korreliert signifikant positiv mit den antizipierten kognitiven Einstellungen zu Übergewicht ( $r=.17$ ,  $p<.05$ ). Die Korrelation zu den individuellen kognitiven Einstellungen verfehlt mit  $.15$  ( $p=.067$ ) knapp die Signifikanzschwelle von 5%. Für das Einstellungsobjekt Homosexualität sind lediglich die fehlerbasierten EAST-Werte von Bedeutung. Der erste EAST-Block korreliert positiv wiederum mit den antizipierten Einstellungen, jedoch nun mit den *affektiven* Einstellungen ( $r=.21$ ,  $p<.01$ ) und ebenfalls positiv mit der Differenz individueller und antizipierter affektiver Einstellungen ( $r=.16$ ,  $p<.05$ ). Der zweite EAST-Block steht in negativer Beziehung zur Differenz individueller und antizipierter kognitiver Einstellungen ( $r=-.18$ ,  $p<.05$ ). Darüber hinaus unterhält der reaktionszeitbasierte Effekt des zweiten Blocks im EAST-Übergewicht signifikante Zusammenhänge zu den antizipierten Einstellungen gegenüber homosexuellen Menschen (HOKOa:  $r=.24$ ,  $p<.001$ ; HOAFa:  $r=.16$ ,  $p<.05$ ).

Von den durch Mittelung der standardisierten Reaktionszeit- und Fehlereffekte gewonnenen Gesamtmaße für IAT und EAST steht lediglich der IATH in statistisch bedeutsamer und erwartungskonformer Beziehungen zu direkten Einstellungen zu Homosexualität (z.B. IATH, HOAFi:  $r=.23$ ,  $p<.01$ ; IATH, ThHo:  $r=-.16$ ,  $p<.05$ ). Der EASTÜ-Effekt steht dagegen mit HOKOa und HOAFa in signifikantem Zusammenhang (EASTÜ, HOKOa:  $r=.22$ ,  $p<.01$ ; EASTÜ, HOAFa:  $r=.18$ ,  $p<.05$ ).

#### 6.3.4.2 Zusammenhänge expliziter, impliziter und assoziativer Einstellungsmaße

Die Geschichtenmessung bezeichnet die dritte eingesetzte Methode zur Messung von Einstellungen gegenüber übergewichtigen und homosexuellen Menschen. Da die offene Assoziativmessung weder rein direkter, noch indirekter Natur ist, sollten die resultierenden Maße implizite *und* explizite Einstellungsanteile reflektieren. Dies müsste sich in jeweils stärkeren Zusammenhängen zwischen Geschichtenvariablen und expliziten, sowie impliziten Einstellungsmaßen innerhalb eines Einstellungsobjektes niederschlagen.

Die Zusammenhänge werden an dieser Stelle aufgrund der Vielzahl an Maße getrennt für beide Einstellungsobjekte dargestellt. Die Variablen jeweils beider Geschichten werden einzeln und gemittelt betrachtet. Zugunsten einer prägnanten Ergebnisdarstellung wird jedoch lediglich auf die Zusammenhänge für die gemittelten Geschichtenvariablen eingegangen.

#### 6.3.4.3 Einstellungsobjekt Übergewicht

Von insgesamt 60 möglichen Korrelationen zwischen den gemittelten Maßen der offenen Assoziativmessung und den direkten Fragebogenmaßen sind 18 statistisch bedeutsam (30% signifikante Ergebnisse). Die meisten signifikanten Zusammenhänge zu den Fragebogenmaßen ( $N_r=6$ ) weist dabei die Variable individueller Vorurteilsgehalt der Geschichten (ÜbVi) auf. Die Beziehungen zwischen den Geschichtenvariablen und den impliziten Maßen sind weniger stark. Hier bestehen insgesamt lediglich 6 signifikante von insgesamt 48 Korrelationen, deren Beträge deutlich unter denen der Fragebogenzusammenhänge liegen (12.5% statistisch bedeutsame Zusammenhänge).

	Direkte Messung (Fragebogen)										Indirekte Messung (IAT, EAST)							
	ÜBKoi	ÜBKog	ÜBAFi	ÜBAFa	ThÜb	ThNo	ThUn	Th_NÜ	Th_U-Ü	Th_NU	IATÜ	IATERÜ	IATÜ_gesamt	EASTÜ1	EASTÜ2	EASTerÜ1	EASTerÜ2	EASTÜ_gesamt
<b>Ü1A</b>	-.18	-.08	-.13	-.05	<b>.22</b>	.16	.09	-.09	-.10	.02	.09	.14	.14	-.01	.04	.00	.08	.05
<b>Ü1V</b>	<b>-.28</b>	-.19	-.17	-.06	.08	-.08	-.03	-.13	-.09	-.03	-.04	.05	.00	.02	-.02	.12	<b>-.18</b>	-.03
<b>Ü1Vi</b>	<b>.35</b>	.05	<b>.26</b>	-.14	<b>-.22</b>	.19	.01	<b>.34</b>	.19	.12	.06	-.07	-.01	-.09	-.01	-.13	.16	-.03
<b>Ü1Va</b>	-.04	.01	-.05	-.01	.15	.09	.10	-.08	-.03	-.04	-.00	.21	.13	.04	-.06	.17	.08	.10
<b>Ü1K</b>	<b>-.35</b>	-.12	-.19	.05	<b>.29</b>	.09	.14	-.20	-.11	-.07	.10	.17	.17	.07	.04	-.02	.03	.05
<b>Ü1S</b>	<b>.22</b>	.04	.11	-.08	-.07	.17	-.01	.19	.06	.12	.04	.13	.11	-.03	-.02	.01	.19	.07
<b>Ü2A</b>	<b>-.21</b>	-.03	-.17	-.02	<b>.26</b>	<b>.24</b>	.06	-.07	-.16	.11	.11	.17	.18	.01	.09	.10	.06	.11
<b>Ü2V</b>	-.12	.01	.05	.05	.08	.01	.11	-.06	.04	-.10	.03	.02	.03	.01	.03	.07	-.05	.02
<b>Ü2Vi</b>	<b>.30</b>	-.03	.13	-.04	-.14	.09	-.06	.19	.06	.12	.14	-.03	.07	.02	.10	-.05	.06	.05
<b>Ü2Va</b>	-.02	-.05	-.11	-.09	.11	.04	-.04	-.07	-.13	.07	-.04	.12	.05	.01	.01	.04	.11	.07
<b>Ü2K</b>	<b>-.24</b>	-.06	-.08	.02	<b>.21</b>	.15	.13	-.08	-.05	-.03	.08	.14	.14	.06	.14	.18	.04	.17
<b>Ü2S</b>	<b>.25</b>	.04	.05	.00	-.14	.08	-.07	.18	.05	.13	.08	.03	.07	-.05	.06	-.01	.14	.06
<b>ÜbA</b>	<b>-.21</b>	-.05	-.15	-.04	<b>.26</b>	<b>.22</b>	.09	-.08	-.14	.07	.11	.17	.17	-.00	.07	.05	.07	.08
<b>ÜbV</b>	<b>-.24</b>	-.11	-.08	-.00	.10	-.04	.05	-.12	-.04	-.08	-.01	.04	.02	.02	.00	.11	-.14	-.01
<b>ÜbVi</b>	<b>.40</b>	.02	<b>.25</b>	-.11	<b>-.22</b>	.17	-.02	<b>.34</b>	.17	.15	.13	-.07	.04	-.04	.06	-.11	.14	.02
<b>ÜbVa</b>	-.04	-.02	-.09	-.06	.17	.09	.05	-.09	-.09	.01	-.02	<b>.21</b>	.12	.03	-.04	.14	.11	.11
<b>ÜbK</b>	<b>-.34</b>	-.10	-.15	.04	<b>.28</b>	.13	.15	-.16	-.09	-.06	.09	.18	.17	.07	.10	.09	.04	.12
<b>ÜbS</b>	<b>.31</b>	.06	.11	-.05	-.14	.16	-.05	<b>.25</b>	.07	.17	.08	.10	.11	-.05	.03	-.00	<b>.21</b>	.09

**Tabelle 6-36**, Zusammenhänge zwischen Geschichtenvariablen und expliziten sowie impliziten **Einstellungen zu Übergewicht**, Ü1/ 2: erste/ zweite Bildmessung; Üb= Mittelung von Ü1 und Ü2 (grau schattiert), A= Wortanzahl, V= Valenz, Vi= individuelle Vorurteile, Va= antizipierte Vorurteile, K= Kreativität, S= Stereotypenerkennung; weitere Abkürzungen siehe Tab. 6-28, 6-29 und 6-35

Die individuellen Vorurteile in den Übergewichtsgeschichten stehen in positiver Beziehung zu den individuellen kognitiven und affektiven Einstellungen zu Übergewicht (ÜBKOi:  $r=.40$ , ÜBAFi:  $r=.25$ ,  $p<.001$ ), in negativer Beziehung zum Gefühlsthermometer Übergewichtige (ThÜb:  $r=-.22$ ,  $p<.01$ ), in positivem Zusammenhang zum Gefühlsthermometer Normalgewichtige (ThNo:  $r=.17$ ,  $p<.05$ ) und ebenso zur Differenz beider Thermometermaße (Th\_NÜ:  $r=.34$ ,  $p<.001$ ). Jedoch bestehen keinerlei signifikante Korrelationen zwischen den individuellen Vorurteilen der Übergewichtsgeschichten und den impliziten Einstellungsmaßen.

Antizipierte Vorurteile (ÜbVa) in den Geschichten haben kaum Bedeutung für die direkte Messung, hier besteht ein schwach positiver Zusammenhang zur Bewertung Übergewichtiger im Gefühlsthermometer ( $r=.17$ ,  $p<.05$ ). Dagegen korreliert Va positiv und somit in erwartungskonformer Richtung mit den fehlerbasierten Effekten von IAT und EAST2, wobei letzter Zusammenhang jedoch nicht signifikant ist (Va, IATer\_Ü:  $r=.21$ ,  $p<.01$ ; Va, EASTer\_Ü2:  $r=.14$ ,  $p<.09$ ). Die ebenfalls einstellungsrelevante Geschichtenvariable Valenz (ÜbVa) korreliert negativ mit den individuellen kognitiven Einstellungen gemäß Fragebogen ( $r=-.24$ ,  $p<.01$ ), je negativer die Geschichten eingeschätzt wurden, desto höhere Werte haben die Probanden auf den Vorurteilsskalen. Es bestehen keine Zusammenhänge zwischen der Geschichtenvalenz und den impliziten Maßen. Die Erkennung/ Thematisierung des Einstellungsobjekts in den Geschichten (ÜbS) steht in positiver Beziehung insbesondere zu ÜBKOi ( $r=.31$ ,  $p<.001$ ) und der Thermometerdifferenz Normal- vs. Übergewichtige (Th\_NÜ:  $r=.25$ ,  $p<.01$ ). Außerdem erhalten Probanden mit deutlicherer Stereotyperkennung bzw. -Thematisierung höhere Werte im fehlerbasierten EAST2-Effekt ( $r=.21$ ,  $p<.01$ ).

Die beiden Kontrollvariablen Wortanzahl (ÜbA) und Kreativität (ÜbK) unterhalten ähnliche Beziehungen zu den direkten und indirekten Maßen. Beide korrelieren negativ zu ÜBKOi (A:  $r=-.21$ ,  $p<.01$ ; K:  $r=-.34$ ,  $p<.001$ ), positiv zum Gefühlsthermometer Übergewichtige (A:  $r=.26$ , K:  $r=.28$ ,  $p<.001$ ) und ebenfalls positiv zum fehlerbasierten IAT-Effekt (A:  $r=.17$ , K:  $r=.18$ ,  $p<.05$ ), sowie dem standardisierten Gesamtscore für reaktionszeit- und fehlerbasierten IAT-Effekt (A:  $r=.17$ , K:  $r=.18$ ,  $p<.05$ ). Hinsichtlich des negativen Zusammenhanges der Wortanzahl mit den expliziten Vorurteilen, könnte die Geschichtenlänge als ein Ausdruck der Bereitschaft zur Auseinandersetzung mit dem Thema interpretiert werden. Dagegen spricht jedoch der positive Zusammenhang von ÜbA zum IAT-Effekt, bei dem hohe Werte einen Reaktionsvorteil vorurteils*kongruenter* Assoziationen anzeigen sollen.

#### 6.3.4.4 *Einstellungsobjekt Homosexualität*

Zwischen den Geschichtenvariablen und den impliziten Maße für das Einstellungsobjekt Homosexualität bestehen mit 6 signifikanten Korrelationen deutlich weniger Zusammenhänge, als für das Einstellungsobjekt Übergewicht (Anteil signifikanter Ergebnisse: 10%).

Dagegen weisen 9 Korrelationen zwischen der offenen Assoziations- und der indirekten Messung signifikante Beträge auf (Anteil signifikanter Ergebnisse: 18.75%).

	Direkte Messung (Fragebogen)							Indirekte Messung (IAT, EAST)							
	HOKOi	HOKOa	HOAFi	HOAFa	ThHo	ThHe	Th_HH	IATH	IATERH	IATH_gesamt	EASTH1	EASTH2	EASTerH1	EASTerH2	EASTH_gesamt
<b>H1A</b>	-.02	.01	-.04	.07	.18	<b>.24</b>	.02	.04	-.05	-.01	.16	.01	.06	.00	.10
<b>H1V</b>	.04	-.02	-.04	-.00	-.06	-.04	.03	.14	.08	.13	-.01	-.02	-.14	-.01	-.08
<b>H1Vi</b>	<b>.23</b>	.05	.16	.01	<b>-.26</b>	-.04	<b>.25</b>	-.03	-.09	-.07	<b>-.16</b>	-.09	.04	-.05	-.12
<b>H1Va</b>	-.04	.05	-.04	.05	.04	.08	.03	.00	.04	.02	.02	-.03	.06	-.11	-.03
<b>H1K</b>	-.00	.08	-.03	.08	.06	.12	.04	.05	-.06	-.01	.13	.10	<b>.17</b>	-.01	<b>.17</b>
<b>H1S</b>	.04	-.09	-.06	-.05	.04	.06	.01	-.15	-.14	<b>-.18</b>	-.15	-.08	-.08	.05	-.12
<b>H2A</b>	-.09	-.11	-.09	-.07	.19	.17	-.05	.03	.03	.04	<b>.22</b>	-.05	.04	-.03	.08
<b>H2V</b>	-.10	-.05	-.03	-.06	.07	-.05	-.13	.03	.11	.09	-.13	.10	.06	.15	.07
<b>H2Vi</b>	.14	.14	.16	.08	<b>-.22</b>	-.01	<b>.24</b>	-.06	-.07	-.08	-.13	-.07	-.13	<b>-.17</b>	<b>-.21</b>
<b>H2Va</b>	-.08	-.01	.01	.03	.07	.06	-.02	-.04	.06	.01	.10	-.04	.11	-.10	.03
<b>H2K</b>	-.00	.02	-.02	.05	.10	.11	-.02	.17	.08	.15	.19	.11	.19	.04	<b>.23</b>
<b>H2S</b>	<b>-.21</b>	.00	-.12	.04	.01	-.19	-.18	-.19	-.05	-.15	-.12	-.06	.01	-.09	-.11
<b>HoA</b>	-.07	-.06	-.09	.00	<b>.22</b>	<b>.24</b>	-.02	.04	-.01	.01	<b>.21</b>	-.02	.06	-.01	.11
<b>HoV</b>	-.06	-.05	-.05	-.05	.02	-.06	-.08	.10	.13	.14	-.10	.06	-.03	.11	.02
<b>HoVi</b>	<b>.21</b>	.12	.20	.06	<b>-.29</b>	-.03	<b>.29</b>	-.06	-.09	-.09	-.17	-.09	-.06	-.14	<b>-.20</b>
<b>HoVa</b>	-.08	.01	-.01	.04	.07	.08	-.01	-.03	.06	.02	.09	-.05	.12	-.12	.02
<b>HoK</b>	-.01	.05	-.03	.08	.10	.13	.01	.12	.01	.08	.18	.12	<b>.20</b>	.02	<b>.23</b>
<b>HoS</b>	-.13	-.06	-.12	-.01	.03	-.11	-.13	<b>-.25</b>	-.14	<b>-.24</b>	-.20	-.10	-.05	-.03	-.17

**Tabelle 6-37**, Zusammenhänge zwischen Geschichtenvariablen und expliziten sowie impliziten **Einstellungen zu Homosexualität**, H1/ 2: erste/ zweite Bildmessung; Ho= Mittelung von H1 und H2 (grau schattiert), Abkürzungen siehe Tab. 6-28, 6-29, 6-35 und 6-36

Auch in den Homosexualitätsgeschichten weist der individuelle Vorurteilsgehalt (HoVi) die größte Bedeutung für die Fragebogenmaße auf. Je mehr Vorurteile beide Geschichten eines Probanden widerspiegeln, desto negativer sind die individuellen kognitiven und affektiven Einstellungen zu Homosexualität (HOKOi:  $r=.21$ ,  $p<.01$ ; HOAFi:  $r=.20$ ,  $p<.05$ ), als desto unangenehmer werden homosexuelle Personen empfunden (ThHo:  $r=-.29$ ,  $p<.001$ ) und desto größer ist die Thermometerdifferenz hetero- vs. homosexuelle Menschen (Th\_HH:  $r=.29$ ,  $p<.001$ ). Außerdem korreliert HoVi negativ zum Reaktionszeiteffekt des EAST1 ( $r=-.17$ ,  $p<.05$ ) sowie zum Gesamteffekt des EAST ( $r=-.20$ ,  $p<.01$ ). Diese negative Korrelation spiegelt einen hypothesenkonträren Zusammenhang zwischen beiden Einstellungsmaßen wider. Die Geschichtenlänge (HoA) steht, wie bereits für Übergewicht beobachtet wurde, in positiver Beziehung zu beiden Gefühlsthermometern (ThHo:  $r=.22$ , ThHe:  $r=.24$ ,  $p<.01$ ).



Außerdem korreliert HoA positiv mit den Reaktionszeiten des ersten EAST-Blocks ( $r=.21$ ,  $p<.01$ ). Dieser positive Zusammenhang repliziert den Befund positiver Korrelationen zwischen Wortanzahl und impliziten Assoziationen innerhalb des Einstellungsobjekts Übergewicht.

Die eingeschätzte Kreativität weist lediglich für die indirekten Maße statistische Bedeutsamkeit auf. Kreativere Geschichtenschreiber weisen höhere EAST-Effekte auf, und zwar speziell für Reaktionszeiten und Fehler des ersten Blocks (EASTH1:  $r=.18$ ,  $p<.05$ ; EATerH1:  $r=.20$ ,  $p<.01$ ), jedoch auch für den Gesamteffekt des EAST ( $r=.23$ ,  $p<.01$ ). Auch in diesen Zusammenhängen zeigt sich eine gewisse Stabilität der Beziehungsmuster zwischen beiden Einstellungsobjekten.

Dagegen hat die Stereotyperkennung bzw. -Thematisierung für Einstellungen zu Homosexualität eine andere Bedeutung als für Einstellungen zu Übergewicht. Die Stereotyperkennung steht zum Reaktionszeit-IAT, dem Gesamteffekt des IAT, dem reaktionszeitbasierten EAST-Effekt des ersten Blocks und ebenfalls dem Gesamteffekt des EAST in jeweils *negativer* Beziehung (IATH:  $r=-.25$ , IATH *gesamt*:  $-.24$ ,  $p<.01$ ; EASTH1:  $r=-.20$ , EAST *gesamt*:  $r=-.17$ ,  $p<.05$ ). Je deutlicher also der Stereotyp erkannt/benannt wird, desto geringer sind die impliziten Assoziationsdifferenzen für homo- und heterosexuelle Stimuli.

#### 6.3.4.5 Gesamtbetrachtung

Zum Überblick wird an dieser Stelle eine gemeinsame Betrachtung der Zusammenhänge zwischen den relevanten Einstellungsmaßen pro Einstellungsobjekt ermöglicht, womit zugleich die allgemeine Multitrait-Multimethod-Matrix abgebildet ist (Tab. 6-38, mehr dazu im nächsten Abschnitt).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. <b>ÜBKO<i>i</i></b>															
2. <b>ÜBAF<i>i</i></b>	<b>.46</b>														
3. <b>TH_NÜ</b>	<b>.48</b>	<b>.43</b>													
4. <b>ÜbV<i>i</i></b>	<b>.40</b>	<b>.25</b>	<b>.33</b>												
5. <b>IATÜ</b>	.06	.10	.04	.13											
6. <b>IATerÜ</b>	-.05	-.09	-.10	-.07	<b>.27</b>										
7. <b>EASTÜ</b>	.04	.01	.05	.01	.17	.15									
8. <b>EASTerÜ</b>	.05	.05	-.03	.02	.08	.16	.22								
9. <b>HOKO<i>i</i></b>	<b>.29</b>	<b>.23</b>	<b>.27</b>	.00	.08	.04	-.01	-.09							
10. <b>HOAF<i>i</i></b>	<b>.21</b>	<b>.47</b>	<b>.20</b>	.04	-.01	.05	.05	-.02	<b>.64</b>						
11. <b>TH_HH</b>	<b>.24</b>	<b>.24</b>	<b>.47</b>	.14	.04	.13	-.05	-.05	<b>.61</b>	<b>.52</b>					
12. <b>HoV<i>i</i></b>	<b>.27</b>	.18	<b>.24</b>	<b>.29</b>	.06	-.06	-.02	.01	<b>.21</b>	<b>.20</b>	<b>.29</b>				
13. <b>IATH</b>	.06	.11	.09	.01	<b>.35</b>	<b>.25</b>	.07	.05	<b>.30</b>	<b>.34</b>	<b>.29</b>	-.06			
14. <b>IATerH</b>	.04	-.01	-.04	-.02	.06	.14	.12	.02	-.08	.04	-.09	-.09	<b>.34</b>		
15. <b>EASTH</b>	<b>-.21</b>	-.15	-.08	-.07	-.11	.05	.05	-.05	.06	.02	-.04	-.18	.02	.02	
16. <b>EASTerH</b>	-.09	-.12	-.09	-.03	-.13	.06	-.04	-.01	.09	-.01	-.03	-.14	.12	-.09	<b>.21</b>

**Tabelle 6-38**, Korrelationen der relevanten Einstellungsmaße, Abkürzungen siehe Tab. 6-28, 6-29, 6-35 und 6-36

Neben den bereits berichteten Beziehungen insbesondere zwischen expliziten Einstellungen und den Geschichtenvariablen *innerhalb* eines Einstellungsobjekts (Monotrait-Heteromethod-Korrelationen), bestehen hierfür ebenfalls substantielle Zusammenhänge *zwischen* den Einstellungsobjekten (Heterotrait-Heteromethod-Korrelationen). Diese liegen betragsmäßig leicht unter den Monotrait-Heteromethod-Korrelationen und zeigen die Existenz von Methodeneffekte an.

## 6.4 Ergebnisse II: Latente Multitrait-Multimethod-Analyse

Das zentrale Forschungsziel vorliegender Arbeit ist die messfehlerbereinigte, konstrukt- und methodenspezifische Untersuchung der Einstellungsmaße im Multitrait-Multimethod-Design mit konfirmatorischer Faktorenanalyse (CFA). Dabei ist neben der konvergenten und diskriminanten Validität der Einstellungskonstrukte die Bedeutung der Methodenfaktoren von großem Interesse. Es liegen Einstellungsmaße gemäß folgender vier Messmethoden vor:

- Direkte Einstellungsmessung via Selbstbericht (kognitiv, affektiv, Gefühlsthermometer)
- offene Assoziativmessung (z.B. individueller Vorurteilsgehalt)
- Indirekte Einstellungsmessung via IAT (Reaktionszeiten- und Fehlereffekte)
- Indirekte Einstellungsmessung via EAST (Reaktionszeiten- und Fehlereffekte)

### 6.4.1 Voraussetzungen und Evaluation latenter Strukturgleichungsmodelle

Die Analyse von Strukturgleichungsmodellen (SEM) erfolgt über Schätzalgorithmen, die in aller Regel die multivariate Normalverteilung der einbezogenen Variablen verlangen, eine Voraussetzung, die einerseits bei sozial- und verhaltenswissenschaftlichen Daten im allgemeinen selten gegeben (Micceri, 1989) und andererseits schwer zu prüfen ist. Voraussetzung für die multivariate Normalverteilung ist zumindest die univariate Normalverteilung. Wie Tabelle 6-39 entnommen werden kann, liegen jedoch für die Mehrzahl der Indikatoren Voraussetzungsverletzungen vor. Der Kolmogorov-Smirnov-Test weist für 21 von insgesamt 30 relevanten Indikatoren signifikante Abweichung von der Normalverteilung nach, wobei keine deutliche Häufung der Normalverteilungsverletzung innerhalb einer Methode besteht. Die Verteilungen der Variablen sind in vielen Fällen deutlich links- bzw. rechtssteil. Außerdem sind die Standardabweichungen im Verhältnis zum Mittelwert einiger Variablen sehr groß, was insbesondere die Fehlereffekte des EAST betrifft.

		<b>N</b>	<b>Min</b>	<b>Max</b>	<b>M (SD)</b>	<b>KS-Z</b>	<b>p</b>	<b>Schiefe</b>	<b>SE</b>	<b>Kurtosis</b>	<b>SE</b>
Selbstbericht	<b>ÜBKO<sub>i</sub>_A</b>	162	1.70	4.50	2.88 (.59)	1.03	.241	.46	.191	.02	.379
	<b>ÜBKO<sub>i</sub>_B</b>	162	1.11	3.78	2.38 (.56)	.75	.621	.13	.191	-.29	.379
	<b>ÜBAF<sub>i</sub>_A</b>	161	1.20	4.40	2.96 (.48)	2.12	.000	-.82	.191	1.79	.380
	<b>ÜBAF<sub>i</sub>_B</b>	161	1.20	4.20	2.83 (.60)	1.18	.124	-.21	.191	-.10	.380
	<b>Th_NÜ</b>	162	-8.00	75.00	16.38 (17.06)	2.08	.000	1.17	.191	.99	.379
	<b>HOKO<sub>i</sub>_A</b>	162	1.44	3.56	2.13 (.39)	1.88	.002	.82	.191	1.16	.379
	<b>HOKO<sub>i</sub>_A</b>	162	1.00	3.78	1.97 (.52)	1.58	.013	.95	.191	1.00	.379
	<b>HOAF<sub>i</sub>_A</b>	161	1.00	4.60	2.73 (.72)	1.68	.007	-.46	.191	-.06	.380
	<b>HOAF<sub>i</sub>_A</b>	161	1.00	4.75	2.75 (.76)	1.83	.002	-.29	.191	.20	.380
	<b>Th_HH</b>	162	-30.00	75.00	7.75 (13.98)	3.39	.000	1.92	.191	5.53	.379
off. Assoz.	<b>Ü1V<sub>i</sub></b>	161	1.00	5.00	2.27 (1.10)	1.73	.005	.48	.191	-.87	.380
	<b>Ü2V<sub>i</sub></b>	160	1.00	4.67	1.98 (.95)	2.07	.000	.79	.192	-.19	.381
	<b>H1V<sub>i</sub></b>	161	1.00	3.33	1.11 (.39)	6.29	.000	4.19	.191	18.25	.380
	<b>H2V<sub>i</sub></b>	160	1.00	5.00	1.20 (.48)	4.98	.000	4.44	.192	26.53	.381
IAT	<b>IATÜ1</b>	160	-.15	1.08	.41 (.24)	.80	.537	.14	.192	-.593	.381
	<b>IATÜ2</b>	160	-.07	1.20	.20 (.20)	2.04	.001	1.54	.192	3.249	.381
	<b>IATerÜ1</b>	160	-20.83	50.00	9.51 (11.63)	1.64	.009	1.00	.192	2.23	.381
	<b>IATerÜ2</b>	160	-6.25	20.83	4.30 (5.27)	1.66	.008	.58	.192	.26	.381
	<b>IATH1</b>	158	-.18	.75	.15 (.14)	.77	.600	.79	.193	2.06	.384
	<b>IATH2</b>	158	-.13	.48	.07 (.09)	1.43	.030	1.41	.193	4.68	.384
	<b>IATerH1</b>	158	-12.50	25.00	2.43 (5.76)	2.29	.000	.45	.193	1.61	.384
	<b>IATerH2</b>	158	-6.25	16.67	1.15 (3.42)	2.22	.000	.99	.193	3.27	.384
EAST	<b>EASTÜ1</b>	159	-.36	.38	-.01 (.10)	.60	.862	.03	.192	1.37	.383
	<b>EASTÜ2</b>	159	-.19	.29	.03 (.10)	.80	.551	.43	.192	-.07	.383
	<b>EASTERÜ1</b>	159	-55.94	21.68	-6.44 (12.65)	1.55	.017	-.69	.192	.87	.383
	<b>EASTERÜ2</b>	159	-33.33	66.67	4.35 (13.45)	1.96	.001	.52	.192	2.45	.383
	<b>EASTH1</b>	160	-.32	.19	-.01 (.09)	.592	.874	-.37	.192	.79	.381
	<b>EASTH2</b>	160	-.23	.23	.02 (.09)	.426	.993	-.06	.192	-.12	.381
	<b>EASTERH1</b>	160	-58.33	33.33	-2.19 (12.35)	2.47	.000	-.86	.192	2.93	.381
	<b>EASTERH2</b>	160	-41.67	41.67	3.85 (12.81)	2.15	.000	.09	.192	1.14	.381

**Tabelle 6-39.** Voraussetzungsprüfung für Strukturgleichungsmodellierung der (individuellen) Einstellungsmaße; <sub>A/</sub><sub>B</sub>= Split-Half-Indikatoren der Fragebogenskalen; KS-Z: Kolmogorov-Smirnov-Test auf Normalverteilung, Schiefe und Kurtosis bei normal verteilten Daten= 0; Abweichungen von (+/-) 2\*SE gelten als signifikant; N kompletter Datensätze= 146; Abkürzungen siehe Tab. 6-28, 6-29, 6-35 und 6-36

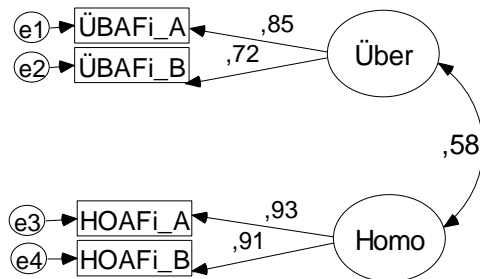
Verletzungen der Normalverteilungsannahme können zu instabilen  $\chi^2$ -Statistiken führen und somit zu fehlerhaften Evaluationen der globalen Modellgüte. Die Maximum-Likelihood-Prozedur wird jedoch als relativ robust gegen Verletzungen der Normalverteilungsannahme eingeschätzt (McDonald & Ho, 2002). Zudem resultieren aus nicht gegebener Normalverteilung *überschätzte*  $\chi^2$ -Werte (Hoyler, 1995), die zu einer fälschlichen Ablehnung des Modells führen. Da in der Strukturgleichungsmodellierung die  $H_0$  (keine signifikanten Abweichungen zwischen empirischer und modellimplizierter Kovarianzstruktur) die Wunschhypothese ist, handelt es sich demnach um konservativere Schätzungen. Wenn auch im Falle nicht normal verteilter Indikatoren für die Interpretation eines einzelnen Modells Vorsicht geboten ist, bleibt ein Vergleich ineinander genesteter Modelle auf der Basis der  $\chi^2$ -Statistik aufgrund gleicher Datengrundlage durchführbar. Zur Evaluation der globalen Güte eines einzelnen Modells kann zudem der „Root Mean Square Error of Approximation“ (RMSEA, Browne & Cudeck, 1993) herangezogen werden, der einen auch bei nicht normal verteilten Daten robusten Fit-Index darstellt. Der RMSEA gibt die Abweichung des zu testenden Modells vom saturierten Modell mit  $df=0$  an und sollte unter .10 liegen. Ein RMSEA von unter .08 wird als *guter* Fit interpretiert, ein RMSEA unter .05 indiziert ein *sehr gutes* Modell (McDonald & Ho, 2002; Tabachnik & Fidell, 2001).

Der Umfang der Stichprobe für robuste Schätzergebnisse in SEM wird mit ca.  $N=200$  Untersuchungseinheiten angegeben (siehe Ullmann, 2001). Die vorliegende Stichprobe ist mit  $N=165$  unvollständigen ( $N_{\text{valide}}=146$ ) somit relativ klein, was zu einer fehlerhaften  $p$ -Wert Berechnung führen kann. Neben dem RMSEA kann ein ebenfalls häufig betrachtetes Gütekriterium – der „Normed Fit Index“ (NFI, Bentler & Bonett, 1980) – in kleinen Stichproben zu einer konservativen Modellevaluation und somit zum fälschlichen Verwerfen des Modells führen (Tabachnik & Fidell, 2001; der NFI resultiert aus einem  $\chi^2$ -Verhältnis zwischen zu testendem und unabhängigem Modell, dessen Wertebereich zwischen 0 und 1 liegt, wobei Werte größer als .90 einen guten Modell-Fit indizieren). Bei der Evaluation eines Modells sollte deshalb insbesondere das Verhältnis von  $\chi^2$ -Wert und Freiheitsgraden ( $df$ ) betrachtet werden. Hier indiziert ein  $\chi^2/df$ -Verhältnis von weniger als 2 einen guten Modell-Fit. Die  $\chi^2$ -Statistik, der RMSEA und der NFI werden als adäquat und ausreichend zur Gesamtevaluation der Modelle betrachtet. Neben den globalen Fit Indizes sollten die vom Modell implizierten Parameter auf Plausibilität geprüft werden. Zusätzlich empfehlen Bagozzi & Yi (1999) insbesondere auch für Multitrait-Multimethod-Modelle die Berücksichtigung der geschätzten Fehlervarianzen, wobei hohe Werte einen relativ schlechten Modell-Fit indizieren.

Um die geeigneten Indikatoren für die latente MTMM-Analyse auszuwählen, werden im folgenden zunächst die Messmodelle für die Modellierung zweier Einstellungstrait (Einstellungen zu Übergewicht = „Über“ und Einstellungen zu Homosexualität = „Homo“) innerhalb der vier Messmethoden geprüft. Die Strukturgleichungsmodellierung in aktueller Arbeit erfolgt in AMOS 5.0 (Arbuckle, 2003).

### 6.4.2 Entscheidung über die Messmodelle

Für die direkte Einstellungsmessung (Selbstbericht) werden drei verschiedene Modelle getestet. Zu-



**Abbildung 6-31**, Beispiel für Messmodell, hier: „Direkt affektiv“

nächst werden alle Einstellungsmaße der Fragebogenerhebung, also die individuellen kognitiven (ÜBKO $_i$ , HOKO $_i$ ) und affektiven Einstellungen (ÜBAFi, HOAF $_i$ ) sowie die Gefühlsthermometerdifferenz (Th\_NÜ, Th\_HH) als Indikatoren betrachtet (Modell „*Direkt gesamt*“). Des Weiteren werden für kognitive und affektive Einstellungen getrennte Messmodelle berechnet, wofür die Skalen in gleiche Hälften geteilt werden, um die latenten Einstellungsfaktoren schätzen zu können (Modell „*Direkt kognitiv*“ bzw. „*Direkt affektiv*“, siehe Abb. 6-31).

Für die offene Assoziativmessung werden ebenfalls drei Messmodelle getestet. In Modell „*Offen gesamt*“ stellen alle einstellungsrelevanten Variablen, nämlich Valenz (z.B. Ü1V, H1V), individuelle Vorurteile (z.B. Ü1Vi, H1Vi), Kreativität (z.B. Ü1K, H1K) und Stereotyperkennung (z.B. Ü1S, H1S) die Indikatoren zur Modellierung der beiden latenten Einstellungsfaktoren dar, wobei Valenz und Kreativität zu diesem Zwecke umkodiert werden, so dass hohe Werte negative Valenz bzw. Unkreativität anzeigen. Darüber hinaus wird ein Modell mit Valenz und Vorurteilen gemäß der vier Geschichten als Indikatoren der Einstellungsfaktoren geprüft („*Offen Valenz & Vorurteile*“). Schließlich testet ein Modell die Modellierung der Einstellungstraits über die beiden Vorurteilsvariablen je Einstellungsobjekt als alleinige Indikatoren („*Offen Vorurteile*“). Jeweils drei Messmodelle werden getrennt für IAT und EAST-Daten berechnet. Je ein Modell prüft Reaktionszeit- und Fehlereffekt gemeinsam als Indikatoren („*IAT gesamt*“, „*EAST gesamt*“), die anderen beiden Modelle testen Reaktionszeit- und Fehlereffekte gesondert („*IAT Reaktionszeiten*“, „*IAT Fehler*“ usw.).

Entscheidungskriterien für geeignete Indikatoren zur Schätzung der latenten Einstellungsfaktoren sind ein angemessener Modell-Fit und eine sinnvolle Parameterlösung. Auf eine ausführliche Dokumentation der einzelnen Analysen wird an dieser Stelle zugunsten einer übersichtlichen Ergebnisdarstellung verzichtet (siehe Tab. 6-40).

Mess-ebene	Modelle	$\chi^2$	df	$\chi^2/df$	p	RMSEA	NFI	$r_{\text{Über, Homo}}$
direkt	<i>Direkt gesamt</i>	62.78	8	7.85	.000	.204	.822	.54***
	<i>Direkt kognitiv</i>	4.35	3	1.45	.226	.052	.986	.36***
	<i>Direkt affektiv</i>	3.76	3	1.25	.286	.039	.988	.58***
offen	<i>Offen gesamt</i>	640.64	109	5.88	.000	.172	.253	.81***
	<i>Offen Valenz &amp; Vorurteile</i>	93.91	25	3.76	.000	.129	.794	.89***
	<i>Offen Vorurteile</i>	1.39	1	1.39	.238	.049	.997	.56***
indirekt IAT	<i>IAT gesamt</i>	115.80	25	4.63	.000	.148	.533	.49***
	<i>IAT Reaktionszeiten</i>	16.19	3	5.40	.001	.163	.887	.49***
	<i>IAT Fehlerraten</i>	4.93	3	1.64	.177	.062	.834	.31 n.s.
indirekt EAST	<i>EAST gesamt</i>	51.69	25	2.07	.001	.080	.144	.18 n.s.
	<i>EAST Reaktionszeiten</i>	12.22	3	4.07	.007	.136	.408	.19 n.s.
	<i>EAST Fehlerraten</i>	2.71	3	.90	.438	.000	.549	-.11 n.s.

**Tabelle 6-40,** Messmodelle zur Auswahl der MTMM-Indikatoren,  $r_{\text{Über, Homo}}$ : Korrelation der latenten Einstellungsfaktoren

Da die Messmodelle „*Direkt kognitiv*“ und „*Direkt affektiv*“ vergleichbar gute Passung an die beobachteten Daten aufweisen, für die *kognitiven* Einstellungen jedoch konsistentere Reliabilitäten und Zusammenhänge zu den weiteren Einstellungsmaßen bestehen, werden die kognitiven Einstellungen (Messmodell „*Direkt kognitiv*“) als Indikatoren der direkten Einstellungsmessung ausgewählt.<sup>2</sup> Die Entscheidung über die geeigneten Indikatoren für die offene Assoziativmessung der Geschichten fällt aufgrund der hervorragenden Fit-Indizes von Messmodell „*Offen Vorurteile*“ relativ leicht. Als Indikatoren für die indirekte Messung müssen gemäß den Messmodellen sowohl für IAT, als auch für EAST, die Fehlereffekte ausgewählt werden (Messmodelle „*IAT Fehlerraten*“ und „*EAST Fehlerraten*“). Für die impliziten Maße sollte auch zu Vergleichbarkeitszwecken ein Datenformat betrachtet werden. Ein auffälliges Ergebnis ist die deutlich geringere Konvergenz von Reaktionszeit- und Fehlereffekten des IAT im Vergleich zum EAST (siehe „IAT gesamt“ vs. „EAST gesamt“). Reaktionszeit- und Fehlerbasierte IAT-Maße legen demnach keine gemeinsame Modellierung nahe. Für den EAST ist nach wie vor unklar, ob vorrangig die Reaktionszeit- oder die Fehlereffekte die kritische Messung darstellen (vgl. de Jong & Huijding, 2005; Ellwart, Rinck & Becker, 2005; Huijding, 2005; Perugini, O’Gorman & Connor, 2005), wobei das konventionelle IAT-Maß auf Reaktionszeiten basiert.

Für die latente MTMM-Analyse werden also Split-Half-Indikatoren direkt erfasster kognitiver Einstellungen, die Vorurteilsvariablen der offenen Assoziativmessung sowie die Fehlereffekte gemäß IAT und EAST ausgewählt. Die beabsichtigte latente MTMM-Analyse dient der separaten Betrachtung

<sup>2</sup> Die Gefühlsthermometermessung je Einstellungsobjekt eignet sich für sich genommen nicht zur latenten Modellierung, da es sich um ein nicht teilbares Differenzmaß handelt. Für die MTMM-Analyse werden Th\_NÜ und Th\_H-H daher nicht weiter berücksichtigt.

trait- und methodenspezifischer Varianz der messfehlerbereinigten Einstellungsmaße. Traitspezifische Varianz sollte entsprechend der beiden Einstellungsobjekte Übergewicht und Homosexualität in zwei latenten Einstellungsfaktoren resultieren. Prinzipiell kann für jede der eingesetzten Messmethoden (Fragebogen, offene Assoziation, IAT und EAST) methodenspezifische Varianz isoliert werden, was bei hoher eigener Varianz der Methode zu vier Methodenfaktoren führen sollte. Es ist fraglich, inwiefern tatsächlich N Methodenfaktoren modelliert werden können. Da IAT und EAST gleichsam implizite Assoziation erfassen sollten, wäre hier beispielsweise die Konvergenz in einem Methodenfaktor zu erwarten.

Im Folgenden wird die Multitrait-Multimethod-Matrix der ausgewählten Einstellungsmaße abgebildet, die die Grundlage des informativen Ansatzes MTMM-Analyse nach Campbell & Fiske (1959) darstellt (Tab. 6-41). Auch wenn bereits an dieser Stelle die Bedeutung von Methodenvarianz erkennbar wird, ermöglicht erst der Ansatz konfirmatorischer Faktorenanalyse (CFA) zur latenten MTMM-Analyse Schätzungen für die einzelnen Varianzanteile (Trait vs. Methoden) sowie die Überprüfung theoriegeleiteter Modellmodifikation.

Da die Reaktionszeiten das konventionell analysierte Datenformat impliziter Maße, insbesondere für den IAT, sind, werden die Reaktionszeiteffekte von IAT und EAST mit dokumentiert. Die Bedeutung dieser Maße soll auch in vorliegender Arbeit nicht ungeprüft bleiben.

Methode	Trait	direkt		offen			iat			east		
		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
direkt	1. <b>ÜBKO<sub>i</sub></b>											
	2. <b>HOKO<sub>i</sub></b>	<b>.29</b>										
offen	3. <b>ÜbV<sub>i</sub></b>	<b>.40</b>	-.00									
	4. <b>HoV<sub>i</sub></b>	<b>.27</b>	.21	<b>.29</b>								
iat (rt)	5. <b>IATÜ</b>	.06	.08	.13	.06							
	6. <b>IATH</b>	.06	<b>.30</b>	.01	-.06	<b>.35</b>						
iat (er)	7. <b>IATerÜ</b>	-.05	.04	-.07	-.06	<b>.27</b>	<b>.25</b>					
	8. <b>IATerH</b>	.04	-.08	-.02	-.09	.06	.14	<b>.34</b>				
east (rt)	9. <b>EASTÜ</b>	.04	-.01	.01	-.02	.17	.15	.07	.12			
	10. <b>EASTH</b>	-.21	.06	-.07	-.18	-.11	.05	.02	.02	.05		
east (er)	11. <b>EASTerÜ</b>	.05	-.09	.02	.01	<b>.08</b>	.16	.05	.02	.22	-.05	
	12. <b>EASTerH</b>	-.09	.09	-.03	-.14	-.13	.06	.12	-.09	-.04	-.01	.21

**Tabelle 6-41.** MTMM-Matrix der relevanten Einstellungsmaße; (rt): Reaktionszeiten, (er): Fehlerraten

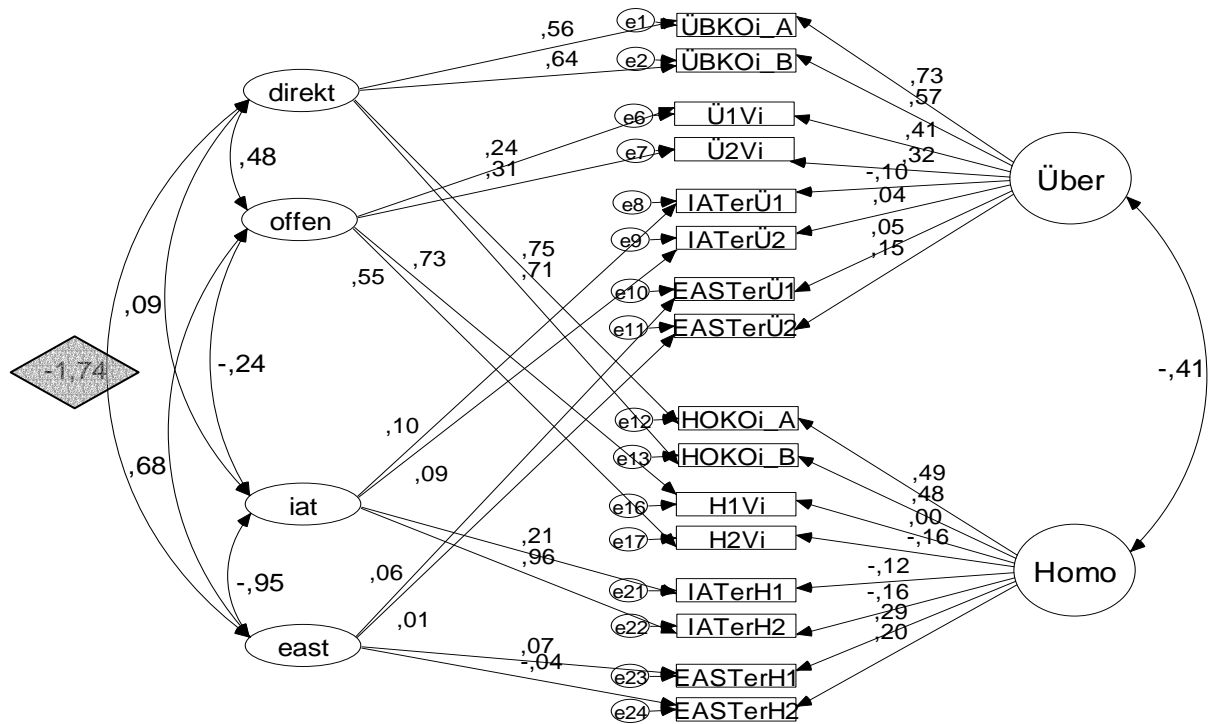
### 6.4.3 Grundlagen der latenten MTMM-Analyse

Prinzipiell sind verschiedene Strategien zur Konstruktvalidierung über konfirmatorische Faktorenanalyse anwendbar. Bedeutsam in einer latenten MTMM-Analyse ist neben dem globalen Vergleich verschiedener Modellierungen die Betrachtung der Parameter des endgültigen Modells. Notwendig für konvergente Validität der interessierenden Konstrukte, hier *Einstellungen zu Übergewicht* und *Homosexualität*, sind signifikante Ladungen der latenten Trait-Faktoren auf die jeweiligen Indikatoren, die in ihrem Betrag die Ladungen der Methodenfaktoren übersteigen sollten. Die Produkte dieser Ladungen entsprechen den Schätzungen der konvergenten Validitätskoeffizienten (z.B. Pedhazur & Schmelkin, 1991). Als diskriminant valide werden die Konstrukte bezeichnet, wenn die Korrelation der latenten Trait-Faktoren signifikant verschieden von 1 ist, bzw. die Konvergenz in einem gemeinsamen Faktor keine signifikante Modellverbesserung herbeiführt (Schmitt & Stults, 1986; Widaman, 1985). Cote, Buckley & Best (2001) raten zusätzlich zur Betrachtung des Anteils erklärter Varianz an der Gesamtvarianz durch die Trait- und Methodenfaktoren. Diese sollte größer sein, als die nicht erklärte Varianz, die sich in den Fehlertermen niederschlägt.

Die Systematik hierarchisch genesteter Strukturgleichungsmodelle von Widaman (1985) ermöglicht durch sukzessive Vergleiche ineinander genesteter Modelle die Validierung von Trait- und Methodenfaktoren. Dabei wird der Nutzen einer Einbuße von Freiheitsgraden aufgrund der Modellierung mehrerer Trait- bzw. Methodenfaktoren getestet. Ist der Verlust an Freiheitsgraden durch die Einführung von Modellkomponenten im Sinne der Konstruktvalidität gerechtfertigt, so können die modellierten Faktoren als konvergent valide bezeichnet werden. Der entsprechende Modellvergleich erfolgt über die Differenz der  $\chi^2$ -Werte zweier ineinander genesteter Modelle. Problematisch am CFA-Ansatz zur MTMM-Analyse ist jedoch die *Unteridentifizierung* des vollständigen Modells mit  $m+t$  Informationen (Kenny & Kashy, 1992). Aus diesem Grund treten bei der Modellierung des kompletten Modells akute Identifikations- und Schätzbarkeitsprobleme auf. Die Modelle sind entweder grundsätzlich nicht identifiziert oder sie weisen unplausible Parameterschätzungen wie negative Varianzen oder standardisierte Pfadkoeffizienten größer als 1 – so genannte „Heywood Cases“ – auf (Kenny & Kashy, 1992; Marsh, 1989; Grayson & Marsh, 1994). Die Wahrscheinlichkeit für das Auftreten solcher Heywood cases ist in kleinen Stichproben ( $N < 400$ ) erhöht (Marsh & Bailey, 1991; Lance, Noble & Scullen, 2002).

Auch im vorliegenden Fall wird für das vollständige Modell korrelierter Methoden- und Traitfaktoren „Out-of-Range“-Parameterschätzungen erhalten (Abb. 6-32), die keine gültige Interpretation des Modells erlaubt. Der Vergleich der CT-CM Modelle nach Widaman (1985) ist demzufolge nach hier nicht geeignet zur Konstruktvalidierung der Einstellungsmaße. Die entsprechende Datendokumentation ist Anhang A.6.4 zu entnehmen.





**Abbildung 6-32,** CT-CM-Modell 3C nach Widaman (1985), Rot markiert: Out-of-Range-Korrelation unter Methodenfaktoren; „Über“: Einstellungen zu Übergewicht, „Homo“: Einstellungen zu Homosexualität“;  
Modell-Fit:  $\chi^2=110.38$ ,  $df=85$ ,  $p=.03$ ,  $\chi^2/df=1.3$ ,  $RMSEA=.043$ ,  $NFI=.804$

#### 6.4.4 CT-C(M-1) Modellierungen

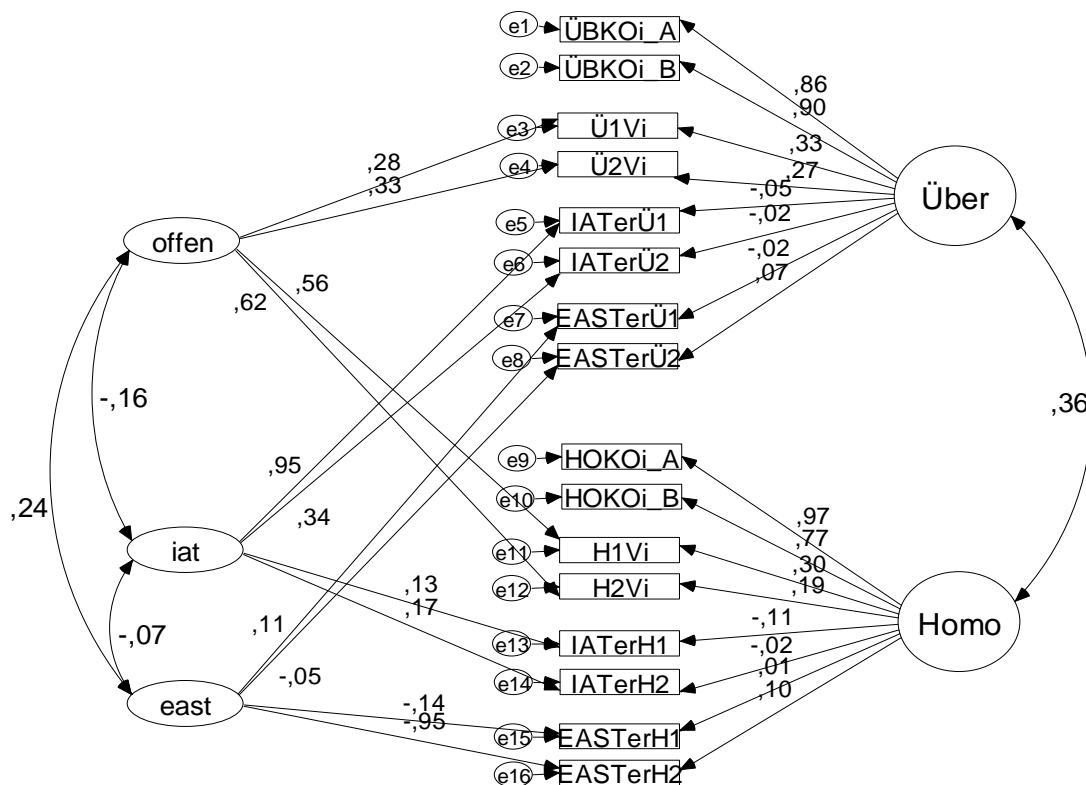
##### 6.4.4.1 Grundlagen

Aufgrund der Identifikationsprobleme der CT-CM Modelle, die auf Überfaktorisierung zurückgeführt werden, schlug Eid (2000) ein Modell mit  $m-1$  Methodenfaktoren vor (CT-C(M-1) Modell), das von Eid, Lischetzke, Nussbeck & Trierweiler (2003) erweitert wurde, um nunmehr auch traitspezifische Methodenvarianz betrachten zu können. Das „Multiple-Indicator CT-C(M-1) Modell“ unterscheidet sich vom ursprünglichen CT-C(M-1) Modell also lediglich in der Modellierung von  $m-1$  Methodenfaktoren für jeden erfassten Trait. Allgemein stellt im CT-C(M-1) Modell die nicht als Methodenfaktor modellierte Messmethode den Vergleichsstandard (Referenzmethode) dar. Dieser muss notwendigerweise im Vorfeld gewählt werden. Im vorliegenden Fall wird der direkte Selbstbericht via Fragebogen als Vergleichsstandard festgelegt. Das bedeutet eine Kontrastierung der direkten Messmethode mit der offenen Assoziation („offen“), dem IAT („iat“) und dem EAST-Verfahren („east“), um die Abweichungen der Assoziativmessung und der beiden indirekten Messmethoden von der direkten Messung betrachten zu können. Das CT-C(M-1) Modell ist nicht symmetrisch; die Bedeutung der Modellparameter hängt von der als Vergleichsstandard gewählten Methode ab.

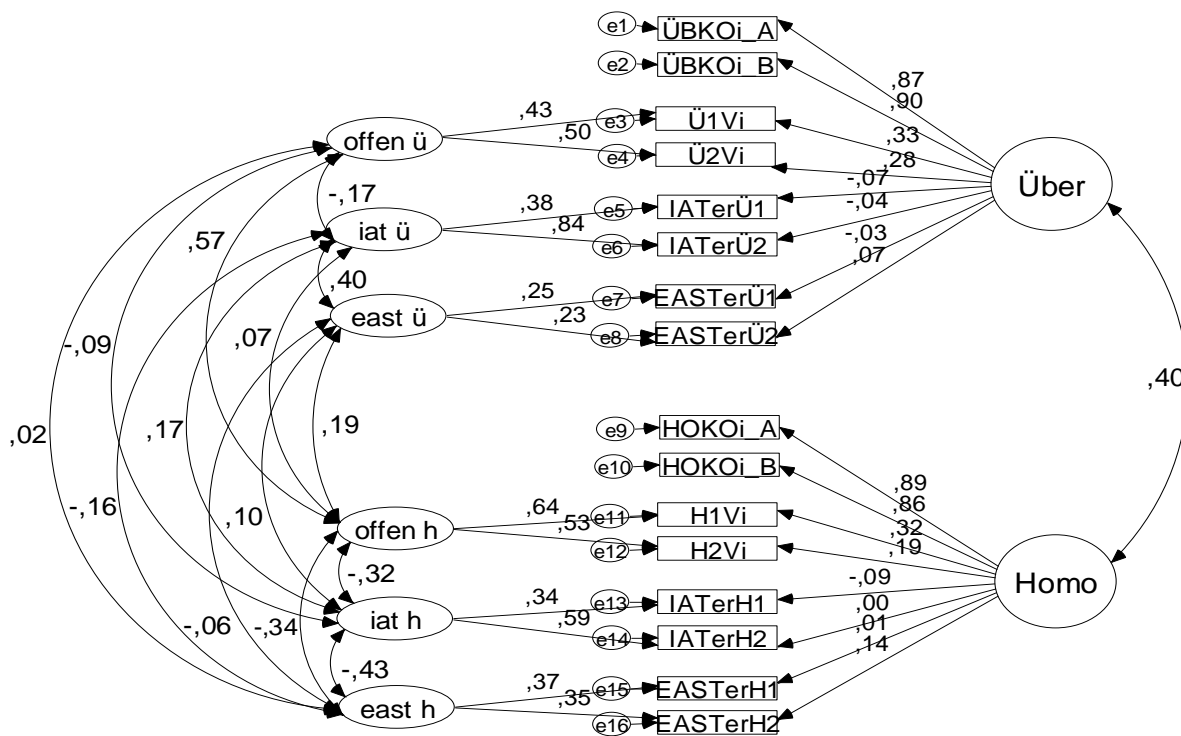
Es werden sowohl das vereinfachte CT-C(M-1) Modell (Abb. 6-33), wie auch das komplexere Multiple Indicator CT-C(M-1) Modell (Abb. 6-34) berechnet, da letzteres mehr Informationen bezüglich der Beziehung zwischen den Methodenfaktoren je Einstellungsobjekt bietet. Zur Identifizierung müssen in beiden Modellen die Traitladungen der Indikatoren, die mit der Referenzmethode erfasst wurden (ÜBKO<sub>i</sub>\_A/B, HOKO<sub>i</sub>\_A/B), auf 1 fixiert werden. Das Multiple Indicator CT-C(M-1) Modell verlangt zusätzliche die Fixierung aller Methodenfaktorladungen (vgl. Eid, 2000; Eid et al., 2003). Neben den grafischen Modelllösungen werden die standardisierten Parameter sowie die Varianzen der latenten Modellkomponenten jeweils mit Signifikanzniveau aufgeführt (Tab. 6-43); die Signifikanzprüfung ist nicht für fixierte Ladungsparameter erhältlich.

#### 6.4.4.2 Ergebnisse

Grundsätzlich werden für beide Modelle vergleichbare Fit-Indizes und Parameterschätzungen erhalten (Abb. 6-33, Abb. 6-34). Der  $\chi^2$ -Statistik zufolge weichen zwar die modellimplizierten Kovarianzmatrizen signifikant von denen der beobachteten Daten ab, dies kann zum einen jedoch Ursache der Nichtnormalverteilung aller Indikatoren sein. Zum anderen weisen beide Modelle gemäß  $\chi^2$ -df-Verhältnis sowie RMSEA und NFI sehr gut bis akzeptable Passung auf.



**Abbildung 6-33**, CT-C(M-1) Modell nach Eid (2000), Direkte Messung als Vergleichsstandard; Modell-Fit:  $\chi^2=127.26$ ,  $df=90$ ,  $p=.01$ ,  $\chi^2/df=1.41$ , RMSEA=.050, NFI=.774



**Abbildung 6-34**, Multiple Indicator CT-C(M-1) Modell nach Eid et al. (2003), Direkte Messung als Vergleichsstandard; Modell-Fit:  $\chi^2=112.41$ ,  $df=83$ ,  $p=.02$ ,  $\chi^2/df=1.35$ , RMSEA=.046, NFI=.800

Auffällig in beiden Modellen ist die geringe konvergente Validität der impliziten Einstellungsindikatoren, für die keine signifikanten Pfadladungen vom jeweiligen Einstellungstrait erhalten werden (vgl. Teil A, Tab. 6-42) bzw. auch keine signifikanten Pfadladungen von den Methodenfaktoren. Die latenten Einstellungsfaktoren „Über“ und „Homo“ werden jeweils aus der gemeinsamen Varianz von Fragebogenskalen und offener Assoziativmessung geschätzt. Statistisch bedeutsame methodenspezifische Varianz besteht lediglich bei traitspezifischer Methodenmodellierung (Multiple Indicator CT-C(M-1)), jedoch auch hierbei nicht für „east ü“ und „east h“ (vgl. Teil B, Tab. 6-42). Insbesondere im Hinblick auf die Varianzschätzungen der latenten Methodenfaktoren werden also Unterschiede zwischen beiden Modellen beobachtet. Während für keinen der drei Methodenfaktoren im CT-C(M-1) Modell eine signifikante Varianzschätzung vorliegt, sind bis auf die beiden „east“-Methoden alle traitspezifischen Methodenfaktoren im Multiple Indicator CT-C(M-1) Modell statistisch bedeutsam. Die hohen Fehlervarianzen der EAST-Indikatoren sind auf die geringen Konsistenzen (traitspezifischer Varianzanteil) bzw. Methodenspezifitäten (methodenspezifischer Varianzanteil) zurückzuführen. Allgemein betrachtet entfällt in beiden Modellen ein jeweils signifikanter Varianzanteil nahezu jedes Indikators auf seinen Fehlerterm.

Der Anteil, durch Trait und Methodenfaktoren erklärter Varianz eines Indikators (multiple quadrierte Korrelation,  $R_i^2$ ) und gibt Aufschluss über die Qualität der Messung. Hier werden zum Teil sehr geringe Varianzaufklärungen erzielt, wogegen jedoch eine hohe Varianzaufklärung durch die latenten

Faktoren und ein entsprechend geringer Anteil unsystematischer Fehlervarianz wünschenswert wären. Die mittlere multiple quadrierte Korrelation der Indikatoren beträgt lediglich .40 im CT-C(M-1) Modell und .39 im Multiple Indicator CT-C(M-1) Modell, woran die EAST-Indikatoren aufgrund von Varianzaufklärungen nahe Null großen Anteil haben. Der Anteil durch einen latenten Faktor erklärter Varianz ( $R^2$ ) wird durch Mittelung seiner quadrierten standardisierten Pfadkoeffizienten erhalten. Die Varianzaufklärung durch die latenten Trait- und Methodenfaktoren im CT-C(M-1)-Modell beträgt sie 39.56% und im Multiple Indicator CT-C(M-1) Modell 39.06%.

Teil A		Parameterschätzung						
		CT-C(M-1)			Multiple Indicator CT-C(M-1)			
Indikator $i$	Trait	$B_i$	$p$	$R^2$		$B_i$	$p$	$R^2$
ÜBKO $_i$ _A	←	.86	/	.74		.87	/	.75
ÜBKO $_i$ _B	←	.90	/	.81		.90	***	.80
Ü1V $i$	←	.33	***	.18		.33	***	.29
Ü2V $i$	←	.28	***	.19		.28	***	.33
IATerÜ1	←	-.05	.537	.91		-.07	.432	.15
IATerÜ2	←	-.02	.779	.12		-.04	.647	.71
EASTerÜ1	←	-.02	.787	.01		-.03	.724	.06
EASTerÜ2	←	.07	.391	.01		.07	.436	.06
HOKO $_i$ _A	←	.77	/	.95		.86	***	.79
HOKO $_i$ _B	←	.97	/	.59		.89	/	.75
H1V $i$	←	.30	***	.40		.32	***	.51
H2V $i$	←	.19	***	.42		.19	***	.31
IATerH1	←	-.11	.181	.03		-.09	.315	.12
IATerH2	←	-.02	.762	.03		-.00	.981	.35
EASTerH1	←	.01	.935	.02		.01	.918	.13
EASTerH2	←	.10	.213	.92		.14	.108	.14
Indikator $i$	Methode	$B_i$	$p$		Indikator $i$	Methode	$B_i$	$p$
Ü1V $i$	←	.28	/		Ü1V $i$	←	.43	/
Ü2V $i$	←	.33	***		Ü2V $i$	←	.50	/
H1V $i$	←	.56	***		H1V $i$	←	.64	/
H2V $i$	←	.62	***		H2V $i$	←	.53	/
IATerÜ1	←	.95	/		IATerÜ1	←	.34	/
IATerÜ2	←	.34	.232		IATerÜ2	←	.59	/
IATerH1	←	.13	.321		IATerH1	←	.38	/
IATerH2	←	.17	.282		IATerH2	←	.84	/
EASTerÜ1	←	.11	/		EASTerÜ1	←	.25	/
EASTerÜ2	←	-.05	.618		EASTerÜ2	←	.24	/
EASTerH1	←	-.14	.301		EASTerH1	←	.37	/
EASTerH2	←	-.95	.650		EASTerH2	←	.35	/

**Tabelle 6-42, Teil A (Teil B nächste Seite)**, Parameter- und Varianzschätzungen für CT-C(M-1)- und Multiple Indicator CT-C(M-1) Modell;  $B_i$ = standardisiertes Pfadgewicht des Indikators  $i$ ,  $R^2$ = Erklärte Indikatorvarianz (entspricht Cronbach's Alpha); \*\*\*,  $p < .05$

<b>Teil B</b>		<b>Varianzschätzung</b>					
		<b>CT-C(M-1)</b>			<b>Multiple Indicator CT-C(M-1)</b>		
<b>latente Modellkomponente</b>	<b>Varianz</b>	<b><i>p</i></b>	<b>R<sup>2</sup></b>	<b>latente Modellkomponente</b>	<b>Varianz</b>	<b><i>p</i></b>	<b>R<sup>2</sup></b>
Über	.26	***	.22	Über	.26	***	.22
Homo	.15	***	.21	Homo	.12	***	.21
offen	.09	.157	.35	offen ü	.22	***	.22
				offen h	.06	***	.35
iat	121.50	.222	.27	iat ü	19.62	***	.43
				iat h	3.97	***	.23
east	1.92	.696	.23	east ü	9.94	.460	.06
				east h	20.28	.104	.13
		<b>CT-C(M-1)</b>			<b>Multiple Indicator CT-C(M-1)</b>		
<b>Fehlerterme <i>e<sub>i</sub></i></b>	<b>Varianz</b>	<b><i>p</i></b>			<b>Fehlerterme <i>e</i></b>	<b>Varianz</b>	<b><i>p</i></b>
e1	.09	***			e1	.09	***
e2	.06	***			e2	.06	***
e3	.97	***			e3	.84	***
e4	.71	***			e4	.59	***
e5	12.41	.900			e5	113.57	***
e6	24.33	***			e6	8.02	.079
e7	157.12	***			e7	148.64	***
e8	178.50	***			e8	169.31	***
e9	.01	.388			e9	.03	***
e10	.10	***			e10	.07	***
e11	.09	***			e11	.08	***
e12	.14	***			e12	.16	***
e13	32.16	***			e13	29.98	***
e14	11.28	***			e14	7.35	***
e15	148.60	***			e15	131.41	***
e16	13.10	.966			e16	139.32	***
<b>Erklärte Varianz an Gesamtvarianz</b>			39.56%				39.06%

**Fortsetzung Tabelle 6-42**

In beiden Modellen bestehen signifikant positive Korrelationen zwischen den latenten Einstellungstrait „Über“ und „Homo“ [CT-C(M-1):  $r=.36$ ; multiple Indicator CT-C(M-1):  $r=.40$ ;  $p<.001$ ]. Die Korrelationen zwischen den Methodenfaktoren sind im CT-C(M-1) Modell prinzipiell als Partialkorrelationen zu verstehen. Sie stellen die für die Differenzen in der Referenzmethode korrigierten Zusammenhänge der Methodenfaktoren dar. Signifikante manifeste Korrelationen zwischen den verschiedenen Messungen, jedoch nicht signifikante Korrelationen der latenten Methodenfaktoren weisen darauf hin, dass mit den modellierten Messmethoden ähnliche Aspekte von Einstellungen erfasst

werden, wie über die direkte Messung via Fragebogen, die als Vergleichsstandard fungiert. Im CT-C(M-1) Modell bestehen keine signifikanten Korrelationen zwischen den drei modellierten Methodenfaktoren und bei Betrachtung der traitspezifischen Methodenfaktoren im Multiple Indicator CT-C(M-1) Modell wird lediglich zwischen den Methodenfaktoren der offenen Assoziativmessung eine signifikante Beziehung erhalten (offen ü, offen h:  $r=.67$ ,  $p<.01$ ).

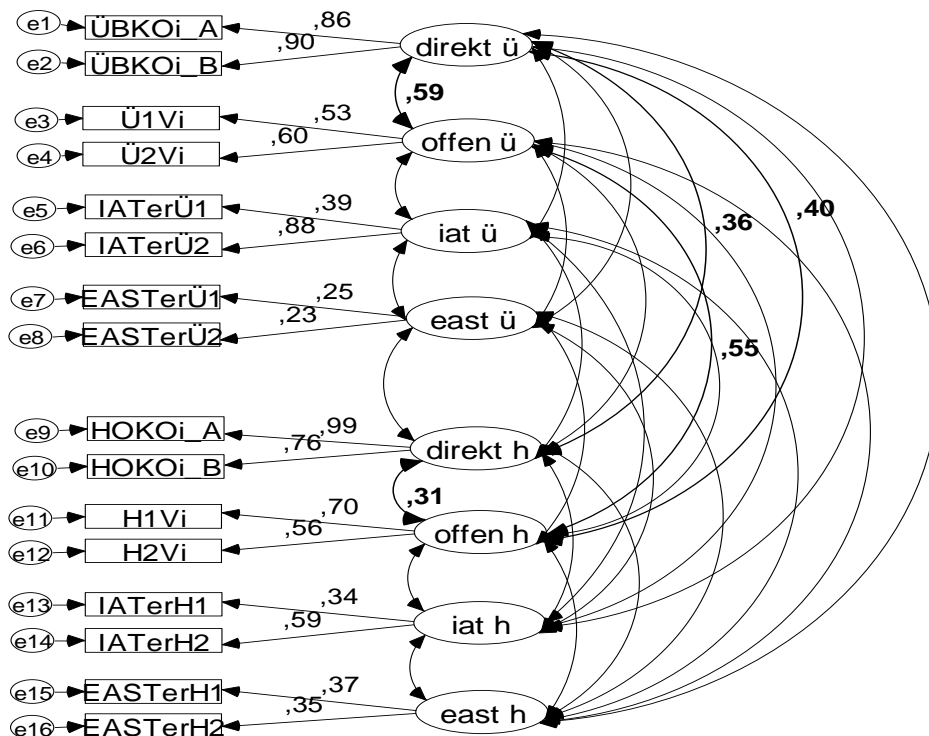
#### 6.4.4.3 *Ergebnisse für Reaktionszeitdaten*

Zur Kontrolle werden die beiden MTMM-Modelle ein zweites Mal mit den reaktionszeit-, statt den fehlerbasierten Indikatoren gemäß IAT und EAST berechnet. Um zunächst die Konvergenz des CT-C(M-1) Modells mit den Reaktionszeitdaten zu gewährleisten, müssen drei zusätzliche Fixierungen vorgenommen werden (eine zusätzliche Pfadladung pro Methodenfaktor). Es ergeben sich jedoch weder drastische Änderungen in der Fit-Statistik ( $\chi^2=149.73$ ,  $df=93$ ,  $p=.00$ ,  $\chi^2/df=1.61$ ,  $RMSEA=.061$ ,  $NFI=.784$ ), noch in der Parameterschätzung im Vergleich zum ursprünglichen Modell mit fehlerbasierten Indikatoren für IAT und EAST. Die Korrelation zwischen den latenten Einstellungstraits beträgt hier beispielsweise  $r=.37$  ( $p<.001$ ). Das aktuelle Modell repräsentiert die empirischen Daten etwas schlechter, als das oben berichtete Modell auf Basis der Fehlereffekte von IAT und EAST. Trait- und Methodenfaktoren erzielen mit 40.00% eine vergleichbare Varianzaufklärung. Das reaktionszeitbasierte Multiple Indicator CT-C(M-1) Modells erzielt dagegen einen leicht besseren Gesamt-Fit, als das fehlerbasierte Modell ( $\chi^2=108.92$ ,  $df=83$ ,  $p=.03$ ,  $\chi^2/df=1.31$ ,  $RMSEA=.044$ ,  $NFI=.843$ ). Jedoch ist die Kovarianzmatrix der Methodenfaktoren nicht positiv definiert, die Modelllösung also nicht akzeptabel. Außerdem sind die Pfadgewichte zwischen dem latenten Einstellungstrait „Homo“ und den IAT-Indikatoren bedeutsam. Es bestehen zudem mehrere signifikante Korrelationen zwischen den Methodenfaktoren (offen h, iat h:  $r=-.30$ , offen h, east h:  $r=-.92$ ,  $p<.05$ ; offen h, offen ü:  $r=.55$ , iat h, iat ü:  $r=.51$ ,  $p<.01$ ). Die Varianzaufklärung durch Trait- und Methodenfaktoren beträgt hier mit 45.93% etwas mehr, als im entsprechenden fehlerbasierten Modell. Auch bei diesem, im Vergleich zu den anderen berichteten Modellen, Maximum an erklärter Varianz handelt sich jedoch um einen relativ geringen Betrag, denn entsprechend 54.07% der Gesamtvarianz bleibt unaufgeklärt.

## 6.4.5 Zwischenresümee

Diese Ergebnisse zusammenfassend können der MTMM-Analyse der erhobenen Maße nicht die erwünschten Aussagen entnommen werden, da bereits auf der Ebene der EAST-Indikatoren keinerlei befriedigenden Konvergenzen im Gesamtmodell existieren: Weder die Pfadladungen zwischen EAST und dem jeweiligen Einstellungsfaktor sind signifikant, was die geringe Reliabilität der Indikatoren aufzeigt, noch liegt ein statistisch bedeutsamer Methodenfaktor für den EAST vor. Die Varianzaufklärung der EAST-Indikatoren durch die latenten Einstellungs- und Methodenfaktoren ist somit mehr oder weniger vernachlässigbar. Die IAT-Indikatoren konvergieren ebenfalls nicht mit der direkten Messung und den Geschichtenvariablen in einem Einstellungsfaktor.

Diese Ergebnisse legen ein reines „State-Modell“ der Einstellungen nahe (Abb. 6-35), das theoretisch dem *Dissociation Approach* (Devine, 1989) zur Einstellungskonzeptualisierung entspricht. Da aufgrund der Datenerhebung die Indikatoren auch durch situative Varianz gekennzeichnet sein könnten, liegt die Bezeichnung „State-Modell“ nahe, obwohl hier zunächst die Annahme vorrangiger Bedeutung von Trait- und Methodenvarianz vor State-Varianz anliegt. Tab. 6-43 enthält die minderungskorrigierten Korrelationen (vgl. Tab. 6-38).



**Abbildung 6-35,** „State-Modell“ der Einstellungen (Fehlerdaten für IAT und EAST);  
Modell-Fit:  $\chi^2=101.93$ ,  $df=84$ ,  $p=.09$ ,  $\chi^2/df=1.21$ ,  $RMSEA=.036$ ,  $NFI=.819$

		1.	2.	3.	4.	5.	6.	7.
1.	<b>direkt ü</b>							
2.	<b>direkt h</b>	.36						
3.	<b>offen ü</b>	<b>.59</b>	-.02					
4.	<b>offen h</b>	.40	.31	<b>.55</b>				
5.	<b>iat (er) ü</b>	-.02	-.12	-.15	.02			
6.	<b>iat (er) h</b>	.12	-.03	-.02	-.32	.16		
7.	<b>east (er) ü</b>	.12	-.26	.08	.11	.37	.11	
8.	<b>east (er) h</b>	-.21	.19	-.10	-.24	-.18	-.45	-.14
		9.	10.	11.	12.	13.	14.	14.
9.	<b>direkt ü</b>							
10.	<b>direkt h h</b>	.37						
11.	<b>offen ü ü</b>	<b>.59</b>	-.02					
12.	<b>offen h h</b>	.39	.32	<b>.54</b>				
13.	<b>iat ü (rt) ü</b>	.09	.12	.24	.12			
14.	<b>iat h (rt) h</b>	.03	.29	.02	-.16	<b>.50</b>		
15.	<b>east ü (rt) ü</b>	.09	.07	.08	.04	.30	.09	
16.	<b>east h (rt) h</b>	<b>-.76</b>	.10	-.23	<b>-.83</b>	-.30	.11	.18

**Tabelle 6-43,** Minderungskorrigierte Zusammenhänge zwischen den Einstellungstrais, oben: fehlerbasierte IAT- und EAST-Indikatoren (er), unten: Reaktionszeitbasierte IAT- und EAST-Indikatoren (rt), ü=Übergewicht, h=Homosexualität

Da „State-„ und MTMM-Modelle nicht ineinander genestet sind, besteht keine Möglichkeit zur direkten Gegenüberstellung. Für einen Augenscheinvergleich können Tabelle 6-44 die Modellstatistiken für die oben dargestellten CT-C(M-1) Modelle und die „State-Modelle“ jeweils auf der Basis der Fehler- bzw. Reaktionszeiteffekte entnommen werden. Die besten globalen Modellpassungen erzielen beide „State-Modelle“, wobei die reaktionszeitbasierte Modellierung einstellungsspezifischer Methodenfaktoren mit 45.23% die höhere Varianzaufklärung erzielt. Die Eingangsmodellierung von jeweils einem Einstellungstrait nach klassischer Konstruktvalidierung steht also in Frage, wobei auch die in jedem Modell erkennbare geringe Reliabilität der EAST-Indikatoren eine kritische Rolle spielen kann.

Modellstatistik	$\chi^2$	df	$\chi^2/df$	p	RMSEA	NFI	R <sup>2</sup>
<b>Fehlerraten</b>							
CT-C(M-1)	127.26	90	1.41	.01	.050	.774	39.56%
MI CT-C(M-1)	112.41	83	1.35	.02	.046	.800	39.06%
„State-Modell“	101.93	84	1.21	.09	.036	.819	39.42%
<b>Reaktionszeiten</b>							
CT-C(M-1)	149.73	93	1.61	.00	.061	.784	40.00%
MI CT-C(M-1)	108.92	83	1.31	.03	.044	.843	45.93%
„State-Modell“	105.95	84	1.26	.05	.040	.847	45.23%

**Tabelle 6-44,** Vergleich von CT-C(M-1)-, Multiple Indikator CT-C(M-1; „MI“- und State-Modell; „er“: Fehlerbasierte IAT- und EAST-Indikatoren, „rt“: Reaktionszeitbasierte IAT- und EAST-Indikatoren



Der zunächst verfolgte CFA-Ansatz zielte, gemäß den beiden betrachteten Einstellungsobjekten, auf zwei latente Einstellungstraits ab. Die Passung einer solchen Modellierung ist für sich genommen nur bedingt aussagekräftig, solange mögliche Alternativmodelle unberücksichtigt bleiben (vgl. Cote, Buckley & Best, 2001). Die Prüfung theoretisch sinnvoller Modellmodifikationen ist nicht nur prinzipiell anzuraten, sondern aufgrund der Inadäquatheit des Eingangsmodells häufig sogar notwendig. Kline (1998) weist entsprechend darauf hin: „And so it often does in CFA that an initial model does not fit the data very well.“ (S. 216). Die eingangs verfolgte CFA-Modellierung zweier Einstellungstraits repräsentiert den *Same Construct Approach* (Wittenbrink, et al. 1997). Der im Rahmen der Einstellungskonzeptualisierung ebenfalls diskutierte *Dissociation Approach* (Devine, 1989) würde dagegen die alternative Modellierung *zweier* Einstellungstraits je Einstellungsobjekt, nämlich eines impliziten und eines expliziten Faktors, nahe legen. Ausgangspunkt eines solchen CFA-Ansatzes ist ein Modell, das einen latenten Faktor je Messung und Einstellungsobjekt annimmt.

#### 6.4.6 Modifikation und Respezifizierung MTMM-Modellierung

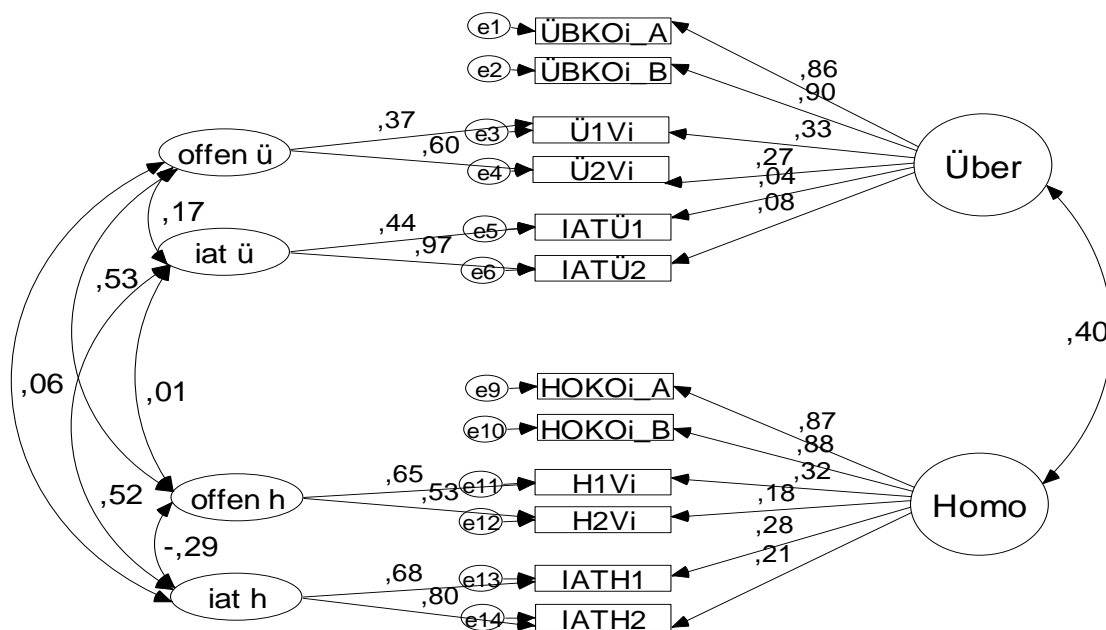
Aufgrund der Ergebnisse der MTMM Analyse werden im Folgenden zwei Modifikationen der ursprünglichen Modellierungen vorgenommen und überprüft:

- 1.) Die Indikatoren des EAST zeigen über alle Modelle hinweg unbefriedigende Reliabilitäten auf, was bereits die Basisanalyse der EAST-Daten erbrachten. Eine Berücksichtigung bzw. Kontrolle des Messfehlers in latenter Strukturgleichungsmodellierung reicht demnach nicht zur Isolation substanzieller systematischer Varianz innerhalb der EAST-Werte eines Einstellungsobjektes. Der EAST ist aus diesem Grund für die Gegenüberstellung mit den drei anderen Einstellungsmessungen als ungeeignet zu bezeichnen und wird daher gerechtfertigt als Datenquelle eliminiert (vgl. Kline, 1998). Im Folgenden wird die latente MTMM Analyse auf der Basis der verbleibenden drei Einstellungsmessungen durchgeführt.
- 2.) Hinsichtlich der IAT-Daten werden entgegen der Messmodelle zunächst die Reaktionszeiteffekte betrachtet, da sie für den IAT das konventionelle Datenformat darstellen und nun die Vergleichbarkeit mit den EAST-Daten nicht mehr notwendig ist.<sup>3</sup> Zudem werden durchgängig höhere Varianzaufklärungen für die Modelle auf Basis der Reaktionszeiten erzielt (vgl. Tab. 6-44). Die Modellergebnisse auf Basis der Fehlerdaten werden zum Vergleich ebenfalls berichtet.

---

<sup>3</sup> Eine gemeinsame Modellierung von Reaktionszeit- und Fehlerdaten wird aufgrund der unterschiedlichen Datenformate und der Messmodellprüfung (siehe Tab. 6-40) nicht vorgenommen.

An dieser Stelle wird allein die modifizierte Multiple Indicator CT-C(M-1) Modellierung der Daten berichtet, die Ergebnisse des systematischen Modellvergleichs nach Widaman (1985) können Anhang A.6.4 entnommen werden. Das Multiple Indicator CT-C(M-1) Modell weist adäquate Fit-Indizes auf und erzielt eine Aufklärung von 58.00% der Varianz der Indikatoren (Abb. 6-36). Jedoch zeigt sich erneut eine verhältnismäßig geringe Trait-Konvergenz. Die gemeinsame Varianz von „Über“ an den Indikatoren für Einstellungen zu Übergewicht beträgt 28.99%, die von „Homo“ an den Einstellungskindikatoren zu Homosexualität 29.81%. Im Vergleich dazu sind die Methodenfaktoren von größerer Bedeutung: „offen ü“ klärt 24.85%, „iat ü“ 56.73%, „offen h“ 35.17% und „iat h“ 55.12% der Varianz an den jeweiligen Indikatoren auf. Jedoch ist ein Methodenfaktor (offen ü) nicht signifikant, was entsprechend auf den relativ hohen Anteil der Referenzmethode „explizit“ an der offenen Assoziativmessung schließen lässt. Die Einstellungsfaktoren korrelieren signifikant positiv, ebenso die traitspezifischen Methodenfaktoren innerhalb einer Messmethode („offen ü“, „offen h“:  $r=.53$ , „iat ü“, „iat h“:  $r=.52$ ,  $p<.05$ ). Die Korrelation zwischen „offen h“ und „iat h“ ist signifikant negativ.

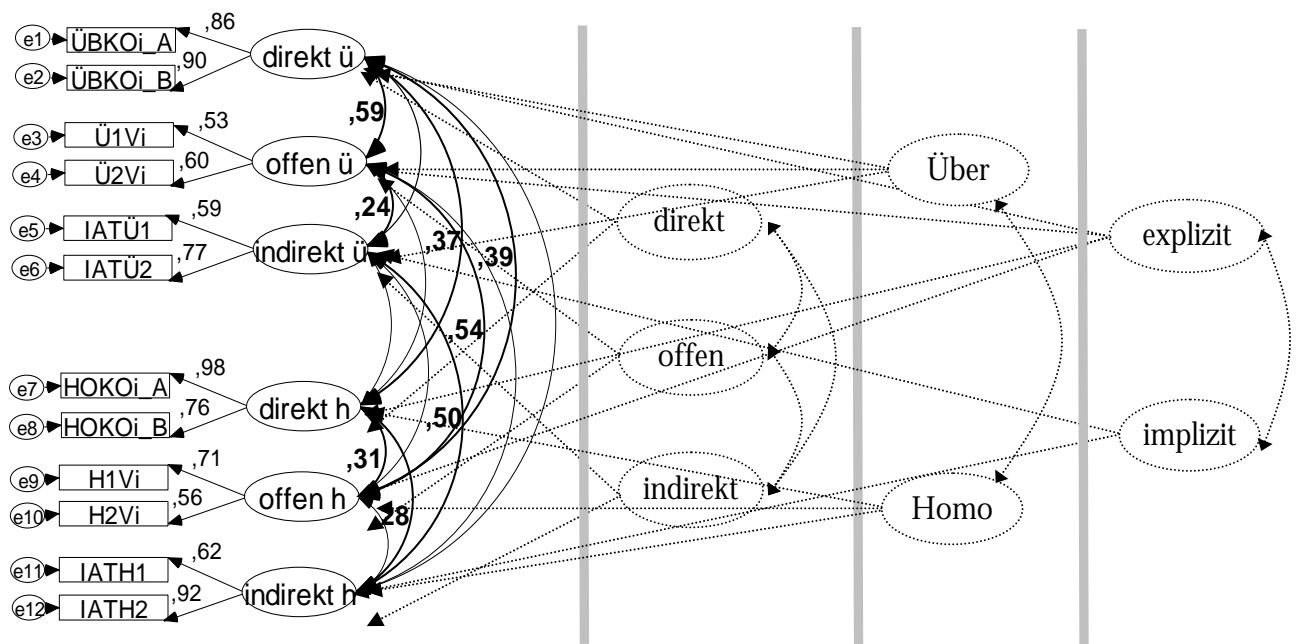


**Abbildung 6-36**, Multiple Indicator CT-C(M-1) Modell mit Maßen aus 3 Datenquellen, direkte Messung: Referenzmethode; Modell-Fit:  $\chi^2=42.71$ ,  $df=39$ ,  $p=.32$ ,  $\chi^2/df=1.10$ ,  $RMSEA=.024$ ,  $NFI=.930$

Das Multiple Indicator CT-C(M-1) Modell stellt eine sinnvolle und akzeptable Anpassung an die empirischen Daten dar, wobei die Pfadladungen der IATÜ-Indikatoren auf den Einstellungsfaktor Übergewicht nicht signifikant sind. Das heißt, der IAT Übergewicht trägt nichts zur Konvergenz von Selbstbericht und Geschichtenmessung bei. Hier hat die Methodenvarianz entsprechend vorrangige Bedeutung.

Zum Vergleich wird erneut ein „State-Modell“ betrachtet (Abb. 6-37), das gemäß der drei Messmethoden für zwei Einstellungsobjekte sechs latente Faktor annimmt, deren Beziehungen untereinander

nicht eingeschränkt sind. Dieses Modell ( $\chi^2=47.69$ ,  $df=45$ ,  $p=.36$ ,  $\chi^2/df=1.06$ ,  $RMSEA=.019$ ,  $NFI=.912$ ;  $R^2=56.08\%$ )<sup>4</sup> zeigt es deutlich: Es bestehen stärkere Zusammenhänge zwischen den gleichen Messmethoden als innerhalb der beiden Einstellungsobjekte. Für Einstellungen zu Übergewicht steht die Geschichtsmessung sowohl zu expliziten, als auch zu impliziten Einstellungen in positiver Beziehung, während kein direkter Zusammenhang zwischen expliziten und impliziten Einstellung besteht. Ein anderes Bild zeigt sich für die Einstellungen zu Homosexualität. Hier korreliert das explizite Maß mit der Geschichtsmessung und dem IAT.



**Abbildung 6-37**, „State-Modell der Einstellungen“ (Reaktionszeitdaten); Modell-Fit:  $\chi^2=47.69$ ,  $df=45$ ,  $p=.36$ ,  $\chi^2/df=1.06$ ,  $RMSEA=.019$ ,  $NFI=.912$

**Abb. 6-37/1**,  
2nd-Order-  
Methodenfaktoren  
direkt, offen, indirekt

**Abb. 6-37/2**,  
2nd-Order-  
Einstellungs-  
faktoren

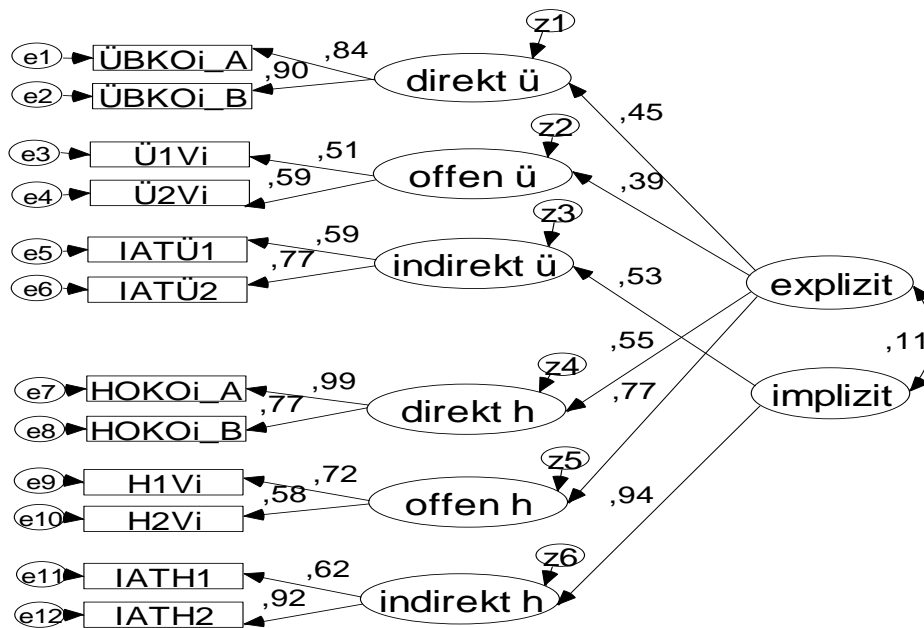
**Abb. 6-37/3**,  
2nd-Order Faktoren  
expliziter und impliziter  
Einstellungen

Faktoren zweiter Ordnung wären also am ehesten innerhalb einer Messmethode zu erwarten (Abb. 6-37/1). Eine entsprechende Aggregation der Faktoren erster Ordnung zu methodenspezifischen Faktoren zweiter Ordnung ist jedoch mit einer signifikanten Modellverschlechterung und einer Out-Of-Range-Schätzung<sup>5</sup> verbunden ( $\chi^2= 84.03$ ,  $df= 54$ ,  $p=.01$ ,  $\chi^2/df= 1.56$ ,  $RMSEA= .058$ ,  $NFI= .861$ ;  $R^2= 50.08\%$ ). Die Modellierung von Einstellungsfaktoren zweiter Ordnung (Abb. 6-37/2) resultiert in einer vergleichsweise stärkeren Fit-Verschlechterung ( $\chi^2= 136.62$ ,  $df= 57$ ,  $p=.00$ ,  $\chi^2/df= 2.40$ ,  $RMSEA= .092$ ,  $NFI= .775$ ;  $R^2= 43.15\%$ ) und einer Korrelationsschätzung von  $r=1.24$  zwischen beiden Second-Order-Faktoren.

<sup>4</sup> Ein fehlerdatenbasiertes Modell fittet die Daten ähnlich gut ( $\chi^2=51.43$ ,  $df=45$ ,  $p=.24$ ,  $\chi^2/df=1.14$ ,  $RMSEA=.039$ ,  $NFI=.896$ ,  $R^2=49.69\%$ ), jedoch werden hier weniger signifikante Beziehungen zwischen den latenten Faktoren beobachtet.

<sup>5</sup> Standardisierte Pfadladung von offen auf offen h von 1.09

Die konsistenten Zusammenhänge zwischen direkter und offener Messung begründen möglicherweise *einen* zugrunde liegenden Faktor expliziter Einstellungen. Ein Modell zweier Second-Order-Methodenfaktoren (Abb. 6-37/3) ist in das Modell gemäß Abb. 6-37/1 genestet. Die Aggregation von direkter und offener Messmethode in einem expliziten Faktor führt zwar zu einer signifikant schlechteren Modellpassung, als die Modellierung von drei Second-Order-Methodenfaktoren, jedoch ist das  $\chi^2$ -df-Verhältnis mit 1.79 nach wie vor gut, ebenso zeigen RMSEA und NFI einen akzeptablen Modell-Fit an und nicht zuletzt tritt keine fehlerhafte Parameterschätzungen auf (Abb. 6-38).



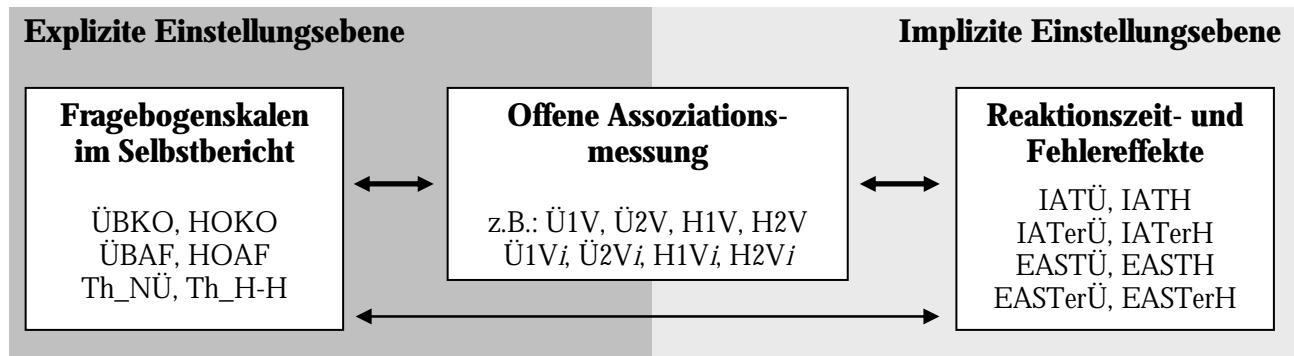
**Abbildung 6-38**, Modellierung nach Abb 10-3, Modell-Fit:  $\chi^2=101.88$ ,  $df=57$ ,  $p=.00$ ,  $\chi^2/df=1.79$ , RMSEA=.069, NFI=.832;  $R^2=50.08\%$

#### 6.4.7 Gesamtresümee der MTMM Analyse

Konstruktvalidität im klassischen Sinne kann für die beiden Einstellungsobjekte nicht angenommen werden, da die ungenügend hohe Konvergenz der Einstellungsindikatoren dazu führt, dass nicht jeweils ein latenter Einstellungsfaktor anzunehmen ist. Es bestehen zwar zum Teil Beziehungen zwischen den verschiedenen Maßen einer inhaltlichen Einstellungsdisposition, diese repräsentieren jedoch nicht zwei zugrunde liegende Traits. Vielmehr müssten explizite und implizite Einstellungen getrennt betrachtet werden, wobei stärkere Beziehungen innerhalb der Messebenen, als innerhalb eines Einstellungsobjekts zu beobachten sind. Diese Ergebnisse liefern eindeutige Hinweise auf die Dissonanz impliziter und expliziter Einstellungen. Die offene Assoziativmessung Für den IAT sollten Reaktionszeiten als Datenformat betrachtet werden. Der EAST erweist sich als wenig geeignetes Maß für die vorliegende Fragestellung.

## 6.5 Ergebnisse III: Explizit-Implizit-Kontinuum der Einstellungen

Die erhobenen Einstellungsmaße können theoretisch auf einer Dimension von expliziten hin zu impliziten Einstellungsanteilen angeordnet werden (Abb. 6-39), wobei sich ein zweidimensionales Kontinuum der Einstellungen durch unterschiedlich enge Beziehungen zwischen den verschiedenen Einstellungsmaßen äußern sollte.



**Abbildung 6-30,** Theoretische Dimensionalität der Datenstruktur

Zur Untersuchung eines Explizit-Implizit-Kontinuums der Einstellungen werden latente Vorhersagemodelle berechnet, wobei die Vorhersage an der theoretischen Dimensionshypothese orientiert ist, weshalb jeweils die Vorhersage relativ direkter Einstellungsmessung durch relativ indirekte Maße erfolgt. Die detaillierte Ergebnisdokumentation der *bimodalen* Vorhersagemodelle, die auch verkreuzte Vorhersagepfade prüfen, erfolgt im Anhang A.6.5 (Prädiktion 1A repliziert beispielsweise die Fragestellung von Gawronski, 2004). Hier wird lediglich die Zusammenfassung berichtet.

Zentraler an dieser Stelle ist die simultane Modellierung der Vorhersage expliziter Einstellungen mit den übrigen Einstellungsmaßen. Zusätzlich zu den Ergebnissen der latenten MTMM-Analyse werden damit Schätzungen der Varianzaufklärung erbracht sowie eine Untersuchung unmittelbarer und medierter Vorhersage von expliziten Einstellungen, durch die offene Assoziativmessung, die sowohl explizite, als auch implizite Einstellungsanteile erfassen sollte und die implizite(n) Messung(en). Da unter Berücksichtigung des Messfehlers bereits substanziellere Methoden-, als Traiteffekte aufgedeckt wurden, kann lediglich eine moderate Varianzaufklärung der Vorhersage erwartet werden.

### 6.5.1 Zusammenfassung der Basismodelle der Vorhersage

Die bimodalen Vorhersagemodelle reflektieren deutliche Unterschiede in der Varianzaufklärung der relativ direkten Kriteriumsmaße durch die relativ indirekten Prädiktormaße. Die Kreuzpfade der Vorhersage sind in keinem Fall statistisch bedeutsam, dennoch übersteigt ihre Bedeutung teilweise die der Vorhersagepfade innerhalb des Einstellungsobjektes (Tab. 6-45, siehe auch Anhang A.6.5).

Prädiktions-Modell	Kriteriums-messung	Prädiktor-messung	R <sup>2</sup> für Einstellungen zu Übergewicht	R <sup>2</sup> für Einstellungen zu Homosexualität
1a	Selbstbericht	IAT	0.8%	9.7%
1b	Selbstbericht	EAST	58.5%	5.7%
2	Selbstbericht	Geschichten-messung	35.2%	15.6%
3a	Assoziativmessung	IAT	6.9%	7.3%
3b	Assoziativmessung	EAST	5.2%	72.1%

**Tabelle 6-45,** Varianzaufklärung für relativ direkte Einstellungsmessung durch relativ indirekte Einstellungsmessung

### 6.5.2 Gesamtmodelle der Vorhersage

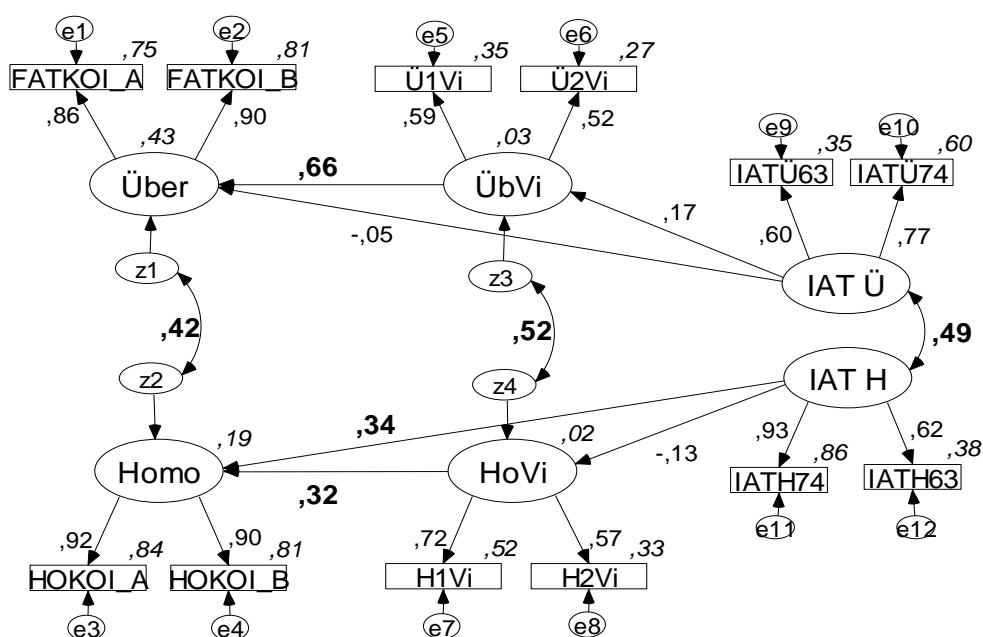
Im Anschluss an die separate Analyse der Vorhersage relativ direkter Einstellungsmaße (z.B. Selbstbericht) durch relativ indirekte Maße (z.B. IAT) erfolgt die simultane Modellierung der Vorhersage der Einstellungsmaße. Dabei werden direkte und vermittelte Vorhersagepfade auf ihre statistische Bedeutsamkeit geprüft, woraus Aussagen über die inkrementelle Validität der Einstellungsprädiktoren (hier IAT, EAST und Geschichtenmessung) zur Vorhersage der expliziten Fragebogenmessung abgeleitet werden können. Die simultane Vorhersage expliziter Einstellungen berücksichtigt eventuelle Multikollinearitäten unter den Prädiktoren und abhängigen Variablen, gibt Aufschluss über die maximale Varianzaufklärung der abhängigen Variablen und ermöglicht die Betrachtung direkter und mediierter Zusammenhänge zwischen den Einstellungsmessungen, die auf einem Kontinuum von direkter zu indirekter Methodik anzuordnen sind. Das Modell entspricht den Beziehungshypothesen für die Einstellungsmessungen (Abb. Abb. 6-39).

Aufgrund der geringen internen Konsistenz EAST und der schwachen bzw. erwartungskonträren Zusammenhänge zu den anderen Einstellungsmaßen erfolgt die Gesamtmodellierung zunächst unter Ausschluss der EAST-Daten. Die Bedeutung einer Einbeziehung des EAST wird jedoch ebenfalls geprüft und im Anschluss berichtet. Der EAST wird im Vergleich zum IAT als Messung mit deutlich indirekterem Charakter betrachtet und bezeichnet entsprechend den impliziten Pol des theoretischen Explizit-Implizit-Kontinuums.

## 6.5.2.1 Vorhersagemodell für direkte Messung, Geschichtenmessung und IAT

Die simultane Betrachtung von IAT, offener Assoziativmessung und Selbstbericht resultiert ungeachtet der eliminierten Kreuzpfade *zwischen* den Einstellungsobjekten in einer gesteigerten Varianzaufklärung für die expliziten Einstellungsfaktoren (vgl. Prädiktion 1b und 2). Die Fixierung der Kreuzpfade auf Null führt zwar zu einem höheren  $\chi^2$ -Wert (47.57,  $df=51$ ,  $p=.45$ ,  $\chi^2/df=1.01$ , RMSEA=.009, NFI=.926), dieser zeigt jedoch keine signifikante Modellverschlechterung an (Gewinn von 4 Freiheitsgraden bei  $\chi^2$ -Einbußen von 7.07). Aus diesem Grund wird also das sparsamere Modell ohne Kreuzprädiktion angenommen (Abb. 6-40).

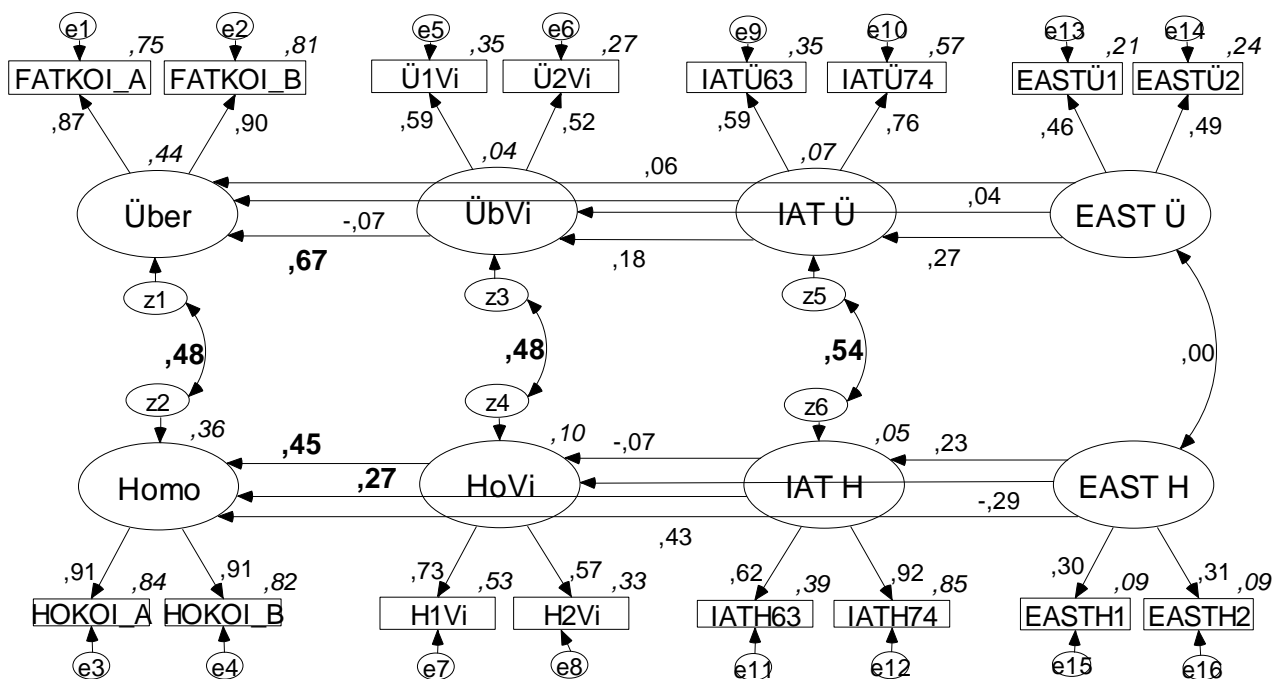
Die signifikanten Pfadgewichte zur Vorhersage expliziter Einstellungen zu Übergewicht und Homosexualität zeigen erwartungsgemäß positive Zusammenhänge an. Während explizite Einstellungen zu Übergewicht lediglich durch die Geschichtenmessung für Übergewicht vorhergesagt werden, sind sowohl die Homosexualitätsgeschichten, als auch der IAT Homosexualität für die Vorhersage expliziter Einstellungen zu Homosexualität von Bedeutung. Jedoch wird eine mehr als doppelt so hohe Varianzaufklärung für explizite Einstellungen zu Übergewicht ( $R^2=43.0\%$ ) im Vergleich zu Homosexualität ( $R^2=19.3\%$ ) erzielt. Die beiden latenten Einstellungsfaktoren der Geschichtenmessung werden zu relativ unbedeutende Varianzanteilen durch die IAT-Maße aufgeklärt ( $R^2=2.9\%$  für Übergewicht bzw.  $R^2=1.6\%$  für Homosexualität). Bis auf eine Fehlervarianz (e11) werden für alle latenten Variablen signifikante Varianzanteile geschätzt. Zwischen den latenten Einstellungstrait einer Messmethode (IAT) bzw. den jeweiligen Residuen der Vorhersage (Geschichtenmessung, Selbstbericht) bestehen signifikant positive Zusammenhänge. Die Korrelationen zwischen z1 und z2 sowie z3 und z4 bezeichnen partielle Zusammenhänge der nicht vorhergesagten Varianz der latenten Einstellungstrait.



**Abbildung 6-40,** Direkte und medierte Vorhersage von expliziten Einstellungen  
 Modell-Fit:  $\chi^2=54.64$ ,  $df=51$ ,  $p=.34$ ,  $\chi^2/df=1.07$ , RMSEA=.021, NFI=.916;  
 Kursive Parameter: quadrierte multiple Korrelationen

## 6.5.2.2 Vorhersagemodell für direkte Messung, Geschichtenmessung, IAT und EAST

Die Einbeziehung des EAST in das Gesamtmodell kann nur Unkorreliertheitsannahme der latenten EAST-Faktoren erfolgen.<sup>6</sup> Da die manifesten Korrelationen zwischen den beiden EAST-Effekten jedoch nahezu Null ist ( $r=.05$ ,  $r>.05$ ) kann diese Restriktion als annehmbar bezeichnet werden. Bis auf den latenten EAST-Faktor für Homosexualität (EAST H) und eine Fehlervarianz (e12) werden signifikante Varianzschätzungen für alle latenten Faktoren ermittelt. Durch die Hinzunahme der EAST-Daten zur Vorhersage wird eine zusätzliche Varianzaufklärung von 16.8% an den expliziten Einstellungen zu Homosexualität erzielt. Die entspricht knapp dem doppelten Anteil erklärter Varianz im Vergleich zum Vorhersagemodell unter Ausschluss des EAST. Für die direkte Einstellungsmessung zu Übergewicht ist die Veränderung im  $R^2$  vernachlässigbar ( $R^2= 1.4\%$ ). Die Geschichtenmessung ÜbVi wird zu 3.6% und HoVi zu 9.9% aufgeklärt, der IAT Ü zu 7.0% und der IAT H zu 5.2% (Abbildung 6-41).



**Abbildung 6-41,** Direkte und medierte Vorhersage von expliziten Einstellungen mit allen Maßen

Modell-Fit:  $\chi^2=121.08$ ,  $df=97$ ,  $p=.05$ ,  $\chi^2/df=1.25$ ,  $RMSEA=.039$ ,  $NFI=.835$ ;

Kursive Parameter: quadrierte multiple Korrelationen

<sup>6</sup> Wird der Zusammenhang zwischen EAST Ü und EAST H nicht restringiert, konvergiert des Modells in keiner zulässigen Lösung.



### 6.5.3 Resümee zur Einstellungsprädiktion

Zwischen den Maßen der verschiedenen Einstellungsmessungen bestehen keine durchgängig bedeutsamen Zusammenhänge. Der EAST weist gravierende psychometrische Mängel auf, weshalb die Validität der resultierenden Daten anzuzweifeln ist. Nichtsdestotrotz wird eine Verdopplung der Varianzaufklärung für explizite Einstellungen zu Homosexualität durch Einbeziehung des EAST erzielt, der EAST besitzt also ein gewisses Vorhersagepotential. *Zwischen* den Einstellungsobjekten bestehen jedoch häufig erwartungskonträre Zusammenhänge seitens des EAST. Darüber hinaus sind keine Korrelation zwischen dem EAST-Übergewicht und -Homosexualität zu verzeichnen. Die IAT-Maße reflektieren dagegen systematische Varianz. Für das Einstellungsobjekt Homosexualität besitzt der IAT inkrementelle Validität hinsichtlich expliziter Einstellungen. Zudem unterhalten beide IAT eine positive Beziehung zueinander. Da jedoch kein signifikanter Zusammenhang zwischen IAT-Übergewicht und expliziten Einstellungen zu Übergewicht bestehen, ist anzunehmen, dass die Korrelation zwischen beiden IATs insbesondere messmethodische Varianz widerspiegelt.

Explizite Einstellungen korrespondieren konsistent für beide Einstellungsobjekte mit dem Vorurteilsgehalt der Geschichten. Hierbei stellt sich die Frage, ob die Quelle dieser Kovarianz inhaltlicher oder methodischer Natur ist. Handelt es sich also um eine Vorurteilsneigung gegenüber den betrachteten Einstellungsobjekten, die von beiden Messungen erfasst wird, oder spiegeln sowohl Selbstbericht, als auch die Geschichtenmessung sozial erwünschtes Antwortverhalten wider. Methodisch sind beide Messungen lediglich hinsichtlich des verbalen Formats und der Möglichkeit, die eigenen Antworten zu kontrollieren, vergleichbar. Inwiefern diese Kontrollmöglichkeit, die für IAT und EAST in dieser Form nicht gegeben ist, Ursache der hohen positiven Korrelationen zwischen Fragebogen und Geschichtenmessung ist, wird mittels Moderatoranalysen in Abschnitt 6.7 geprüft.

## 6.6 Ergebnisse IV: Bedeutung individueller & antizipierter Einstellungen

An dieser Stelle wird die Bedeutung einer relativen Abweichung bzw. Übereinstimmung zwischen den individuellen und den gesellschaftlich antizipierten Einstellungen auf den Zusammenhang expliziter und impliziter Maße untersucht. Sollte der IAT vor allem Umweltassoziationen erfassen (Karpinski & Hilton, 2001), wären hier höhere Korrelationen zu den *antizipierten* Einstellungen zu erwarten. Dies ist nicht der Fall, wie Tabelle 6-35 zu entnehmen ist. Nun bietet das Erhebungsdesign von Studie III die Möglichkeit, Personen, die relativ hohe Abweichungen zwischen den individuellen und den antizipierten Einstellungen von Personen mit hoher Übereinstimmung zwischen eigener und gesellschaftlicher

Meinung zu unterscheiden und für sie die Korrespondenz impliziter und expliziter Maße zu vergleichen. Zunächst ist festzuhalten, dass die Differenzmaße der Einstellungsskalen enge Beziehungen zu einander unterhalten, insbesondere innerhalb der Einstellungsobjekte (Tab. 6-45).

	<b>ÜBKO</b>	<b>ÜBAF</b>	<b>HOKO</b>
<b>ÜBKO</b>			
<b>ÜBAF</b>	<b>.65</b>		
<b>HOKO</b>	<b>.37</b>	<b>.30</b>	
<b>HOAF</b>	<b>.40</b>	<b>.48</b>	<b>.61</b>

**Tabelle 6-45,** Zusammenhänge der Differenzmaße individueller und antizipierter Einstellungen

Um die Bedeutung der Abweichung zwischen individuellen und antizipierten Einstellungen für die Korrespondenz expliziter und impliziter Maße zu untersuchen, wird die Stichprobe mittels Mediansplit für jede der vier Differenzmaße zwischen individuellen und antizipierten Einstellungen (ÜBKO, ÜBAF, HOKO, HOAF) in zwei Gruppen geteilt, in diejenigen, die relativ geringe Unterschiede zwischen eigenen und gesellschaftlichen Einstellungen aufzeigen und diejenigen, für die relativ hohe Divergenzen bestehen. Nach den Ergebnissen von Nosek (2004) wären höhere Korrelationen zwischen expliziten und impliziten Einstellungen bei Personen zu erwarten, die eine stärkere Abweichung zwischen ihren eigenen Einstellungen und denen der durchschnittlichen Anderen wahrnehmen.

### 6.6.1 Einstellungsobjekt Übergewicht

Für die Korrespondenz impliziter und expliziter Einstellungen zu Übergewicht spielt die Abweichung zwischen individuellen und antizipierten Einstellungen eine eher untergeordnete Rolle (Tab. 6-46).

	<b>ÜB- KO<sub>i</sub></b>	<b>ÜB- KO<sub>a</sub></b>	<b>ÜB- AF<sub>i</sub></b>	<b>ÜB- AF<sub>a</sub></b>	<b>Th_ NÜ</b>	<b>ÜB- KO<sub>i</sub></b>	<b>ÜB- KO<sub>a</sub></b>	<b>ÜB- AF<sub>i</sub></b>	<b>ÜB- AF<sub>a</sub></b>	<b>Th_ NÜ</b>
	<b>Abweichung in ÜBKO gering</b>					<b>Abweichung in ÜBKO hoch</b>				
IATÜ	.11	-.02	.11	.17	-.04	.03	.11	.13	-.01	-.19
IATerÜ	.02	.04	-.14	-.05	.02	-.15	.12	-.04	.12	.10
EASTÜ	.14	.18	.06	-.02	.05	-.06	.14	-.04	.06	-.18
EASTerÜ	.04	.09	.14	.08	.12	-.03	.02	-.08	.06	-.03
	<b>Abweichung in ÜBAF gering</b>					<b>Abweichung in ÜBAF hoch</b>				
IATÜ	.07	.02	.15	.16	.05	.08	.04	.11	.03	.05
IATerÜ	-.01	-.05	-.08	-.10	.03	-.04	.13	-.05	.09	-.23
EASTÜ	.06	-.01	-.07	.01	.20	.02	.24	.03	.06	-.09
EASTerÜ	.06	-.00	.09	.11	-.01	-.00	.04	-.05	.05	-.12

**Tabelle 6-46;** Bedeutung der Abweichung individueller von antizipierten Einstellungen zu **Übergewicht** für die implizit-explizit-Korrespondenz, ÜBKO= kognitive Einstellungen (Median= 1.05), ÜBAF=affektive Einstellungen (Median= .50); *i/a*: individuelle/ antizipierte Einstellungen; Th\_NÜ= Differenz der Gefühlsthermometer Normal- und Übergewichtiger; IATÜ/EASTÜ=Reaktionszeiteffekte von IAT und EAST; IATerÜ/EASTerÜ=Fehlereffekte von IAT und EAST

Lediglich für zwei Korrelationen zwischen expliziten und impliziten Einstellungen wird ein Effekt der Selbst-Andere-Divergenz beobachtet. Bei hoher Divergenz zwischen individuellen und antizipierten *affektiven* Einstellungen zu Übergewicht besteht eine negative Beziehung zwischen Th\_NÜ und IATerÜ. Das bedeutet also, wenn die persönlichen affektiven Einstellungen als abweichend von der Allgemeinheit wahrgenommen werden, geht eine hohe Thermometerdifferenz mit einem geringeren IAT-Fehlereffekt einher. Dieser negative Zusammenhang muss allgemein als erwartungskonträr bezeichnet werden. Da beide Maße im Sinne einer Ablehnung von übergewichtigen Personen gleich gerichtet sind, sollten hier prinzipiell positive Beziehungen bestehen.

Für Probanden, die eine hohe Selbst-Andere-Divergenz in den affektiven Einstellungen aufweisen, besteht zudem eine signifikante Korrelation zwischen den antizipierten kognitiven Einstellungen zu Übergewicht und dem Reaktionszeiteffekt des EASTÜ ( $r=.24, p<.05$ ), während für Personen mit geringen Abweichungen zwischen eigener und gesellschaftlicher Haltung in den affektiven Einstellungen kein entsprechender Implizit-Explizit-Zusammenhang besteht ( $r=-.01, p=.930$ ).

### 6.6.2 Einstellungsobjekt Homosexualität

Für die Einstellungen zu Homosexualität bestehen im Gegensatz zum Einstellungsobjekt Übergewicht deutliche Unterschiede im Beziehungsmuster expliziter und impliziter Einstellungsmaße in Abhängigkeit von der Selbst-Andere-Divergenz der expliziten Einstellungen (Tab. 6-47).

	<b>HO-KO<i>i</i></b>	<b>HO-KO<i>a</i></b>	<b>HO-AF<i>i</i></b>	<b>HO-AF<i>a</i></b>	<b>Th_- HH</b>	<b>HO-KO<i>i</i></b>	<b>HO-KO<i>a</i></b>	<b>HO-AF<i>i</i></b>	<b>HO-AF<i>a</i></b>	<b>Th_- HH</b>
	<b>Abweichung in HOKO gering</b>					<b>Abweichung in HOKO hoch</b>				
IATH	<b>.30</b>	.14	<b>.43</b>	<b>.32</b>	<b>.32</b>	.26	.22	.19	.08	.14
IATerH	-.03	.00	-.05	.06	-.02	-.07	-.14	.19	-.21	-.16
EASTH	-.01	.03	-.06	.04	-.04	.19	.21	.08	.19	-.10
EASTerH	.01	.05	.03	.00	-.11	.12	.07	-.10	<b>.31</b>	-.01
	<b>Abweichung in HOAF gering</b>					<b>Abweichung in HOAF hoch</b>				
IATH	<b>.33</b>	.06	<b>.45</b>	<b>.36</b>	<b>.33</b>	.18	.17	.26	.13	.14
IATerH	-.04	.04	.06	.09	-.07	-.14	-.05	.07	-.20	-.12
EASTH	.04	.02	.02	.06	.05	.14	.09	.05	.14	-.18
EASTerH	.17	-.09	.13	.11	.07	.12	.01	.02	.08	-.11

**Tabelle 6-47;** Bedeutung der Abweichung individueller von antizipierten Einstellungen zu **Homosexualität** für die implizit-explizit-Korrespondenz, HOKO= kognitive Einstellungen (Median= 1.53), HOAF=affektive Einstellungen (Median= 1.11); *i/a*: individuelle/ antizipierte Einstellungen; Th\_HH= Differenz der Gefühlsthermometer Hetero- und Homosexueller; IATH/EASTH=Reaktionszeiteffekte von IAT und EAST; IATerH/EASTerH=Fehlereffekte von IAT und EAST

Sowohl für kognitive und affektive Einstellungen, als auch die Gefühlsthermometerdifferenz resultiert eine geringe Divergenz zwischen individueller und antizipierter Disposition in höheren Zusammenhängen zwischen expliziten Maßen und dem Reaktionszeiteffekt des IATH, während für die Gruppe der Hochdivergenten deutlich geringere und größtenteils nicht signifikante Korrelationsbeträge für IATH und Fragebogenmaß beobachtet werden. Ausnahme sind hierbei die antizipierten kognitiven Einstellungen zu Homosexualität, für die eine Selbst-Andere-Abweichung keinen bedeutsamen Einfluss auf den Zusammenhang zum IATH. Darüber hinaus besteht für den Zusammenhang zwischen den antizipierten affektiven Einstellungen zu Homosexualität und den Fehlereffekt des EASTH ein Unterschied zwischen hoch- und niedrigdivergenten Probanden. Personen mit geringen Unterschieden zwischen eigenen und gesellschaftlich antizipierten kognitiven Einstellungen weisen keine Beziehung zwischen HOAFa und EASTerH ( $r=.00$ ,  $p=.986$ ). Im Gegensatz dazu besteht ein signifikant positiver Zusammenhang zwischen beiden Maßen ( $r=.31$ ,  $p<.01$ ).

Insgesamt ist also festzuhalten, dass die Divergenz individueller und antizipierter expliziter Einstellungen kaum Bedeutung für die Explizit-Implizit-Korrespondenz innerhalb der Einstellungen zu Übergewicht besitzt. Dagegen wirkt eine relative Divergenz zwischen individuellen und antizipierten Einstellungen zu Homosexualität auf die Beziehung zwischen affektiven bzw. kognitiven Einstellungen und IAT zusammenhangsmindernd.

## 6.7 Ergebnisse V: Moderatoranalysen Motivation zur Vorurteilsfreiheit

In diesem Abschnitt wird zunächst untersucht, inwieweit eine Motivation, Vorurteile zu kontrollieren bzw. sich vorurteilsfrei zu präsentieren, Auswirkungen auf die Zusammenhänge der verschiedenen Einstellungsmaße hat. Im Hinblick auf die geringen bzw. nicht bedeutsamen Zusammenhänge zwischen impliziten (IAT, EAST) und expliziten Einstellungen können Moderatoreffekte von Vorurteilkontrollmotivation dergestalt vorliegen, dass speziell hoch motivierte Personen ihre expliziten Einstellungen entsprechend färben während niedrig motivierte Personen dies nicht tun. Nach Devine (1989) ist jedoch auch davon auszugehen, dass insbesondere eine intrinsische Motivation als eine tief greifende Überzeugung Einfluss auf relativ schwer kontrollierbare implizite Messungen wie den IAT ausüben kann. In Bezug auf die relativ hohen Zusammenhänge zwischen den direkten Einstellungen und der Geschichtenmessung ist zu prüfen, inwieweit diese auf die relativ zu den impliziten Messungen höhere Kontrollierbarkeit der eigenen Reaktionen zurückzuführen ist. In diesem Fall sollten die Korrelationen zwischen Selbstbericht und Geschichtenmessung deutlich sinken bzw. nicht länger signifikant sein.

### 6.7.1 Moderatoreffekte der Beziehung zwischen impliziten und expliziten Einstellungen

Als Moderator der Explizit-Implizit-Beziehung wird die Motivation zur Vorurteilsfreiheit diskutiert. Zur Untersuchung der Moderatorwirkungen von vorurteilsbezogener Motivation wird der extrahierte individuelle Motivationsfaktors IVMF zunächst mit dem klassischen Moderatoransatz hierarchischer Regressionsmodelle untersucht (z.B. Banse & Gawronski, 2003). Im Anschluss erfolgt eine varianz-analytische Untersuchung der Einstellungen in Abhängigkeit von der Ausprägung von Internaler und Externaler Motivation zur Vorurteilsfreiheit (IMS, EMS) nach Devine et al. (2002).

#### 6.7.1.1 Hierarchische Regression

Mittels hierarchischer Regressionsanalysen kann die Bedeutung eines Einflussfaktors auf die Beziehung zwischen verschiedenen Maßen über die inkrementelle Vorhersagevalidität des Moderatorterms geprüft werden. Um die statistische Vergleichbarkeit der betrachteten Variablen zu gewährleisten, werden Kriterium und Prädiktoren jeweils z-standardisiert, der Moderatorterm wird schließlich als Produkt der standardisierten Prädiktoren gebildet (vgl. Cohen & Cohen, Aiken & West, 2003; Friedrich, 1982). Wichtige Parameter der Regressionsanalyse sind die inkrementelle Varianzaufklärung des Moderatorterms, angezeigt durch die Änderung in  $R^2$  sowie das standardisierte Regressionsgewicht des jeweiligen Moderatorterms, das angibt, um welchen Betrag sich jeweils der Zusammenhang zwischen expliziten und impliziten Einstellungen pro Standardabweichungseinheit des Moderators ändert. Entsprechend der verschiedenen Einstellungsmaße je Einstellungsobjekt können Moderatoreffekte für eine Reihe möglicher Vorhersagemodelle untersucht werden. In vorliegender Arbeit ist insbesondere die *individuelle* Vorurteilsneigung von Interesse, entsprechend betrachtete Moderatoren sind die individuellen Motivationen zur Vorurteilsfreiheit, welche vom motivationalen Faktor IVMF repräsentiert werden. Die expliziten Einstellungen stellen das Kriterium dar, das im ersten Schritt durch implizite Einstellungen und Motivation zur Vorurteilsfreiheit vorhergesagt wird. Im zweiten Schritt wird der Produktterm beider Prädiktoren in die Vorhersage miteinbezogen und auf inkrementelle Validität und Zusammenhang zum Kriterium hin geprüft. Die Untersuchung von Moderationseffekten in der Beziehung zwischen impliziten und expliziten Einstellungen führt entsprechend der beiden impliziten (IAT und EAST) und der drei expliziten Maße je Einstellungsobjekt (kognitive und affektive Einstellungen, Gefühlsthermometerdifferenz) zu insgesamt zwölf Moderationsanalysen (Tab. 6-48).

Von den acht berechneten Moderationsmodellen liegt lediglich für die Vorhersage affektiver Einstellungen zu Homosexualität (HOAF) durch IAT Homosexualität eine signifikante Moderatorwirkung der Motivation zur Vorurteilsfreiheit vor (Änderung in  $R^2 = .03$ ;  $p < .05$ ). Das positive Vorzeichen des Moderatorterms zeigt seine zusammenhangssteigernde Wirkung an. Wird die Stichprobe anhand des

Motivationsfaktors in hoch und niedrig motivierte Probanden geteilt<sup>7</sup>, findet sich ein entsprechend unterschiedliches Zusammenhangsmuster zwischen den affektiven Einstellungen zu Homosexualität und dem IAT-Homosexualität. Für die Personen mit geringer Motivation zur Vorurteilstkontrolle beträgt die Korrelation  $r=.24$  ( $p<.05$ ,  $N= 72$ ). Für Personen, die angeben, hoch motiviert zu sein, beträgt der Zusammenhang dagegen  $r=.44$  ( $p<.001$ ,  $N= 84$ ).

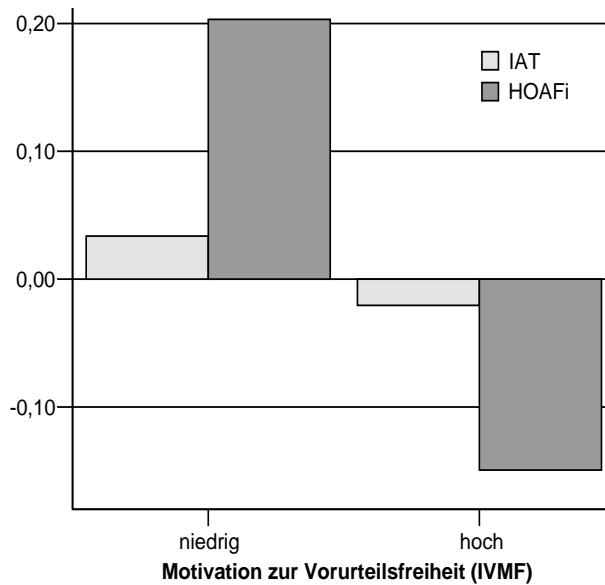
	IAT						EAST							
	R <sup>2</sup>	ΔR <sup>2</sup>	p	Beta	T	p	R <sup>2</sup>	ΔR <sup>2</sup>	p	Beta	T	p		
ÜBKOi	.18	.01	.23	I	.07	.90	.37	.18	.01	.30	I	.09	1.26	.21
				Mo	-.39	-5.31	.00				Mo	-.43	-5.73	.00
				I*Mo	.09	1.21	.23				I*Mo	-.08	-1.05	.30
ÜBAFi	.09	.00	.79	I	.10	1.29	.20	.09	.00	.81	I	.04	.55	.58
				Mo	-.28	-3.62	.00				Mo	-.30	-3.85	.00
				I*Mo	.02	.27	.79				I*Mo	-.02	-.24	.81
Th_NÜ	.05	.01	.29	I	.06	.70	.49	.05	.02	.10	I	.07	.81	.42
				Mo	-.18	-2.30	.02				Mo	-.20	-2.43	.02
				I*Mo	.09	1.07	.29				I*Mo	-.13	-1.68	.10
HOKOi	.15	.02	.09	I	.32	4.19	.00	.06	.01	.28	I	.06	.78	.44
				Mo	-.22	-2.88	.01				Mo	-.22	-2.75	.01
				I*Mo	.13	1.73	.09				I*Mo	-.09	-1.09	.28
HOAFi	.17	.03	.02	I	.37	4.93	.00	.02	.00	.48	I	.02	.22	.82
				Mo	-.16	-2.12	.04				Mo	-.14	-1.79	.08
				I*Mo	.18	2.35	.02				I*Mo	-.06	-.71	.48
Th_HH	.11	.01	.18	I	.31	3.95	.00	.02	.00	.63	I	-.02	-.26	.79
				Mo	-.13	-1.76	.08				Mo	-.15	-1.82	.07
				I*Mo	.10	1.34	.18				I*Mo	.04	.48	.63

**Tabelle 6-48.** Moderatorwirkung Individueller Vorurteilstmotivation (IVMF) auf Implizit-Explizit-Korrespondenz; ÜBKOi/ BAFi// HOKOi/ HOAFi=kognitive/ affektive Einstellungen zu Übergewicht// Homosexualität; Th\_NÜ/ Th\_HH= Gefühlsthermometerdifferenz Normal-Übergewichtige/ Homo-Heterosexuelle; I=implizites Maß, Mo=Moderator IVMF, I\*Mo= Produktterm der Moderation; ΔR<sup>2</sup>= Änderung in R<sup>2</sup> durch I\*M

Ein Mittelwertsvergleich affektiver Einstellungen zu Homosexualität und dem IAT Homosexualität für hoch und niedrig motivierte Personen weist erwartungsgemäß *positivere* affektive Einstellungen zu Homosexualität bei hoch motivierten Personen nach (IVMF niedrig: HOAFi= 2.86; IVMF hoch: HOAFi= 2.63, Abb. 6-42). Dieser Unterschied ist zwar gering, jedoch statistisch bedeutsam (T=2.06, df= 158,  $p<.05$ ). Der IAT-Effekt hoch motivierter Probanden ist ebenfalls kleiner, als der niedrig motivierter Probanden (IVMF niedrig: IATH= .11 ms ( $\cong$  77.01 ms); IVMF hoch: HOAFi= .10 ms ( $\cong$  63.94 ms), wobei dieser Unterschied nicht signifikant ist (T=.44, df= 136.43,  $p=.661$ )<sup>8</sup>.

<sup>7</sup> Da der Faktor IVMF z-standardisiert ist, wird die Stichprobe anhand des Mittelwertes 0 in zwei nahezu gleich große Gruppen geteilt.

<sup>8</sup> Gemäß Levene-Test auf Gleichverteilung der Varianzen liegen signifikante Varianzunterschiede vor (F=4.86,  $p<.05$ ).



**Abbildung 6-42.** Implizite und explizite Einstellungen als Funktion der Motivation zur Vorurteilsfreiheit (z-standardisierte Werte)

Das bedeutet, dass hoch motivierte Personen sowohl positivere explizite, als auch implizite Einstellungen gegenüber Homosexualität haben, was die Ergebnisse von Devine et al. (2002) bestätigt, die zeigt, dass hoch intrinsisch vorurteilsfrei motivierte Personen (IMS) geringere Vorurteilslevel auch in den impliziten Maße aufweisen. Dies wird auf ihrer Fähigkeit zur Selbstregulation zurückgeführt. Hier bestehen deutlich Unterschiede zwischen extrinsischer und intrinsischer Motivation zur Vorurteilsfreiheit. Im Folgenden wird die Differenzierbarkeit beider Orientierungen zur Vorurteilskontrolle hinsichtlich ihrer Bedeutung für die direkte und indirekte Einstellungsmessung untersucht.

### 6.7.1.2 Varianzanalyse

Nach Devine et al. (2002) wird ein Median-Split für die Maße  $IMS_i$  (Med.=4.0) und  $EMS_i$  (Med.=2.5) vorgenommen.  $IMS$  und  $EMS$  sind, wie bereits oben dokumentiert, unkorreliert ( $r=-.09$ ,  $p=.281$ ). Zunächst wird die Bedeutung beider vorurteilsbezogener Motivationen für alle expliziten Maße beider Einstellungsobjekte (kognitive und affektive Einstellungen, Gefühlsthermometerdifferenz) in einer 2 ( $IMS_i$  hoch vs. niedrig) x 2 ( $EMS_i$  hoch vs. niedrig) multivariaten Varianzanalyse (MANOVA) geprüft. Anschließend erfolgt die Untersuchung der Bedeutung von  $IMS$  und  $EMS$  für den Vorurteilsgehalt der Geschichten und schließlich wird der Einfluss der beiden distinkten Motivationen auf die implizite Einstellungsmessung analysiert. Die entsprechenden Mittelwerte und Standardabweichungen sind in Tabelle 6-49 dargestellt.

		Maß	IMS niedrig		IMS hoch	
			EMS niedrig	EMS hoch	EMS niedrig	EMS hoch
			M (SD)	M (SD)	M (SD)	M (SD)
Selbstbericht	Über	ÜBKO <i>i</i>	2.48 (.45)	2.43 (.60)	2.85 (.50)	2.83 (.50)
		ÜBAF <i>i</i>	2.80 (.43)	2.81 (.57)	3.04 (.38)	3.07 (.35)
		Th_NÜ	12.51 (14.05)	15.02 (19.44)	17.29 (16.03)	22.64 (17.64)
	Homo	HOKO <i>i</i>	1.75 (.44)	1.76 (.58)	1.00 (.47)	1.87 (.58)
		HOAF <i>i</i>	2.67 (.65)	2.69 (.86)	2.84 (.61)	2.75 (.73)
		Th_HH	4.71 (11.57)	8.12 (17.32)	8.89 (13.67)	9.79 (12.82)
off. Assoz.	Über	Ü1V <i>i</i>	2.86 (1.14)	2.17 (1.08)	2.26 (1.04)	1.88 (.97)
		Ü2V <i>i</i>	2.23 (1.09)	2.01 (.86)	2.13 (.99)	1.60 (.78)
	Homo	H1V <i>i</i>	1.21 (.55)	1.10 (.33)	1.16 (.47)	1.00 (.00)
		H2V <i>i</i>	1.22 (.39)	1.25 (.53)	1.20 (.65)	1.13 (.28)
IAT	Über	IATÜ	.30 (.17)	.34 (.22)	.30 (.17)	.28 (.16)
		IATerÜ	5.81 (7.43)	7.19 (7.46)	6.36 (7.00)	8.23 (6.83)
	Homo	IATH	.09 (.09)	.12 (.11)	.10 (.11)	.10 (.09)
		IATerH	1.92 (4.87)	1.85 (3.71)	1.89 (2.84)	1.22 (2.95)
EAST	Über	EASTÜ	.00 (.08)	-.01 (.08)	.02 (.07)	.02 (.09)
		EASTerÜ	-3.15 (11.84)	-2.83 (7.27)	.83 (9.55)	-.36 (8.94)
	Homo	EASTH	-.02 (.06)	.02 (.06)	.02 (.06)	.00 (.06)
		EASTerH	2.15 (9.12)	1.32 (9.52)	-.33 (8.13)	1.77 (8.38)

**Tabelle 6-49**, Explizite Einstellungen als Funktion von Internaler und Externale Motivation, ohne Vorurteile zu reagieren (IMS, EMS); M=Mittelwert, SD=Standardabweichung; ÜBER/ HOMO: Einstellungsobjekt Übergewicht/ Homosexualität; ÜBKO*i*/ ÜBAF*i*// HOKO*i*/ HOAF*i* = kognitive/ affektive Einstellungen zu Übergewicht// Homosexualität; Th\_NÜ/ Th\_HH=Gefühlsthermometerdifferenz Normal-Übergewichtige/ Homo-Heterosexuelle, *Fortsetzung nächste Seite*

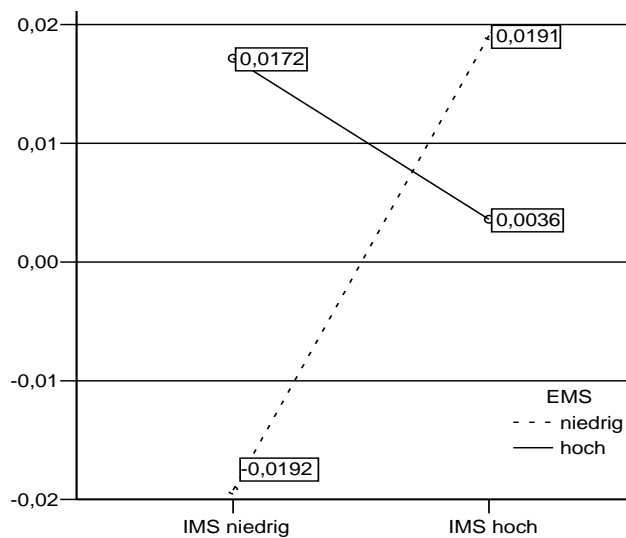
In Bezug auf *explizite Einstellungen* liefert die Varianzanalyse einen Haupteffekt für IMS,  $F(6,151)=4.67$ ,  $p<.000$ , der auf geringere Vorurteilsausprägung in allen expliziten Maße für hoch intrinsisch vorurteilsfrei motivierte Personen im Vergleich zu niedrig intrinsisch vorurteilsfrei motivierte Personen zurückzuführen ist. Mit Ausnahme von HOAF*i* ( $F=.92$ ,  $p=.340$ ) und Th\_HH ( $F=1.71$ ,  $p=.193$ ) sind diese Unterschiede für alle abhängigen Variablen signifikant, wie die Einzeltests ergeben. Weder für EMS, noch die Interaktion von IMS und EMS werden signifikante Unterschiede in den expliziten Einstellungsmaßen beobachtet.

Die Untersuchung von IMS und EMS hinsichtlich des Vorurteilsgehalts der vier *Geschichtenmessungen* zeigt einen Haupteffekt von EMS  $F(4,151)=4.24$ ,  $p<.000$  auf, der sich durch einen höheren Vorurteilsgehalt in den Geschichten gering extrinsisch vorurteilsfrei motivierter Personen ergibt. Die Einzeltests legen den Einfluss von EMS für drei der vier Geschichten offen (nicht signifikant ist die Wir-



kung von EMS für die zweite Bildmessung zum Einstellungsobjekt Homosexualität,  $F=.06$ ,  $p=.814$ ). Außerdem liegt für die erste Geschichtenmessung zum Einstellungsobjekt Übergewicht ein Effekt für IMS vor ( $F=6.80$ ,  $p<.05$ ), der sich durch höheren Vorurteilsgehalt von Personen, die wenig intrinsisch zur Vorurteilsfreiheit motiviert sind. Es wird wiederum keine Interaktion zwischen IMS und EMS beobachtet.

Im Hinblick auf *implizite Einstellungen* bestehen weder für IMS oder EMS, noch für die Interaktion beider Motivationsmaße signifikante Effekte. Die Einzeltests zeigen lediglich einen signifikanten Unterschied im Fehlereffekt des EAST Übergewicht in Abhängigkeit von IMS ( $F=4.40$ ,  $p<.05$ ), der erwar-



**Abbildung 6-43**, EASTH-Effekte (Reaktionszeiten) als Funktion von IMS und EMS

tungskonträr auf höhere EAST-Effekte intrinsisch hoch vorurteilsfrei motivierter Personen zurückgeht, und eine signifikante Interaktion von IMS und EMS für den Reaktionszeiteffekt des EAST Homosexualität ( $F=6.46$ ,  $p<.05$ ). Hohe IMS und niedrige EMS resultiert, völlig entgegen der Erwartungen, in hohen EAST-Effekten, also negativerer impliziter Bewertung homosexueller im Vergleich zu heterosexuellen Personen, während bei geringen Ausprägungen von IMS und EMS sogar negative EAST-Effekte beobachtet werden (siehe Abb. 6-43).

### 6.7.2 Moderatoreffekte der Beziehung zwischen offener Assoziativmessung und expliziten Einstellungen

Die engen Zusammenhänge zwischen dem Vorurteilsgehalt der Geschichten und den expliziten Einstellungen könnten auf der gemeinsamen Beziehung beider Messungen zur Motivation zur Vorurteilskontrolle bzw. sozialer Erwünschtheit beruhen (vgl. Abschnitt 6.8) und würden in diesem Falle keine inhaltliche Basis haben. Zur Analyse dieser Fragestellung wird ein Vergleich der Produkt-Moment-Korrelationen von Fragebogenmaßen und Vorurteilsgehalt der Geschichten mit den für Motivation zur Vorurteilsfreiheit und soziale Erwünschtheit kontrollierten Partialkorrelationen vorgenommen. Dazu wird innerhalb eines Einstellungsobjekts der Zusammenhang zwischen dem gemittelten Vorurteilsgehalt der beiden Geschichten und den drei expliziten Einstellungskriterien für alle Fragebogenmaße motivierter Vorurteilskontrolle und sozialer Erwünschtheit kontrolliert. Für Motivation zur

Vorurteilsfreiheit werden entsprechend der vier individuellen Motivationsmaße interne und externe Motivation zur Vorurteilsfreiheit ( $IMS_i$ ,  $EMS_i$ ), Motivation zu vorurteilsfreiem Verhalten ( $MVV_i$ ) und die Vorurteilstkontrolle ( $VKS_i$ ) betrachtet. Allgemeine soziale Erwünschtheit repräsentieren die Maße der Sozialen Erwünschtheitskala (SES) sowie der Selbst- und Fremdtäuschungsskala gemäß dem Balanced Inventory of Desirable responding (BIDR\_ST und \_FT).

Die Analyse der partiellen Zusammenhänge zeigt eine stabile Beziehung zwischen den expliziten Einstellungen zu Übergewicht und dem Vorurteilsgehalt der Übergewichtsgeschichten. Die Beträge der Korrelationen sinken nur leicht ab und bleiben signifikant (Tab. 6-50). Für das Einstellungsobjekt Homosexualität dagegen sinken die Korrelationen zwischen den Vorurteilen innerhalb der Geschichten und den kognitiven und affektiven Einstellungen gemäß Fragebogen auf ein nicht signifikantes Niveau zurück, wenn für die vorurteilsbezogene Motivation kontrolliert wird. Soziale Erwünschtheit scheint auch hier den Zusammenhang nicht zu moderieren. Relativ stabil bleiben die Zusammenhänge zwischen dem Geschichtenmaß und der Differenz der Gefühlsthermometer für hetero- und homosexuelle Personen. Das bedeutet, für das Einstellungsobjekt Übergewicht beruhen die Zusammenhänge zwischen der Geschichtenmessung und den Fragebogenmaße nicht darauf, dass die Probanden gleichermaßen ihre Antwort gemäß ihrer Motivation zur Vorurteilsfreiheit oder allgemeiner sozialer Erwünschtheit kontrollieren. Für das Einstellungsobjekt Homosexualität liegen bereits weniger enge Zusammenhänge zwischen den Geschichten und den Fragebogenmaßen vor, die bei Kontrolle der Vorurteilsmotivation der Probanden zwar immer noch positiver Natur sind, jedoch deart absinken, dass sie nicht mehr statistisch bedeutsam sind. Das gilt nicht für das Gefühlsthermometermaß.

	Produkt-Moment-Korrelation mit $V_i$	Partielle Zusammenhänge mit $V_i$		
		kontrolliert für: Motivation zur Vorurteilsfreiheit N=153	kontrolliert für: Soziale Erwünschtheit N=154	kontrolliert für: Motivation und Soziale Erwünschtheit N=150
ÜBKO $_i$	<b>.40</b> **	<b>.38</b> **	<b>.37</b> **	<b>.37</b> **
ÜBAF $_i$	<b>.25</b> **	<b>.22</b> **	<b>.22</b> **	.20 *
Th_NÜ	<b>.34</b> **	<b>.31</b> **	<b>.31</b> **	<b>.30</b> **
HOKO $_i$	<b>.21</b> **	.13	<b>.23</b> **	.16 *
HOAF $_i$	.20 *	.14	<b>.20</b> **	.15
Th_HH	<b>.29</b> **	<b>.25</b> **	<b>.27</b> **	<b>.22</b> **

**Tabelle 6-50**, Beziehung zwischen expliziten Einstellungen und Vorurteilsgehalt der Geschichten ( $V_i$ ), kontrolliert für Motivation zur Vorurteilsfreiheit:  $IMS_i$ ,  $EMS_i$ ,  $MVV_i$ ,  $VKS_i$ ; für Soziale Erwünschtheit: SES, BIDR\_ST, BIDR\_FT und gesamt ÜBKO $_i$ / ÜBAF $_i$ // HOKO $_i$ / HOAF $_i$  = kognitive/ affektive Einstellungen zu Übergewicht// Homosexualität; Th\_NÜ/ Th\_HH=Gefühlsthermometerdifferenz Normal-Übergewichtige/ Homo-Heterosexuelle; \*\*:  $p < .01$ ; \*:  $p < .05$

## 6.8 Ergebnisse VI: Das nomologische Netzwerk von Einstellungen

In diesem Abschnitt erfolgt die Analyse des nomologischen Netzwerkes der Einstellungen gemäß Selbstbericht, IAT und EAST sowie offen Assoziationsmessung getrennt nach den Einstellungsmessungen. Zudem wird ggf. mithilfe multipler Regressionsanalysen geprüft, inwieweit die Einstellungsmaße mithilfe der Motivations- und Prädiktormäße *vorhergesagt* werden können. Zur Übersichtlichkeit werden für die korrelativen Analysen teilweise die faktorenanalytisch gewonnenen Motivations- und Prädiktionsfaktoren einbezogen.

### 6.8.1 Bedeutung expliziter Einstellungen

#### 6.8.1.1 Korrelationsanalysen zum nomologischen Netzwerk expliziter Einstellungen

Von insgesamt 147 Korrelationen zwischen den selbstberichteten Einstellungen und den Prädiktormäßen bzw. -faktoren weisen 33 einen signifikanten Betrag auf. Das entspricht mit 22.45% einem Anteil weit über der Zufallswahrscheinlichkeit (Tab. 6-51).

	Motivation			Einstellungsprädiktoren				MRS <i>i</i>
	AVMF	IVMF	EMSF	SEF	RSF	KGF	LSF	
<b>ÜBKO<i>i</i></b>	-.01	<b>-.39</b>	.16	<b>-.22</b>	<b>.44</b>	.10	.04	<b>.45</b>
<b>ÜBKO<i>a</i></b>	<b>-.24</b>	.11	.12	-.02	.09	.02	-.01	.03
<b>ÜBKO</b>	-.17	<b>.42</b>	-.05	.18	<b>-.31</b>	-.07	-.04	<b>-.36</b>
<b>ÜBAF<i>i</i></b>	.02	<b>-.29</b>	.01	-.19	.19	.04	-.01	<b>.29</b>
<b>ÜBAF<i>a</i></b>	-.09	.15	.10	-.07	.01	.00	.02	-.07
<b>ÜBAF</b>	-.08	<b>.35</b>	.06	.11	-.15	-.03	.02	<b>-.29</b>
<b>ThÜb</b>	-.15	<b>.28</b>	-.03	.15	<b>-.21</b>	.02	.08	<b>-.26</b>
<b>ThNo</b>	-.01	.09	.00	.11	.12	.07	.03	-.06
<b>ThUn</b>	-.13	.19	-.06	<b>.27</b>	.06	.03	.03	-.13
<b>Th_NÜ</b>	.13	-.19	.03	-.05	<b>.29</b>	.03	-.05	.20
<b>Th_NU</b>	.12	-.12	.06	-.19	.03	.02	-.01	.09
<b>Th_UÜ</b>	.01	-.06	-.03	.14	<b>.24</b>	.01	-.04	.09
<b>HOKO<i>i</i></b>	.10	<b>-.21</b>	.11	.10	<b>.35</b>	.11	-.03	<b>.45</b>
<b>HOKO<i>a</i></b>	<b>-.27</b>	.07	-.09	.02	.02	.09	.10	-.08
<b>HOKO</b>	<b>-.29</b>	<b>.21</b>	-.15	-.06	<b>-.24</b>	-.01	.11	<b>-.39</b>
<b>HOAF<i>i</i></b>	.10	-.15	.07	.03	<b>.19</b>	.03	-.13	<b>.40</b>
<b>HOAF<i>a</i></b>	-.12	.14	.01	-.06	-.05	-.03	-.01	-.05
<b>HOAF</b>	-.17	<b>.22</b>	-.06	-.06	<b>-.21</b>	-.04	.12	<b>-.41</b>
<b>ThHo</b>	-.16	<b>.25</b>	.03	.05	-.17	.05	.09	<b>-.37</b>
<b>ThHe</b>	-.07	.15	.02	.08	.12	.14	.03	-.08
<b>Th_HH</b>	.11	-.14	-.02	.01	<b>.29</b>	.07	-.08	<b>.33</b>

**Tabelle 6-51**, Nomologisches Netzwerk expliziter Einstellungen, Abkürzungen siehe Tab. 6-28, 6-29, 6-54

Die Bedeutung der antizipierten Vorurteilmotivation (AVMF) beschränkt sich auf die antizipierten kognitiven Einstellungen gegenüber übergewichtigen und homosexuellen Menschen und die resultierenden Differenzen aus antizipierten und individuellen Dispositionen (ÜBKO $\alpha$ :  $r=-.24$ , HOKO $\alpha$ :  $r=-.27$ ,  $p<.001$ ; ÜBKO:  $r=-.17$ ,  $p<.05$ , HOKO:  $r=-.29$ ,  $p<.001$ ). Das erwartungskonform negative Vorzeichen dieser Korrelationen bedeutet eine jeweils geringere Vorurteilsausprägung sowie geringere Unterschiede zwischen individueller und antizipierter Einstellungen, je höher die Motivation zur Vorurteilsfreiheit ist. Wenn eine vorurteilsfrei motivierte Gesellschaft antizipiert wird, fallen die antizipierten Einstellungen weniger negativ aus, und stehen somit zu den allgemein positiveren individuellen Einstellungen in geringerem Kontrast.

Die individuelle Motivation, ohne Vorurteile zu sein (IVMF), korreliert entsprechend signifikant negativ mit den individuellen kognitiven und affektiven Einstellungen, mit Ausnahme affektiver Einstellungen zu Homosexualität (ÜBKO $i$ :  $r=-.39$ , ÜBAF $i$ :  $r=-.29$ ,  $p<.001$ ; HOKO $i$ :  $r=-.21$ ,  $p<.01$ , HOAF $i$ :  $r=-.15$ ,  $p>.05$ ). Die positiven Beziehungen von IVMF zu den Differenzen antizipierter und individueller Einstellungen weisen auf eine zunehmende Diskrepanz zwischen allgemein negativer Gesellschaftsmeinung und den eigenen Einstellung hin, je stärker sich eine Person vorurteilsfrei verhalten möchte (ÜBKO:  $r=.42$ , ÜBAF:  $r=.35$ ,  $p<.001$ ; HOKO:  $r=.21$ , HOAF:  $r=.22$ ,  $p<.01$ ). Außerdem werden für IVMF signifikant positive Korrelationen zu den Gefühlsthermometern für übergewichtige und homosexuelle Personen beobachtet (ThÜb:  $r=.28$ , ThHo:  $r=.25$ ,  $r<.001$ ), ebenso wie für untergewichtige Personen (ThUn:  $r=.19$ ,  $p<.05$ ). Dies bewirkt geringere Bewertungsunterschiede in den Gefühlsthermometern über- und normalgewichtiger Personen (Th\_NÜ:  $r=-.19$ ,  $p<.05$ ), für die Differenz hetero- vs. homosexuelle Personen ist die Korrelation zu IVMF nicht signifikant.

Der Faktor external orientierter Motivation, ohne Vorurteile zu sein (EMSF), steht mit Ausnahme der kognitiven individuellen Einstellungen zu Übergewicht zu keinem weiteren Kriteriumsmaß in Beziehung. Die positive, betragsmäßig jedoch sehr geringe Korrelation zu ÜBKO $i$  ( $r=.16$ ,  $p<.05$ ) weist auf leicht negativere individuelle Einstellungen hin, je stärker die externale Motivation zur Vorurteilsfreiheit ausgeprägt ist. Dieser Zusammenhang ist dann nicht erwartungskonträr, wenn die Einstellungsmessung wenig anfällig für soziale Erwünschtheit ist, da bei hoher EMS lediglich der äußere soziale Druck zu einer motivierten Verringerung der Vorurteilsausprägung führt.

Von den Einstellungsprädiktoren ist es insbesondere „RWA und SDO“ (RSF), der ein konsistentes Beziehungsmuster zu den Einstellungskriterien aufzeigt. Der durch autoritäre und sozial dominanzorientierte Einstellungen geprägte Faktor korreliert entsprechend positiv mit allen individuellen Kriteriumsmaßen, wobei besonders starke Beziehungen zu den *kognitiven* Einstellungen bestehen (ÜBKO $i$ :  $r=.44$ , HOKO $i$ :  $r=.35$ ,  $p<.001$ ; ÜBAF $i$ :  $r=.19$ , HOAF $i$ :  $r=.19$ ,  $p<.05$ ). Je stärker die Ausprägung in RWA und SDO, desto größer sind die Differenzen zwischen antizipierten und individuellen Einstellungen, diese Korrelationen sind bis auf die affektiven Einstellungen zu Übergewichtigen signifikant

(ÜBKO:  $r=-.31$ , HOKO:  $r=-.24$ , HOAF:  $r=-.21$ ). Des Weiteren korreliert RSF negativ mit den Gefühlsthermometern für übergewichtige und homosexuelle Personen (ThÜb:  $r=-.21$ ,  $p<.01$ ; ThHo:  $r=-.17$ ,  $p<.05$ ) und positiv mit den Thermometerdifferenzen Normal- vs. Übergewichtige ( $r=.29$ ,  $p<.001$ ), Unter- vs. Übergewichtige ( $r=.24$ ,  $p<.001$ ) sowie Hetero- vs. Homosexuelle ( $r=.29$ ,  $p<.001$ ). Ein vergleichbares und zudem noch stabileres Beziehungsmuster zu den expliziten Einstellungsmaßen unterhält individueller Modern Racism (MRS $i$ ). MRS korreliert für beide Einstellungsobjekte in erwartungskonformer Richtung mit den kognitiven Einstellungen (ÜBKO $i$ :  $r=.45$ , HOKO $i$ :  $r=.45$ ,  $p<.001$ ), den affektiven Einstellungen (ÜBAF $i$ :  $r=.29$ , HOAF $i$ :  $r=.40$ ,  $p<.001$ ) und den Thermometerdifferenzen (Th\_NÜ:  $r=.20$ ,  $p<.05$ ; Th\_HH:  $r=.33$ ,  $p<.001$ ).

Der varianzstärkste, soziale Erwünschtheit repräsentierende Prädiktionsfaktor (SEF) unterhält ähnliche Beziehungen zu den Vorurteilskriterien wie die individuelle Vorurteilmotivation (IVMF). Beide Faktoren sind thematisch ähnlich und teilen 11.56% gemeinsamer Varianz (siehe Tab. 6-27). Abgesehen von den nahezu durchgehend geringeren Korrelationsbeträgen von SEF und den Einstellungsmaßen, findet sich ein weiter Unterschied zwischen IVMF und SEF hinsichtlich der Bewertung untergewichtiger Personen im Gefühlsthermometer. Diese fällt deutlich positiver aus, je sozial erwünschter ein Proband angibt zu sein ( $r=.29$ ,  $p<.001$ ) im Vergleich zur vorurteilsbezogenen Motivation. Zudem steht die Thermometerdifferenz Normal- vs. Untergewichtige in negativem Zusammenhang zu SEF. In diesen Beziehungen äußert sich vermutlich eine die Gesellschaft prägende Schlankheitsnorm, die somit Relevanz für soziale Erwünschtheit besitzt. Die Faktoren „kognitive Geschlossenheit“ (KGF) und „Liberalismus und Selbsttäuschung“ (LSF) korrelieren mit keinem der Kriteriumsmaße.

#### 6.1.1.1 Regressionsanalysen zur Vorhersage expliziter Einstellungen

Die in Form der Motivations- und Prädiktionsfaktoren betrachteten Variablen im nomologischen Netzwerk expliziter Einstellungen sind nicht unabhängig voneinander. An dieser Stelle soll ihr *gemeinsames* Vorhersagepotential an expliziten Einstellungen betrachtet werden. Als Kriterien liegen für jedes Einstellungsobjekt die individuellen kognitiven und affektiven Maße (z.B. ÜBKO $i$ , ÜBAF $i$ ) sowie die Gefühlsthermometerdifferenz (z.B. Th\_NÜ) vor. Die Analyse der Vorhersagekraft von vorurteilsbezogener Motivation (IMS $i$ , EMS $i$ , MVV $i$ , VKS $i$ ), sozialen und kognitiven Einstellungsdispositionen (RWA, SDO, NFC, UGTS, KONS, SES, BIDR\_ST, BIDR\_FT) und Modern Racism (MRS $i$ )<sup>1</sup> für die expliziten Einstellungen erfolgt demgemäß über jeweils drei schrittweise multiple Regressionen je Einstellungsobjekt (siehe Tab. 6-52 und 6-53).

<sup>1</sup> Als Prädiktoren der expliziten Einstellungen werden die einzelnen Fragebogenskalen anstelle der Faktoren betrachtet, um zusätzlich zur Korrelationsanalyse detaillierte Aussagen über ihre Bedeutung ableiten zu können.

## a) Einstellungsobjekt Übergewichtige Menschen

Die größte Varianzaufklärung wird für die kognitiven Einstellungen zu Übergewicht ( $R^2=34.4\%$ ) erhalten. Die in die Regressionsgleichung aufgenommenen Prädiktoren sind Modern Racism (MRS*i*), Intrinsische Motivation (IMS*i*), ohne Vorurteile zu reagieren, Right-Wing-Autoritarismus (RWA) und die BIDR-Skala Selbsttäuschung (ST). Während höhere Werte in MRS*i* und RWA mit einem Anstieg in ÜBK*O**i* korrespondieren, gehen höhere Werte in IMS*i* und BIDR\_ST mit geringeren Ausprägungen in ÜBK*O**i* einher. Affektive Einstellungen zu Übergewicht werden zu 11.8% durch die Motivation zu vorurteilsfreiem Verhalten und wiederum die BIDR-Skala Selbsttäuschung vorhergesagt. Beide Prädiktoren stehen in negativer Beziehung zum Kriterium. An der Varianz der Gefühlsthermometerdifferenz werden schließlich lediglich 9.1% vorhergesagt. Die aufgenommenen Prädiktoren sind IMS*i* und Konservatismus. Eine höhere intrinsische Motivation, ohne Vorurteile zu reagieren, bedeutet eine geringere Differenz, stärkerer Konservatismus eine größere Differenz der Gefühlsthermometer Normal- und Übergewichtiger.

	AV	R	R <sup>2</sup>	Korr. R <sup>2</sup>	Aufgenommene UV	Beta	T	p
1	ÜBK <i>O</i> <i>i</i>	.587	.344	.327	MRS <i>i</i>	.26	3.44	.001
					IMS <i>i</i>	-.23	-3.20	.002
					RWA	.23	3.23	.002
					BIDR_ST	-.15	-2.15	.033
2	ÜBAF <i>i</i>	.344	.118	.107	MVV <i>i</i>	-.25	-3.28	.001
					BIDR_ST	-.18	-2.33	.021
3	Th_NÜ	.302	.091	.080	IMS <i>i</i>	-.25	-3.29	.001
					KONS	.19	2.44	.016

**Tabelle 6-52**, Schrittweise multiple Regression zur Vorhersage von **Einstellungen zu Übergewicht**; AV= Abhängige Variablen, AV 1: Kognitive Einstellungen, 2: Affektive Einstellungen, 3: Thermometerdifferenz Normal-Übergewichtiger; UV= Unabhängige Variablen: IMS*i*, EMS*i*, MVV*i*, VKS*i*, RWA, SDO, NFC, UGTS, KONS, SES, BIDR\_ST, BIDR\_FT und MRS*i*

## b) Einstellungsobjekt Homosexuelle Menschen

Ähnlich zum Einstellungsobjekt Übergewicht wird für Homosexualität die größte Varianzaufklärung an den *kognitiven* Einstellungen erzielt ( $R^2=26.2\%$ ). Die relevanten Prädiktoren sind wiederum Modern Racism (MRS*i*) und Right-Wing-Autoritarismus (RWA), beide stehen in positiver Beziehung zum Kriterium. Von den affektiven Einstellungen zu Homosexualität wird mit 16.0% ein etwas größerer Varianzanteil aufgeklärt als für die affektiven Einstellungen zu Übergewicht. Hierbei verbleibt von allen unabhängigen Variablen allein MRS*i* als Prädiktor. Von der Differenz der Gefühlsthermometer

Hetero- und Homosexuelle wird deutlich mehr Varianz aufgeklärt im Vergleich zu Th\_NÜ ( $R^2=17.0\%$ ). Modern Racism, RWA und extrinsische Motivation, ohne Vorurteile zu reagieren verbleiben als signifikante Prädiktoren im Regressionsmodell, wobei erstere beiden in positivem Zusammenhang zum Kriterium stehen, während EMS ein negatives Beta-Gewicht aufweist.

AV	R	R <sup>2</sup>	Korr. R <sup>2</sup>	Aufgenommene UV	Beta	T	p
1 HOKO <sub>i</sub>	.512	.262	.253	MRS <sub>i</sub>	.34	4.57	.000
				RWA	.27	3.62	.000
2 HOAF <sub>i</sub>	.400	.160	.155	MRS <sub>i</sub>	.40	5.45	.000
				MRS <sub>i</sub>	.25	3.18	.002
3 Th_HH	.413	.170	.154	RWA	.25	3.04	.003
				EMS <sub>i</sub>	-.16	-2.07	.040

**Tabelle 6-53**, Schrittweise multiple Regression zur Vorhersage von **Einstellungen zu Homosexualität**; AV=Abhängige Variablen, AV 1: wie oben, siehe Tab. 6-52

## 6.1.2 Bedeutung impliziter Einstellungen

### 6.1.2.1 Korrelationsanalysen zum nomologischen Netzwerk impliziter Einstellungen

Im Gegensatz zum Zusammenhangsmuster für explizite Einstellungen sind lediglich 7 von den 136 Korrelationen signifikant. Das entspricht mit 5.55% nicht wesentlich mehr als der Zufallswahrscheinlichkeit bei dieser Anzahl an statistischen Tests. Jedoch sind die Zusammenhänge zum größten Teil sinnvoll interpretierbar (Tab. 6-54).

Aufgrund des geringen Anteils statistisch bedeutsamer Zusammenhänge wird von einer detaillierten Dokumentation der *Vorhersage* der verschiedenen impliziten Einstellungsmaße durch die expliziten Fragebogenvariablen analog zu oben abgesehen. Die regressionsanalytischen Ergebnisse der Vorhersage der Latenz-Fehlerwerte von IAT und EAST zu Übergewicht und Homosexualität werden im Anschluss an die korrelative Untersuchung des nomologischen Netzwerkes impliziter Einstellungen in aller Kürze berichtet.

	IAT				EAST			
	IATÜ	IATerÜ	IATH	IATerH	EASTÜ	EASTerÜ	EASTH	EASTerH
<b>AVMF</b>	-.08	-.12	.00	-.09	-.11	-.10	-.03	.12
<b>IVMF</b>	-.02	.12	-.03	-.01	.14	.18	.07	-.04
<b>EMSF</b>	.07	.02	.07	-.08	-.08	-.03	-.03	.05
<b>SEF</b>	.04	<b>.25</b>	.03	-.07	.19	.06	.16	-.01
<b>RSF</b>	.17	-.05	.07	-.10	-.02	-.03	-.10	-.07
<b>KGF</b>	.02	-.02	.00	-.03	.09	.03	-.01	-.08
<b>LSF</b>	.03	-.04	-.04	-.03	<b>.24</b>	.00	-.04	-.04
<b>MRS</b>	.09	-.06	.16	.01	-.07	.03	-.12	.04
<b>KS-Real</b>	.05	.02	.00	-.11	-.11	-.05	.04	.09
<b>KS-Ideal</b>	.11	.03	.03	-.07	-.03	-.14	.02	.08
<b>KS_diff</b>	.04	.00	.03	.06	.10	-.06	-.03	-.03
<b>BMI</b>	.06	-.04	.07	-.10	-.11	.00	-.06	.15
<b>Sex-O</b>	-.03	.07	<b>.20</b>	.14	-.14	-.05	.08	.13
<b>Sex-ID</b>	.12	.16	.05	.11	.03	-.01	-.05	-.04
<b>FreundÜ</b>	.00	-.01	.11	.05	-.04	-.05	-.06	.06
<b>FreundH</b>	.06	-.01	-.13	.01	.05	-.07	-.13	.04
<b>MWT-IQ</b>	-.05	.02	-.09	.00	.11	.03	.01	-.05
<b>Ekel</b>	-.05	.08	-.03	.13	.14	.00	.00	-.06

**Tabelle 6-54.** Nomologisches Netzwerk **impliziter Einstellungen**, AVMF= Antizipierte-, IVMF= Individuelle-, EMSF= External orientierte Vorurteilmotivation, SEF= Soziale Erwünschtheit, RSF= Autoritarismus und Soziale Dominanzorientierung, KGF= Kognitive Geschlossenheit, LSF= Liberalismus und Selbsttäuschung; MRS=Modern Racism; KS-Real= eigenes Körperschema, KS-Ideal=Körper-idealbild, KS\_diff=Differenz), BMI=Body-Mass-Index; Sex-O= sexuelle Orientierung, Geschlechts-ID= Geschlechtsidentität, FreundÜ/H= Anzahl übergewichtiger/ bi- oder homosexueller Freunde; MWT-IQ=Mehrfachwortwahltest; Ekel=Ekelempfindlichkeit, weitere Abkürzungen siehe Tab. 6-35

Der Faktor soziale Erwünschtheit (SEF) korreliert positiv mit drei impliziten Maßen, was eine Zunahme der impliziten Vorurteile mit höherer sozialer Erwünschtheit anzeigt. Ebenso verhält es sich mit der individuellen Motivation zur Vorurteilsfreiheit (IVMF), die in positivem Zusammenhang zum fehlerbasierten Effekt des EAST-Übergewicht steht ( $r=.18$ ,  $p<.05$ ) und mit dem Faktor Liberalismus und Selbsttäuschung (LSF), der positiv mit dem Reaktionszeiteffekt des EAST-Übergewicht korreliert ( $r=.24$ ,  $p<.01$ ). Die beiden noch nicht beschriebenen Zusammenhänge bestehen zwischen dem IATH und Modern Racism (MRS,  $r=.16$ ,  $p<.05$ ) und der Sexuellen Orientierung (Sex-O:  $r=.20$ ,  $p<.05$ ). Höhere Werte in MRS und eine stärkere heterosexuelle Orientierung gehen mit einer ausgeprägteren impliziten Ablehnung homosexueller Personen bzw. einer relativen Aufwertung heterosexueller Personen gemäß IAT einher.

Aufgrund der geringfügigen Korrespondenzen zwischen impliziten Einstellungen und den faktorisierten Einstellungsprädiktoren werden an dieser Stelle nochmals die Primärkonstrukte hinsichtlich ihrer Bedeutung für implizite Einstellungen betrachtet (Tab. 6-55). Hier wird ein Anteil von 9.4% signifikanter Korrelationen ermittelt.



	IAT				EAST			
	IATÜ	IATerÜ	IATH	IATerH	EASTÜ	EASTerÜ	EASTH	EASTerH
<b>RWA</b>	.15	.00	.13	-.02	.07	-.08	-.05	-.03
<b>SDO</b>	.02	<b>-.25</b>	-.04	<b>-.21</b>	-.09	.03	-.12	-.09
<b>NFC</b>	-.04	.08	.03	.02	.03	-.01	.02	.03
<b>UGTS</b>	.05	.01	.02	-.06	.07	.01	-.05	-.11
<b>KONS</b>	.06	.07	.08	.02	<b>-.22</b>	-.04	-.02	-.02
<b>SES</b>	.05	<b>.28</b>	-.02	-.08	.12	.12	.12	-.02
<b>BIDR_ST</b>	.12	.07	.02	-.01	.19	-.02	-.02	-.03
<b>BIDR_FT</b>	.01	.13	.03	-.02	.20	.07	.18	.00
<b>IMS<i>i</i></b>	-.01	.11	-.07	-.09	.04	.14	.05	-.08
<b>EMS<i>i</i></b>	.04	.03	.10	-.05	-.07	-.09	-.02	.04
<b>MVV<i>i</i></b>	.02	.17	-.05	.10	.08	.15	.08	-.04
<b>VKS<i>i</i></b>	.00	.10	.10	.04	.19	.12	.06	.04

**Tabelle 6-55.** Nomologisches Netzwerk **impliziter Einstellungen**. RWA=Right-Wing Authoritarianism, SDO=Soziale Dominanzorientierung, NFC=Need for Cognition, KONS=Konservatismus, BIDR\_ST=Erwünschtheitsverhalten Selbsttäuschung, BIDR\_FT=Erwünschtheitsverhalten Fremdtäuschung; IMS*i*/ EMS*i*=Internal/ External Motivation to Respond Without Prejudice; MVV*i*= Motivation zu vorurteilsfreiem Verhalten, VKS*i*=Vorurteilstrollskala weitere Abkürzungen siehe Tab. 6-35

Die Reaktionszeiteffekte des IAT stehen zu keinem Prädiktor in signifikanter Beziehung, lediglich die Korrelation zwischen IAT Übergewicht und RWA ist tendenziell bedeutsam ( $r=.15$ ,  $p<.07$ ). Für die IAT-Fehlereffekte sind dagegen einige signifikante Zusammenhänge festzustellen. Für beide Einstellungsobjekte besteht ein negativer Zusammenhang zwischen dem Fehler-IAT und SDO (IATerÜ:  $r=-.25$ ,  $p<.001$ ; IATerH:  $r=-.21$ ,  $p<.01$ ). Das bedeutet, je stärker der Fehlereffekt auf eine Bevorzugung normalgewichtiger bzw. heterosexueller Personen hindeutet, desto weniger sozial dominanzorientiert gibt ein Proband an zu sein. Darüber hinaus korreliert der Fehler-IAT Übergewicht positiv mit SES ( $r=.28$ ,  $p<.001$ ). Diese drei Korrelationen für die Fehlereffekte der IATs erscheinen erwartungskontinär. Im EAST sind es die Reaktionszeiteffekte, insbesondere für Übergewicht, die eine Bedeutung für die Einstellungsprädiktoren haben. Konservatismus korreliert negativ ( $r=-.22$ ,  $p<.01$ ), die Selbst- und Fremdtäuschung des BIDR dagegen positiv zu EASTÜ (ST:  $r=.19$ , FT:  $r=.20$ ,  $p<.05$ ). Außerdem besteht auch zwischen dem Reaktionszeiteffekt des EAST Homosexualität und der Fremdtäuschung gemäß BIDR eine schwach positive Beziehung ( $r=.18$ ,  $p<.05$ ).

Die vorurteilsrelevante Motivation hat keine systematische Bedeutung für die impliziten Einstellungsmaße. Ein jeweils leicht positiver Zusammenhang besteht zwischen dem Fehler-IAT Übergewicht und der Motivation zu vorurteilsfreiem Verhalten (MVV*i*:  $r=.17$ ,  $p<.05$ ) sowie zwischen dem Reaktionszeit EAST Übergewicht und der Vorurteilstrollskala (VKS*i*:  $r=.19$ ,  $p<.05$ ). Diese Korrelationen bedeuten negativere implizite Einstellungen zu Übergewicht bei höherer Motivation zu Vorurteilsfreiheit.

### 6.1.2.2 *Regressionsanalysen zur Vorhersage impliziter Einstellungen*

Regressionsanalytische Untersuchung der Vorhersagbarkeit impliziter Einstellungen untermauern, dass durch die selbstberichteten Einstellungsprädiktoren lediglich an IAT und EAST-Übergewicht<sup>2</sup> signifikante Varianzanteile aufgeklärt werden können (Modelle analog zu Abschnitt 6.8.1.2). Für den IAT-Übergewicht verbleibt allein Soziale Erwünschtheit (SES) als signifikant positiver Prädiktor (Beta: .19,  $T=2.38$ ,  $p=.018$ ) und klärt 3.6% der Varianz impliziter Einstellungen zu Übergewicht gemäß IAT auf. Die Varianz des EAST-Übergewichts-Scores kann zu 7.1% durch die Vorurteilskontrollskala (VKS<sub>i</sub>, Beta=.21,  $T=2.70$ ,  $p=.005$ ) und Konservatismus (Beta=-.17,  $T=2.716$ ,  $p=.032$ ) aufgeklärt werden.

Zusammengenommen deuten die Ergebnisse darauf hin, dass für implizite Einstellungen zum Teil gegenteilige Zusammenhänge zu den betrachteten Prädiktormmaßen als für explizite Einstellungen anzunehmen sind.

### 6.1.3 Bedeutung der Variablen der offenen Assoziativmessung

#### 6.1.3.1 *Korrelationsanalysen zum nomologischen Netzwerk der Geschichtenvariablen*

Mit 32 statistisch bedeutsamen von insgesamt 216 Korrelationen beträgt der Anteil signifikanter Ergebnisse 14.81% und liegt somit deutlich über der Zufallswahrscheinlichkeit von 5% (Tab. 6-56). Für die offene Assoziativmessung werden die über beide Geschichten je Einstellungsobjekt hinweg gemittelten Variablen betrachtet.

An dieser Stelle soll nochmals darauf hingewiesen werden, dass die offene Assoziativmessung mit dem Selbstbericht im Fragebogen prinzipiell das verbale Format und die Kontrollierbarkeit der eigenen Reaktionen teilt, dass jedoch auch nach wie vor deutliche methodische Unterschiede zwischen beiden Messungen bestehen.

---

<sup>2</sup> Gesamtscores gemäß gemittelter z-standardisierter Reaktionszeit- und Fehlereffekte von IAT und EAST

	Übergewicht						Homosexualität					
	ÜbA	ÜbV	ÜbVi	ÜbVa	ÜbK	ÜbS	HoA	HoV	HoVi	HoVa	HoK	HoS
<b>AVMF</b>	-.06	-.06	.00	-.01	.07	-.08	.03	.00	.11	.06	.03	.07
<b>IVMF</b>	.18	.19	<b>-.23</b>	.02	<b>.22</b>	-.15	.17	.00	<b>-.24</b>	.05	.12	-.13
<b>EMSF</b>	.01	.02	-.09	-.05	-.18	-.06	.02	.10	-.01	-.02	-.06	-.02
<b>SEF</b>	.20	.08	-.18	.01	.16	-.10	.24	-.01	<b>-.21</b>	.03	.16	<b>-.23</b>
<b>RSF</b>	-.13	.00	.17	-.05	<b>-.23</b>	.15	-.14	-.07	<b>.22</b>	-.02	<b>-.23</b>	.04
<b>KGF</b>	.05	-.04	.14	.12	-.01	.09	.04	.05	.07	-.01	.11	.06
<b>LSF</b>	-.11	-.17	.00	-.05	-.07	-.08	-.06	-.07	.04	.09	-.06	.09
<b>MRS</b>	<b>-.21</b>	-.14	.13	-.07	-.14	.06	<b>-.22</b>	-.06	<b>.25</b>	-.07	-.11	.06
<b>KS-Real</b>	.08	.02	-.02	-.05	.01	-.03	.01	.01	-.02	-.01	-.04	.18
<b>KS-Ideal</b>	-.08	.02	-.06	-.12	-.01	-.13	-.08	.01	.05	-.10	.00	-.01
<b>KS_diff</b>	-.15	.00	-.03	-.04	-.02	-.06	-.08	-.01	.05	-.07	.05	<b>-.21</b>
<b>BMI</b>	-.08	.02	-.01	-.11	.02	-.08	-.09	.07	.11	-.04	-.04	.18
<b>Sex-O</b>	-.09	-.16	.11	.18	-.14	<b>.21</b>	-.09	-.09	-.05	-.08	-.13	-.16
<b>Sex-ID</b>	-.05	-.04	.12	.04	-.13	.12	-.09	-.08	.07	-.04	<b>-.19</b>	.02
<b>Freund Ü</b>	-.15	.07	<b>-.17</b>	-.03	.03	-.14	-.09	.01	.06	-.08	.00	-.03
<b>Freund H</b>	-.09	.04	-.11	-.14	-.03	-.15	-.04	.10	.06	.02	-.01	<b>.23</b>
<b>MWT-IQ</b>	.06	.02	.04	-.09	.02	-.07	.08	.09	.08	-.01	.04	-.01
<b>Ekel</b>	.09	-.17	.05	.10	-.01	.07	.10	-.08	<b>-.19</b>	-.04	-.07	-.08

**Tabelle 6-56.** Nomologisches Netzwerk der Geschichtenmessung, Ü**ü**=Übergewichtsgeschichten, Ho=Homosexualitätsgeschichten, A=Wortanzahl, V=Valenz, Vi=individuelle Vorurteile, Va=antizipierte Vorurteile, S=Stereotyperkennung, K= Kreativität

Insbesondere Individuelle Vorurteilmotivation (IVMF), Soziale Erwünschtheit (SEF) sowie Autoritarismus und Soziale Dominanzorientierung (RSF) spielen entscheidende Rollen für die offene Assoziativmessung. Vorurteilsfrei motivierte Probanden schreiben längere Geschichten mit geringerem individuellem Vorurteilsgehalt (ÜbA:  $r=.18$ , HoA:  $r=.17$ ,  $p<.05$ ; Ü**ü**Vi:  $r=-.23$ , HoVi:  $r=-.24$ ,  $p<.01$ ). Für die Übergewichtsgeschichten geht die Motivation zur Vorurteilsfreiheit mit einer positiveren Valenz und höherer Kreativität einher (ÜbV:  $r=.19$ ,  $p<.05$ ; Ü**ü**K=.22,  $p<.01$ ). Die individuelle Vorurteilmotivation (IVMF) ähnelt hinsichtlich ihrer Bedeutung für die offene Assoziativmessung prinzipiell der Sozialen Erwünschtheit (SEF), wobei ein stärkeres Beziehungsmuster zwischen IVMF und den Variablen der Übergewichtsgeschichten besteht.

Probanden mit höheren Werten auf sozialer Erwünschtheit schreiben konsistent längere und kreativere Geschichten (ÜbA:  $r=.20$ , HoA:  $r=.24$ ; Ü**ü**K:  $r=.17$ , HoK:  $r=.16$ ,  $p<.05$ ) mit geringerem individuellem Vorurteilsgehalt (ÜbVi:  $r=-.18$ ,  $p<.05$ ; HoVi:  $r=-.21$ ,  $p<.01$ ). Außerdem erkennen sozial erwünschtere Probanden das jeweilige Einstellungsobjekt weniger deutlich bzw. vermeiden sie eine eindeutige Thematisierung, wobei dieser Zusammenhang nur für die Homosexualitätsgeschichten signifikant ist (HoS:  $r=-.23$ ,  $p<.01$ ). Autoritarismus und Soziale Dominanzorientierung (RSF) stehen über beide Einstellungsobjekte hinweg in positivem Zusammenhang zu den individuellen Vorurteilen (ÜbVi:  $r=.17$ ,  $p<.05$ ; HoVi:  $r=.22$ ,  $p<.01$ ) und in negativer Beziehung zur Kreativität der Geschichten

(ÜbK:  $r=-.23$ , HoK:  $r=-.23$ ,  $p<.01$ ). Modern Racism korreliert negativ mit der Anzahl der Geschichten (ÜbA:  $r=-.21$ , HoA:  $r=-.22$ ,  $p<.01$ ) und positiv mit den individuellen Vorurteilen der Homosexualitätsgeschichten (HoVi:  $r=.25$ ,  $p<.001$ ).

Weder die Körperschematavariablen, noch der Body-Mass-Index der Probanden stehen in Zusammenhang zu den Übergewichtsgeschichten. Dafür wird ein leicht positiver Zusammenhang zwischen dem eigenen Körperbild und der Stereotyperkennung der Homosexualitätsgeschichten beobachtet ( $r=.18$ ,  $p<.05$ ). Die sexuelle Orientierung hat dagegen Relevanz für die Übergewichtsgeschichten. Stärkere heterosexuelle Orientierung geht mit negativerer Valenz (ÜbV:  $r=-.16$ ,  $p<.05$ ), höherem antizipierten Vorurteilsgehalt (ÜbVa:  $r=.18$ ,  $p<.05$ ) und deutlicheren Stereotyperkennung einher (ÜbS:  $r=.21$ ,  $p<.01$ ). Inwiefern sich die Probanden mit ihrem eigenen Geschlecht identifizieren steht einzig zur Kreativität der Homosexualitätsgeschichten in Beziehung ( $r=-.19$ ,  $p<.05$ ), insofern als dass stärker identifizierte Personen weniger kreative Geschichten verfassten. Die Anzahl übergewichtiger Freunde und Bekannte korreliert negativ mit ÜbVi ( $r=-.17$ ,  $p<.05$ ), während die Anzahl homosexueller Personen im Freundes- und Bekanntenkreis in positiver Beziehung zur Stereotyperkennung/ -Thematisierung steht ( $r=.23$ ,  $p<.01$ ). Die Ekelempfindlichkeit korreliert negativ mit der Valenz der Übergewichtigkeitsgeschichten ( $r=-.17$ ,  $p<.05$ ). Je ekelempfindlicher die Probanden sind, desto negativer fallen also ihre Übergewichtsgeschichten aus. Der Punktwert im MWT steht zu keiner der Geschichtenvariablen in statistisch bedeutsamer Beziehung.

#### 6.1.3.2 *Regressionsanalysen zur Vorhersage des Vorurteilsgehalts der Geschichten*

Analog zum Vorgehen für explizite Einstellungen wird das Prädiktorpotential von vorurteilsbezogener Motivation, sozialen und kognitiven Einstellungsdispositionen und Modern Racism in Bezug auf den Vorurteilsgehalt der Übergewichts- und Homosexualitätsgeschichten mit schrittweisen linearen Regressionsanalysen untersucht.

##### a) Einstellungsobjekt Übergewichtige Menschen

Der Vorurteilsgehalt der beiden Übergewichtsgeschichten wird zu 14.4% aufgeklärt, und zwar insbesondere durch motivationale Variablen der Vorurteilkontrolle (IMS<sub>i</sub>, EMS<sub>i</sub>) bzw. sozial erwünschter Reaktionen (BIDR\_FT) und durch RWA. Je stärker eine Versuchsperson intrinsisch und extrinsisch motiviert ist, kein Vorurteile zu haben, je eher er oder sie sich zur Fremdtäuschung sozial erwünscht verhält und umso weniger autoritär er oder sie angibt zu sein, desto weniger Vorurteile spiegeln die beiden Geschichten zum Einstellungsobjekt Übergewicht wider (Tab. 6-57).

AV	R	R <sup>2</sup>	Korr. R <sup>2</sup>	Aufgenommene UV	Beta	T	p
ÜbVi	.379	.144	.122	EMS <sub>i</sub>	-.23	-2.91	.004
				BIDR_FT	-.18	-2.30	.023
				RWA	.18	2.28	.024
				IMS <sub>i</sub>	-.16	-2.03	.044

**Tabelle 6-57**, Schrittweise multiple Regression zur Vorhersage des **Vorurteilsgehalts der Übergewichtsgeschichten**; AV=Abhängige Variable, UV= Unabhängige Variablen: IMS<sub>i</sub>, EMS<sub>i</sub>, MVV<sub>i</sub>, VKS<sub>i</sub>, RWA, SDO, NFC, UGTS, KONS, SES, BIDR\_ST, BIDR\_FT und MRS<sub>i</sub>

#### b) Einstellungsobjekt Homosexuelle Menschen

Für die beiden Homosexualitätsgeschichten wird eine Varianzaufklärung von 16.9% erzielt. Den Vorurteilsgehalt dieser Geschichten klären ausnahmslos motivationale Variablen der Vorurteilstkontrolle und sozialen Erwünschtheit auf. Während die Motivation zu vorurteilsfreiem Verhalten und die Fremdtäuschung in negativer Beziehung zum Kriterium stehen, geht im Zusammenspiel mit den anderen Prädiktoren von der intrinsischen Motivation zur Vorurteilsfreiheit ein positiver Einfluss auf den Vorurteilsgehalt der Homosexualitätsgeschichten aus (Tab. 6-58).

AV	R	R <sup>2</sup>	Korr. R <sup>2</sup>	Aufgenommene UV	Beta	T	p
HoVi	.411	.169	.153	MVV <sub>i</sub>	-.35	-4.08	.000
				BIDR_FT	-.26	-3.31	.001
				IMS <sub>i</sub>	.19	2.11	.037

**Tabelle 7-1**, Schrittweise multiple Regression zur Vorhersage **Vorurteilsgehalts der Homosexualitätsgeschichten**; AV=Abhängige Variable, UV= Unabhängige Variablen: IMS<sub>i</sub>, EMS<sub>i</sub>, MVV<sub>i</sub>, VKS<sub>i</sub>, RWA, SDO, NFC, UGTS, KONS, SES, BIDR\_ST, BIDR\_FT und MRS<sub>i</sub>

#### 6.1.4 Resümee

Während für die expliziten Einstellungen zu Übergewicht und Homosexualität motivationale Aspekte und soziale Einstellungen eine Rolle spielen, liegt das Hauptgewicht in der Vorhersage der Geschichtsmessung bei den motivationalen Variablen. Möglicherweise führt die Methodik der Geschichtsmessung, dem nur begrenzt Hinweise über den Sinn der Messung entnommen werden können, gemeinsam mit dem verbalen Format zu einer stärkeren Kontrolle der eigenen Antworten. Die impliziten Einstellungsmaße stehen in keinem systematischen Zusammenhang zu den Fragebogenmaßen der Motivation und sozialen Einstellungen.

## 6.2 Diskussion

Mit Studie III ist ein umfangreicher Datenpool sozialer Einstellungen und weiterer relevanter Konstrukten angelegt worden, um verschiedene zentrale Fragestellungen zur Konstruktvalidierung, Konzeptualisierung und Bedeutung von expliziten und impliziten Einstellungen zu untersuchen. Bevor die Fragestellungen, die Studie III leiteten, resümierend diskutiert werden, seien einige allgemeine Aspekte vorweg gesagt.

### 6.2.1 Voraussetzung für die Überprüfung inhaltlicher Fragestellungen

Eine Grundvoraussetzung für die empirische Untersuchung von interindividuellen Unterschieden in Merkmalsdispositionen und der Analyse von Beziehungen zwischen verschiedenen psychologischen Variablen sind Messverfahren von hinreichender psychometrischer Güte. Während die Fragebogenskalen der Kriteriumsmessung gewohnt reliabel sind und für die Geschichtenvariablen der offenen Assoziativmessung eine exzellente mittlere Interraterreliabilität ermittelt wird, zeichnet sich für die impliziten Maße ein heterogenes und insgesamt recht unbefriedigendes Bild hinsichtlich der psychometrischen Qualitäten ab. Ungeachtet der fraglichen Validität impliziter Maße, deren Untersuchung auch erst mit reliablen Messwerten nachgegangen werden kann, besteht insbesondere für den EAST vordergründig eine *Reliabilitätsproblematik*. Wie bereits De Houwer (2003a) und – um einiges drastischer – Teige et al. (2004) berichten, zeichnen sich die EAST-Effekte durch eine ausgesprochene Unreliabilität aus. Die höchste im Rahmen von Studie III ermittelte Split-Half-Reliabilität beträgt .36 für die Reaktionszeiteffekte des EAST Übergewicht. Dies lässt sich auch nicht durch eine Konstanthaltung der Stimulussequenz im Experiment positiv beeinflussen, was bereits Studie I aufzeigt.

Für die beiden IAT sind deutliche Unterschiede in der im Mittel jedoch akzeptablen Reliabilität festzustellen. Während der IAT-Homosexualität mit .76 zu einer relativ hohen Split-Half-Reliabilität gelangt, die vergleichbar zu der in Studie I ist, erweist sich der IAT Übergewicht mit .60 als deutlich weniger reliabel.

Allgemein unterscheiden sich die vier impliziten Messungen (IAT und EAST für Übergewicht und Homosexualität) in der Reliabilität der Reaktionszeit- vs. Fehlerdaten voneinander. Prinzipiell ist der EAST nicht als reliable Messung einzustufen, womit die Befunde der wenigen hierfür relevanten Publikationen repliziert werden. Die mangelhafte Reliabilität hat nachteilige Auswirkungen auf die nach inhaltlichen Gesichtspunkten vorgenommene Analyse der Maße, da die Reliabilität maßgeblich die Höhe von Zusammenhängen zwischen verschiedenen Messungen bedingt (vgl. Spearman, 1910). Um dieser Tatsache zu begegnen, wurde neben der Konstanthaltung sämtlicher experimenteller Variationsmöglichkeiten (z.B. Stimulussequenz) von vornherein ein Stichprobenumfang angestrebt, der eine

Berücksichtigung des Messfehlers mittels latenter Strukturgleichungsmodellierung ermöglicht. Die Kontrolle der unsystematischen Messfehlervarianz führt jedoch bei einer Reliabilität von .09 (EAST Homosexualität, Reaktionszeiten) zu varianzschwachen und ergo statistisch unbedeutenden latenten Variablen. Anders ausgedrückt kann auch keine latente Datenmodellierung solch gravierende Reliabilitätsmängel, wie sie für den EAST vorliegen, ausgleichen. Entsprechend stellen die EAST-Effekte für die Untersuchung der verschiedenen Fragestellungen weder auf manifester Ebene, noch in den durchgeführten latenten Analysen verlässliche Maße dar. Dies zeigt letztlich wieder, wie hoch allein die Variabilität der psychometrischen Charakteristika innerhalb der impliziten Messungen ist, wobei nicht zuletzt dadurch über differenzierte Validitäten keine gültige Aussage getroffen werden kann. Selbstredend weisen die wenigen statistisch bedeutsamen Korrelationen der EAST-Maße häufig in erwartungskonträre Richtung, was grundlegende Zweifel an der Validität des EAST erheben muss. Die Zusammenfassung von Reaktionszeit- und Fehlereffekte beider Blöcke resultiert in wenigen signifikanten Zusammenhängen zwischen EAST und expliziten Einstellungen, die jeweils nur *über* die beiden Einstellungsobjekte *hinweg* bestehen und zudem zum Teil negativ gerichtet sind (vgl. Tab. 6-43). Wird eine Mittelung der Blöcke innerhalb der Reaktionszeiten bzw. Fehlerdaten vorgenommen, ist lediglich eine signifikante, erwartungskonträr negativ gerichtete Korrelation zwischen dem EAST Homosexualität und den kognitiven Einstellungen zu Übergewicht zu konstatieren (vgl. Tab. 6-46). Darüber hinaus die Korrelation für das Einstellungsobjekt Homosexualität zwischen dem Gesamt-EAST und der entsprechenden Geschichtenmessung ist schwach, jedoch signifikant negativ (vgl. Tab. 6-47). Da alle Einstellungsmaße gleichermaßen gerichtet sind, sollten hier positive Zusammenhänge bestehen. Die mangelnde interne Konsistenz des EAST-Gesamteffekts erschwert allerdings eine inhaltliche Interpretation der Zusammenhangsmuster. Bei der Betrachtung ungemittelter EAST-Effekte auf der Basis der Messblöcke zeichnet sich ein etwas konsistenteres Beziehungsmuster ab, wobei eine vorrangige Bedeutung gesellschaftlich antizipierte Einstellungen gemäß Fragebogen nahe liegt. Dennoch ist bei allen Ergebnissen zum EAST Anlass zur vorsichtigen Interpretation gegeben, da ein Verfahren zur Messung interindividueller Einstellungsdispositionen nicht solche drastischen internen Schwankungen aufweisen sollte.

Aufgrund der mangelnden Reliabilität und der zweifelhaften Validität des EAST müssen die Studie III leitenden Fragestellungen häufig unter Ausschluss dieses Verfahrens verfolgt und resümiert werden. Das ist jedoch weniger tragisch, insofern der EAST neben dem relativ etablierten IAT eines von zwei eingesetzten impliziten Maßen darstellt, das vorrangig zur weiteren Exploration und Evaluation ausgewählt und zusätzlich erhoben wurde. Die Bedeutung des EAST als implizites Messverfahren interindividuell differierender Merkmale wird stellenweise aufgegriffen umfassend in der Gesamtdiskussion besprochen.

Ein ebenso vorweg zu stellender Diskussionsaspekt betrifft die Voraussetzungen der Datenauswertung mit latenter Strukturgleichungsmodellierung, die für die Untersuchung der ersten beiden Fragestellungen von Studie III zentral sind. Die Analyse latenter Variablen setzt bestimmte Dateneigenschaften voraus. Im Wesentlichen handelt es sich bei diesen Voraussetzungen um eine genügend große Stichprobe und multivariat normalverteilte, für Ausreißerwerte bereinigte Indikatorvariablen zur Schätzung messfehlerbereinigter latenter Variablen.

Hinsichtlich der zugrunde liegenden Stichprobengröße sollte berücksichtigt werden, dass die Stichprobe von etwa  $N=160$  gemessen an der Komplexität der hier berechneten MTMM-Modelle möglicherweise grenzwertig ist. Um ein stabiles Alpha-Niveau zu garantieren, empfiehlt Kline (1998) einen Stichprobenumfang von mindestens  $N=200$ . Dennoch, die zur Konstruktvalidierung von Einstellungen durchgeführte MTMM-Analyse liefert relativ stabile und vergleichbare Modellschätzungen. Die Stichprobengröße erlaubt darüber hinaus keine Gruppenvergleiche, z.B. hinsichtlich des Geschlechtes, für die Zusammenhangsanalysen auf latenter Ebene. Dies könnte aber bei der Untersuchung von Einstellungen zu *Homosexualität* und *Übergewicht* von Bedeutung sein, da für beide Einstellungsobjekte geschlechtsbedingte Unterschiede im Ausmaß (expliziter) Vorurteile berichtet werden. So weisen Frauen durchschnittliche negativere Einstellungen zu Übergewicht auf (Crandall & Biernat, 1990; Tiggemann & Rothblum, 1988), während für Männer im Mittel negativere Einstellungen zu Homosexualität zu verzeichnen sind (Hayes, 1995; Kite & Whitley, 1998, vgl. Kapitel 3.2.2). Damit ist eine mögliche Konfundierung der Einstellungsmessung mit der Variablen Geschlecht zu berücksichtigen (vgl. Pearl, 2000), die aufgrund des unausgewogenen Geschlechterverhältnisses und der bereits diskutierten Stichprobengröße nicht hinreichend kontrolliert werden kann. Die eigentliche Konstruktvalidierung konvergenter und diskriminanter Validität von expliziten und impliziten Einstellungen zu Übergewicht und Homosexualität sollte jedoch von diesen Geschlechtsunterschieden nicht dramatisch beeinflusst werden, da sie letztlich nur Mittelwertsdifferenzen darstellen.

Neben einem ausreichenden Stichprobenumfang zur Sicherung des Alpha-Niveaus stellen die Schätzalgorithmen der Strukturgleichungsmodelle einige Verteilungsanforderungen an die Indikatorvariablen. Diese sind jedoch in Studie III nicht global gegeben. Für die meisten Indikatoren sind Abweichungen von der Normalverteilung und signifikante Schiefe- und Kurtosis-Statistiken festgestellt worden, somit liegt eine Verletzung der Normalverteilungsvoraussetzungen vor. Dies beeinträchtigt die Gültigkeit der Modellschätzung. Bei nicht gegebener Normalverteilungsannahme, liefert die Maximum-Likelihood-Schätzung häufig zu große  $\chi^2$ -Werte (Hoyle, 1995). Daraus resultiert ein erhöhtes Alpha-Fehlerrisiko. Da jedoch bei der Strukturgleichungsmodellierung nicht die  $H_1$ , sondern die  $H_0$  die Wunschhypothese ist, werden die Zusammenhangshypothesen mit nicht normal verteilten Daten *konservativ* getestet, was im Sinne der Vermeidung von Fehlern erster Ordnung ist. Da ungeachtet dessen die Modellschätzungen trotzdem mit gebührender Behutsamkeit zu behandeln und mit anderen Mo-



dellen zu vergleichen sind, bedeutet die Verletzung der Normalverteilungsannahme eine unerwünschte Einschränkung für die Datenauswertung. Für Reaktionszeiten ist das Problem nicht normal verteilter Daten aufgrund geringer Anteile extrem langsamer bzw. extrem schneller Reaktionen und einer insgesamt rechtsschiefen Verteilung bekannt (vgl. Greenwald, McGhee & Schwartz, 1998). Zur Normalisierung der Verteilungen werden Reaktionszeitschranken festgelegt, Ausreißer ausgeschlossen und die Reaktionszeiten häufig logarithmiert, dies führt jedoch – wie hier anschaulich wird – nicht notwendigerweise zu normalverteilten Daten. Nicht nur für die Maße der impliziten Verfahren liegen häufig Voraussetzungsverletzungen vor, sondern auch für die Variablen aus Selbstbericht und Geschichtsmessung. Dies hat mit hoher Wahrscheinlichkeit inhaltliche Ursachen, da sozial sensitive Themen angesprochen sind, zu denen nur wenige Personen entgegen der sozialen Norm stark negative Einstellungen zum Ausdruck bringen. Insbesondere für Bildungsstichproben, wie sie auch Studie III zugrunde liegt, werden eher mit liberalen Einstellungen berichtet (z.B. Six, 1996), wobei anzunehmen ist, dass diese geringeren Vorurteilsausprägungen nicht allein positivere Einstellungen zu Minderheiten reflektieren, sondern auch auf eine bildungsbedingt bessere Durchschaubarkeit und Kontrollierbarkeit von Selbstberichtsmessungen zu sozialen Einstellungen zurückgeführt werden können.

Zusammenfassend ist bei der Beantwortung der Fragestellungen von Studie III ggf. zu berücksichtigen, dass für die EAST-Daten eine gravierende Reliabilitäts- und ergo Validitätsproblematik besteht, dass die Verteilungsvoraussetzungen für die Konstruktvalidierung mit Strukturgleichungsmodellen nicht für alle Variablen gegeben sind und der Stichprobenumfang verhältnismäßig klein ist. Um erstgenanntem Problem zu begegnen, wird in der Hauptsache der IAT als Vertreter der impliziten Einstellungsmessungen betrachtet, während auf den EAST weniger Gewicht gelegt wird. Den Verletzungen der Normalverteilungsvoraussetzungen zufolge ist mit einer überzufälligen Ablehnung der Strukturgleichungsmodelle nach  $\chi^2$ -Statistik zu rechnen, wobei dies einerseits im Sinne konservativen Testens ist und zudem prinzipiell neben dem  $\chi^2$ -Wert auch andere Fit-Indizes, wie RMSEA oder NFI, als Evaluationskriterium herangezogen werden sollten (vgl. Bagozzi & Yi, 1999; McDonald & Ho, 2002). Die Stichprobengröße von etwa  $N=160$  ist lediglich für die Strukturgleichungsmodellierung kritisch. Aufgrund der verschiedenen Kreuzmodellierungen zur MTMM-Analyse, die allesamt nachvollziehbare und vergleichbare  $\chi^2$ -Statistiken aufweisen, sind potentielle Instabilitäten im Alpha-Niveau bei zu kleinen Stichproben jedoch als weniger problematisch zu bewerten. Gemäß den spezifischen Stichprobencharakteristika wird schließlich eine verallgemeinernde Interpretation der Ergebnisse mit den entsprechenden Einschränkungen vorgenommen.

## 6.2.2 Die Fragestellungen von Studie III

### 6.2.2.1 Fragestellung I: Konstruktvalidierung von Einstellungen

Die Unterscheidung zwischen expliziten und impliziten Einstellungen (Banaji & Greenwald, 1995) wurde insbesondere empirisch vorangetrieben. Theoretische Konzeptualisierungsansätze sind erst in der Nachfolge empirischer Untersuchungen diskutiert worden (vgl. Fazio & Olson, 2003). Für den mittleren Zusammenhang zwischen Fragebogenmaßen und entsprechenden IAT-Effekten wurde in einer Metaanalyse ein Wert von .24 ermittelt; gleichzeitig ist eine enorme Streubreite der Korrelationen bis hin zu Beträgen mit negativem Vorzeichen festgestellt worden (-.25 bis .60; Hoffman et al., 2003). Auf theoretischer Ebene stehen sich entsprechend die beiden Konzeptualisierungsansätze von expliziten und impliziten Einstellungen als Aspekte *ein und derselben* Disposition (Same-Construct-Approach, Wittenbrink et al., 1997) bzw. als *unabhängige* Konstrukte (Dissociation-Approach, Devine, 1989) gegenüber. Inwiefern Konvergenzen bzw. Divergenzen zwischen expliziten und impliziten Einstellungen gegenständliche Korrespondenzen bzw. Divergenzen reflektieren oder messmethodische Ursachen haben, ist dabei jedoch noch weitestgehend ungeklärt. Bei Betrachtung der in vielerlei Hinsicht stark voneinander abweichenden Methoden zur Erfassung von expliziten und impliziten Einstellungen fällt es sogar schwer, enge Zusammenhänge zwischen selbstberichteten Einstellungen und kognitiven Assoziationsstärkemaßen zu erwarten. Hier sind neben der inhaltlichen Varianz interindividuell differierender Einstellungsdispositionen und der allgegenwärtigen Messfehlervarianz, die insbesondere für implizite Maße kritisch ist (Cunningham, Preacher & Banaji, 2001), substanzielle Varianzanteile der spezifischen Messmethode zu erwarten. Während die inhaltliche Varianz bei gegebener Validität der Messungen zu einer Konvergenz von selbstberichteten Einstellungen und Assoziationsstärkemaßen führen sollte, wirken die methodenspezifische Varianz und die unsystematische Messfehlervarianz statistisch begründet einer Konvergenz entgegen. Um die inhaltlich bedingte Korrespondenz von Einstellungen untersuchen und quantifizieren zu können, muss also einerseits eine Kontrolle des Messfehlers erfolgen und andererseits eine Trennung zwischen Konstruktvarianz und Methodenvarianz. Dies leistet eine latente Multitrait-Multimethod-Analyse.

Da eine rein theoretische Konzeptualisierung von Einstellungen keinen Anspruch auf empirische Validität hat, sollte eine entsprechend angelegte empirische Untersuchung Aufschluss über die messmethodischen Aspekte der Beziehung zwischen expliziten und impliziten Einstellungen geben. Dementsprechend war Studie III angelegt, eine *latente* Multitrait-Multimethod-Analyse zur Konstruktvalidierung expliziter und impliziter Einstellungen vornehmen zu können. Die Fragestellung dazu lautete:

- (1) *Inwiefern kann für explizite und implizite Einstellungen zu Übergewicht und Homosexualität konvergente und diskriminante Validität angenommen werden?  
Wie bedeutsam sind Effekte der Einstellungsdisposition (Traitvarianz) im Verhältnis zu methodenspezifischen Effekten der expliziten, impliziten und offenen Assoziativmessung (Methodenvarianz)?*

Eine Konstruktvalidierung konvergenter und diskriminanter Validität erfordert ein spezielles Datendesign. Im vorliegenden Fall wurde eine 2-mal-4-Struktur der Daten analysiert, wobei die beiden Einstellungsobjekte Übergewicht und Homosexualität die Inhalte bezeichnen (*Traits*), die mit vier Messmethoden erfasst wurden, nämlich Selbstbericht, Assoziativmessung, IAT und EAST (*Methoden*). Vor der Betrachtung der Ergebnisse der MTMM-Analyse sollen einige grundlegende Aspekte der Indikatorwahl für die MTMM-Analyse besprochen werden.

Jede Messmethode liefert zunächst unterschiedliche Maße je Einstellungsobjekt. Gemäß Selbstbericht sind jeweils ein Maß kognitiver und affektiver Einstellungen, sowie eine Gefühlsthermometermessung verfügbar. Von den kodierten Geschichtenvariablen sind ebenfalls mehrere Maße als einstellungsrelevant ausgewiesen worden, allen voran der Vorurteilsgehalt der Geschichten, darüber hinaus die Valenz und die Wortanzahl. Die beiden impliziten Messungen liefern jeweils Reaktionszeit- und Fehlereffekte. Für die Modellierung der Daten zur Konstruktvalidierung muss also notwendigerweise eine Auswahl geeigneter Einstellungsindikatoren getroffen werden. Aus den für jede Methode separat betrachteten Messmodellen resultierte die Entscheidung für kognitive Einstellungen gemäß Fragebogen (Methode „explizit“), für den Vorurteilsgehalt gemäß Geschichtenmessung (Methode „offen“) sowie für die Fehlereffekte gemäß IAT und EAST (Methode „implizit“). Über diese ausgewählten Einstellungsindikatoren hinaus bestehen natürlich weitere, prinzipiell mögliche Datenkombinationen; von einer Analyse sämtlicher denkbaren Konstellationen wurde aber aufgrund der daraus resultierenden Unübersichtlichkeit und des relativ geringen Informationsgewinns abgesehen. Nicht nur aufgrund der Messmodelle, sondern auch zur Vergleichbarkeit der Maße innerhalb der impliziten Methodengruppe wurden für die erste MTMM-Analyse die Fehlerdaten aus IAT und EAST herangezogen. Zwar ist diese Entscheidung datengeleitet getroffen worden, jedoch bedeutete die Wahl der Fehlereffekte an dieser Stelle eine deutliche Reliabilitätseinbuße für die Indikatoren der IAT-Faktoren beider Einstellungsobjekte. Sowohl für Übergewicht, als auch für Homosexualität erweisen sich die Reaktionszeiteffekte des IAT deutlich reliabler als die Fehlereffekte. Da die Differenzen der Reaktionszeiten auch das konventionell betrachtete IAT-Maß darstellen, sind die Modelle zusätzlich für Reaktionszeitdaten geprüft worden, woraus jedoch keine gravierenden Abweichungen zu den fehlerbasierten Indikatoren resultierten.

Konstruktvalidität im klassischen Sinne ist gegeben, wenn inhaltlich übereinstimmende Maße verschiedener Messmethoden jeweils in einem spezifischen Konstruktfaktor (Trait) konvergieren und unterschiedliche Traits diskriminant valide sind, das heißt möglichst geringe Korrelationen untereinander aufweisen. Die Konstruktvalidität von Einstellungen wäre entsprechend anzunehmen, wenn bei Kon-

trolle der Methoden- und Messfehlervarianz eine Konvergenz expliziter und impliziter Einstellungsmaße zu Übergewicht und Homosexualität in jeweils einem Einstellungsfaktor bestünde, die untereinander gering bis moderat korreliert sind, zumindest aber nicht einen gemeinsamen Faktor repräsentieren. Wünschenswert für eine latente Konstruktvalidierung sind prinzipiell substanzielle Trait-Effekte, die in ihrer Bedeutung die Methodenfaktoren überragen. Bereits bei Betrachtung der manifesten Multitrait-Multimethod-Korrelationen (vgl. Tab. 6-38) zeichnet sich jedoch ein entgegen gerichtetes Verhältnis von Trait- und Methodenbedeutung ab. Ungeachtet des gemessenen Einstellungsobjektes (Übergewicht vs. Homosexualität) bestehen hohe Korrelationen zwischen den Maßen, die aus der gleichen Messmethode resultieren. Für die dominierende Bedeutung der Methodeneffekte liefert die latente MTMM-Analyse schließlich eine Bestätigung.

Um die Konstruktvalidierung expliziter und impliziter Einstellungen vorzunehmen, wurde zunächst ein systematischer Modellvergleich angelegt an Widaman (1985) durchgeführt. Im daraus resultierenden vollständigen Modell (2 Trait-, 4 Methodenfaktoren) trat, wie zu erwarten war, eine Out-of-Range-Schätzung auf. Dies wird als „Heywood-Case“ bezeichnet und ist ein bekanntes Problem der MTMM-Analyse via CFA-Ansatz, bei der das vollständige Modell unteridentifiziert ist (z.B. Kenny & Kashy, 1992, Marsh, 1989). Da also auch hier ein Heywood-Case eine allgemeingültige Interpretation des MTMM-Modells erschwerte, wurden zwei Alternativmodelle nach Eid (2000) und Eid et al. (2003) berechnet, in denen jeweils N-1 Methodenfaktoren angenommen werden, wobei hier der Selbstbericht als nicht modellierte Referenzmethode definiert wurde. Im Single-Indicator CT-C(M-1) Modell, das globale Methodenfaktoren annimmt, werden jedoch keine signifikanten Varianzschätzungen für die modellierten Methodenfaktoren für offene Assoziativmessung und implizite Messung erzielt. Dagegen zeigt sich im Multiple-Indicator CT-C(M-1) Modell, das bis auf die EAST-Faktoren signifikante Varianzschätzungen für die traitspezifischen Methodenfaktoren liefert, dass die Verallgemeinerung der Methodenfaktoren über beide Einstellungsobjekte nicht angeraten ist. Infolgedessen wurde ein State-Modell der Einstellungen überprüft, das eine reine Modellierung messfehlerbereinigter korrelierender Einstellungsmaße darstellt, entsprechend „minderungskorrigierte“ Korrelationen<sup>3</sup> liefert und zu einem besseren Fit gelangt als die vorhergehenden MTMM-Modelle. Die zunächst eher aus technischen Gründen als „State“-Modell bezeichnete Zusammenhangshypothese berücksichtigt den möglicherweise in der Tat relevanten Anteil situativer Varianz in den Messungen, denn zwischen den verschiedenen Maße liegt zum Teil ein einwöchiger Abstand. Dennoch sollte Konstruktvalidität von Einstellungen mit einer gewissen Stabilität einhergehen, weshalb der Aspekt situativer Einflüsse auf die Messungen an dieser Stelle noch vernachlässigt und erst weiter unten diskutiert wird (siehe auch Gesamtdiskussion, Abschnitt 7.2.2).

---

<sup>3</sup> Terminus nach Spearman (1910)

Die oben bereits diskutierten Reliabilitätsschwächen der EAST-Effekte zeichnen sich bei der Strukturgleichungsmodellierung der latenten MTMM-Analyse in durchweg geringen Konvergenzen der betreffenden Indikatorvariablen auf die latenten EAST-Faktoren ab. Hierbei spielt die Gestalt des Messmodells keine Rolle, selbst im State-Modell, der sparsamsten und dabei bestangepassten Datenmodellierung erweisen sich die EAST-Daten als mangelhaft konvergent. Insofern ist es auch wenig verwunderlich, dass für die beiden impliziten Maße IAT und EAST kein gemeinsamer impliziter Methodenfaktor anzunehmen ist. Es bleibt an dieser Stelle jedoch offen, ob dieser Befund auf inhaltliche oder psychometrische Ursachen zurückzuführen ist. Aus diesem Grund wurde der EAST nach einer ersten gemeinsamen Modellierung von der MTMM-Analyse ausgeschlossen und eine erneute MTMM-Analyse mit modifizierten Ausgangsbedingungen durchgeführt. Neben dem Ausschluss des EAST als Einstellungsmessung, wurden gemäß den Konventionen nun die reaktionszeitbasierten IAT-Effekte betrachtet, für die zudem eine befriedigendere Reliabilität zumindest eine hinreichende Konvergenz beider Indikatoren sichert. Im Sinne der Konzeptualisierung expliziter und impliziter Einstellungen wurden an dieser Stelle außerdem mögliche Respezifikationen des klassischen MTMM-Modells vorgenommen, um die konkurrierende Bedeutung von Einstellungs- bzw. Methodenfaktoren zu prüfen.

Das Multiple-Indicator CT-C(M-1) Modell kann zunächst als empirisch bedingtes MTMM-Modell der Wahl für die Daten in Studie III betrachtet werden. Abgesehen davon, dass vor allem für Übergewicht keine allgemeine Konvergenz der Indikatorvariablen zum Einstellungstrait gegeben ist und der Methodenfaktor für die Übergewichtsgeschichten nicht signifikant ist, sind die Modellschätzungen sinnvoll interpretierbar. Da eine Konstruktvalidierung von Einstellungen vorgenommen werden sollte, hat die konfirmatorische MTMM-Modellierung im Vergleich zu einer simplen Messfehlerbereinigung (State-Modell) für den informativen Vergleich minderungskorrigierte MTMM-Korrelationen (Campbell & Fiske, 1959) entscheidende Vorteile: Die Anteile von Trait- im Vergleich zu Methoden- und Fehlervarianz sind abschätzbar, was Rückschlüsse auf die relative Bedeutung von inhaltlicher und methodischer Varianz für die betrachteten Variablen erlaubt. Im vorliegenden Fall wird demgemäß ein herausragender Einfluss methodischer Varianz im Vergleich zu inhaltlicher Einstellungsvarianz aufgezeigt. Die latenten Einstellungsfaktoren Übergewicht und Homosexualität korrelieren moderat positiv. Dies spiegelt eine allgemeine Tendenz der Ablehnung sozialer Minderheiten in *expliziten* Einstellungen wider. Da die Traitfaktoren in CT-C(M-1)-Modellierungen als True-Scores der Referenzmethode – hier Selbstbericht – definiert sind, repräsentieren die Einstellungstrait in der gegenwärtigen Analyse die expliziten Einstellungen zu Übergewicht bzw. Homosexualität. Die modellierten Methodenfaktoren dagegen sind als Residuen der Referenzmethode zu verstehen. Insofern zeigt die relativ hohe Varianz der IAT-Methodenfaktoren und die geringe Konvergenz der IAT-Indikatoren zu inhaltsbedingten Einstellungsfaktoren an, dass die Referenzmethode Selbstbericht einen sehr geringen Varianzanteil der IAT-Messung erklärt. Ganz anders verhält es sich mit der offenen Assoziativmessung, deren Metho-

denfaktoren zum Teil nur halb so varianzstark wie die beiden IATs sind, während die Konvergenz der Geschichtenindikatoren mit den Einstellungstraits vergleichsweise höher ist. Das bedeutet, Selbstbericht und offene Assoziativmessung teilen einen substanziellen Anteil an inhaltlicher und methodischer Varianz (siehe Gesamtdiskussion, Abschnitt 7.1).

Ausgehend von dem am wenigsten restringierten State-Modell sind in einem weiteren Schritt theoretisch entwickelte Respezifikationen im Sinne der Einstellungskonzeptualisierung überprüft worden. Dabei erwies sich ein State-Modell mit einem expliziten und einem impliziten Second-Order-Methodenfaktor als vergleichsweise beste Datenrepräsentation unter den modifizierten Modellen. Der explizite Second-Order-Faktor subsumiert die Einstellungen zu Übergewicht und Homosexualität gemäß Fragebogen und Geschichtenmessung. Der implizite Second-Order-Faktor umfasst entsprechend die IAT-Faktoren erster Ordnung für Übergewicht und Homosexualität. Beide Methodenfaktoren korrelieren geringfügig positiv ( $r=.11$ , n.s.). Dieses Modell stellt jedoch keine bessere Anpassung an die empirischen Daten darstellt, als das einfache State-Modell, nichtsdestotrotz repräsentiert es von den theoretisch plausiblen Modellierung die beste Datenanpassung.

Der MTMM-Analyse zufolge kann also keine Konstruktvalidität für explizite und implizite Einstellungen im klassischen Sinne angenommen werden. Für das Einstellungsobjekt Homosexualität kann zwar eine inhaltliche Konvergenz der drei Einstellungsmaße aus expliziter, impliziter und offener Assoziativmessung angenommen werden, jedoch weisen die IAT-Indikatoren auch für dieses Einstellungsobjekt einen Überhang an methodenspezifischer bzw. nicht durch die Einstellungsfaktoren erklärter Varianz auf. Die Konzeptualisierung von expliziten und impliziten Einstellungen als Ausdrücke ein und derselben zugrunde liegenden Einstellungsdisposition z.B. gemäß *Same-Construct-Approach*, kann demzufolge nach nicht untermauert werden. Dagegen besteht zwischen der Fragebogenmessung und der offenen Assoziativmessung eine relative enge Korrespondenz. Ungeachtet der sehr unterschiedlichen Messmethoden findet sich eine positive Beziehung zwischen dem kodierten Vorurteilsgehalt der Geschichten und der relativen Ablehnung Übergewichtiger bzw. Homosexueller im Selbstbericht, wobei die Beziehung für das Einstellungsobjekt Übergewicht deutlich stärker ist, als für Homosexualität. Dies wird anhand der vergleichsweise geringen Pfadgewichte der Indikatoren für die Homosexualitätsgeschichten an einem gemeinsamen Einstellungstrait mit expliziten und impliziten Einstellungen zu Homosexualität deutlich und ist möglicherweise auf die relative Schwierigkeit zurückzuführen, den Stereotypen Homosexualität in den betreffenden Bildmotiven zu erkennen. Nichtsdestotrotz reflektieren Selbstbericht und Geschichtenmessung eher eine gemeinsame Einstellungsdisposition, als Selbstbericht und IAT oder Geschichtenmessung und IAT.

Der latenten MTMM-Analyse zufolge sind also explizite und implizite Einstellungsmaße nicht ausreichend konvergent valide. Außerdem wird deutlich, dass die methodenspezifische Varianz expliziter und impliziter Einstellungsmessung einflussreicher ist, als die inhaltliche Varianz der Einstellungsmes-

sung, die gleichermaßen im Grad der Ablehnung übergewichtiger bzw. homosexueller Menschen zum Ausdruck kommen sollte. Für Einstellungen zu Homosexualität besteht ein leichter Zusammenhang zwischen IAT und Fragebogen, nicht jedoch für Einstellungen zu Übergewicht. Unter der Annahme einer wahren Konvergenz expliziter und impliziter Einstellungsmessung eröffnen sich dieser Stelle nun zwei unterschiedliche Interpretationswege. Prinzipiell könnten fehlende bzw. inkonsistente Beziehungen zwischen expliziten und impliziten Einstellungen auf den moderierenden Einfluss von Drittvariablen wie der Motivation zur Vorurteilsfreiheit (z.B. Plant & Devine, 1998) zurückzuführen sein. Dies wurde mit Fragestellung 4 untersucht.

Darüber hinaus muss die Durchführung der Datenerhebung als mögliche Ursache der Divergenz expliziter und impliziter Einstellungen speziell zu Übergewicht diskutiert werden. Dies betrifft die oben bereits angedeutete Unterscheidung zwischen überdauernden und situativ variierenden Varianzanteilen der Einstellungsmessungen. Der Durchführung der Datenerhebung von Studie ist geschuldet, dass für jeden Probanden zwischen der Erhebung des EAST und IAT Übergewicht und des Selbstberichts zu Einstellungen zu Übergewicht ein Intervall von sieben Tagen lag, während EAST und IAT Homosexualität zusammen mit dem zuallerletzt vorgelegten Selbstbericht in einer Sitzung erhoben wurden (vgl. Abschnitt 6.2.3). Diese situative Variabilität könnte eine Ursache für die Dissoziation expliziter und impliziter Einstellungen zu Übergewicht sein, was einer Instabilität der impliziten Messung zu Lasten gelegt wird, da für explizite Einstellungsmessungen deutlich höhere Stabilitäten (Test-Retest-Reliabilitäten) dokumentiert sind als für implizite Messungen via IAT (z.B. Cunningham, Preacher & Banaji, 2001; Steffens & Buchner, 2003, siehe Gesamtdiskussion, Abschnitt 7.2).

Die Analyse der globalen diskriminanten Validität von Einstellungen zu Übergewicht und Homosexualität erübrigt aufgrund der fehlenden Konvergenz der verschiedenen Einstellungsindikatoren nahezu. Die moderaten Korrelationen der latenten Einstellungstraits weisen jedoch auf diskriminante Validität expliziter Einstellungen zu Übergewicht bzw. Homosexualität hin. Entsprechend der simultan geprüften Vorhersage von expliziten Einstellungen zu Übergewicht und Homosexualität mit beiden IATs kann außerdem diskriminante Validität für impliziten Einstellungen gemäß IAT angenommen werden, da die beiden Kreuzpfade der Vorhersage nicht statistisch bedeutsam sind. Das heißt, der IAT Übergewicht sagt explizite Einstellungen zu Homosexualität ebenso wenig vorher, wie der IAT Homosexualität explizite Einstellungen zu Übergewicht. Jedoch besteht unerwünschter Weise auch eine eindeutig Dissoziation zwischen expliziten und impliziten Einstellungsmaße zu Übergewicht, da hier ebenso wenig *innerhalb* des Einstellungsobjektes ein Zusammenhang zwischen IAT und Selbstbericht besteht, wie *zwischen* beiden Einstellungsobjekten (siehe unten).

### 6.2.2.2 Fragestellung II: Implizit-Explizit-Dimension der Einstellungen

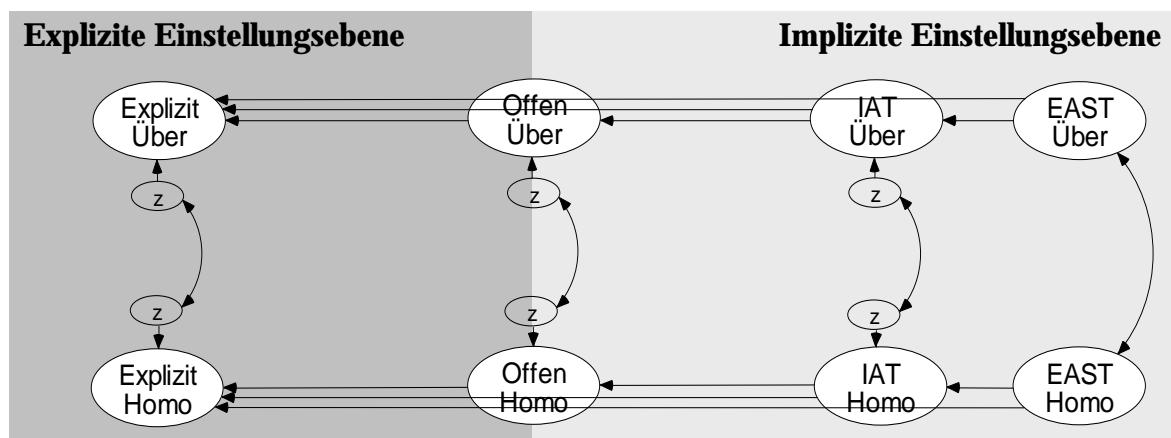
Die Unterscheidung zwischen expliziten und impliziten Einstellungen legt eine Art Kontinuum nahe (siehe Abb. 1-1), dessen Pole durch direkte und indirekte Einstellungsmessungen charakterisiert sind und vorrangig hinsichtlich Bewusstheit und Steuerbarkeit der Einstellungen bzw. der Resultate direkter und indirekter Einstellungsmessungen einander gegenüberstehen. Da eine Reinheit der Messung hinsichtlich expliziter und impliziter Einstellungsanteile weder für den Selbstbericht, noch für die gängigen impliziten Maße angenommen werden kann (vgl. De Houwer, in press), handelt es sich hierbei um ein theoretisches Kontinuum praktisch ungewissen Ausmaßes. Ob bzw. inwiefern explizite und implizite Einstellungen tatsächlich als Kontinuum konzipiert werden sollten, thematisierte die zweite Fragestellung von Studie III, wobei die Untersuchung der Dimensionalität expliziter und impliziter Einstellungen mit vorliegender Arbeit nicht hinreichend erfolgen kann, da allgemein die Validität impliziter Einstellungsmaße und die Funktionsweise der Geschichtenmessung in Bezug auf eine Erfassung expliziter und impliziter Einstellungsanteile noch ungeklärt ist (vgl. Abb. 6-39):

(2) *Inwieweit und durch welche der anderen Einstellungsmessungen können explizite Einstellungen vorhergesagt werden. Ist es zulässig, eine Explizit-Implizit-Dimension der Einstellungen anzunehmen?*

Den Überlegungen zur Untersuchung eines Explizit-Implizit-Kontinuums zufolge, wurde eine zusätzliche Einstellungsmessung entwickelt, die potentiell eine Mediatorfunktion für die Beziehung zwischen expliziten und impliziten Einstellungen hat (vgl. Studie II). Das als „offene Assoziativmessung“ bezeichnete Messverfahren bedient sich der Methodik des Thematischen Apperzeptionstests, der in seiner ursprünglichen Konzeption als projektives Testverfahren insbesondere zur impliziten Motivmessung eingesetzt wird (Campbell & Williams, 2001; McClelland, 1980; McClelland, Koesters & Weinberger, 1989; Schultheiss & Brunstein, 2001; Spangler, 1992; vgl. Schmalt & Sokolowski, 2000). Es wird angenommen, dass das die relativ undurchschaubare Messintention und das offene Antwortformat des Verfahrens eine Erfassung impliziter Einstellungsanteile zu ermöglichen, wogegen die Bildmotive, die mehr oder weniger zumindest Rückschlüsse auf die Thematik erlauben, und der verbale, frei kontrollierbare Charakter der erforderlichen Antworten eher explizite Einstellungsanteile ansprechen sollte. Dementsprechend sollte untersucht werden, ob für die offene Assoziativmessung eine Art „Zwischenstellung“ zwischen Selbstbericht und IAT (bzw. EAST) besteht. Zwar teilen Fragebogen und Geschichtenmessung das verbale und kontrollierbare Antwortformat, dennoch weiß eine Person, wenn sie eine Geschichte zu einem Bildmotiv schreibt, nicht, nach welchen Kriterien diese ausgewertet wird, während es relativ einfach ist, sich auf einer eindimensionalen Einstellungsskala in bestimmter Weise zu „verhalten“. Genauso ist anzunehmen, dass viele Versuchspersonen die Messintention des IAT durchschauen oder zumindest erahnen (vgl. De Houwer in press) und dennoch wenig Möglichkeiten haben, diese zu beeinflussen (z.B. Banse, Seise & Zerbes, 2001; Steffens, 2004).



Die Explizit-Implizit-Dimension der Einstellungen wird anhand der Beziehungsmuster der Einstellungsmaße untersucht. Sollten explizite und implizite Einstellungen die beiden Pole eines Kontinuums darstellen, müssten jeweils höheren Korrelationen zwischen expliziten Einstellungen und dem Vorurteilsgehalt der Geschichten sowie zwischen impliziten Einstellungen und dem Vorurteilsgehalt der Geschichten bestehen, während die Beziehungen zwischen expliziten und impliziten Einstellungen vergleichsweise schwächer sein sollten. Die Korrelationen zwischen Selbstbericht und Geschichtenmessung sollten mehr oder weniger vergleichbar zu den Korrelationen zwischen IAT (bzw. EAST) und Geschichtenmessung sein, zumindest aber sollten sie betragsmäßig höher sein, als die Korrelationen zwischen Selbstbericht und IAT (bzw. EAST). Demgemäß sind bei der Vorhersage expliziter Einstellungen durch offene Assoziativmessung und implizite Einstellungen hohe direkte Pfadladungen zwischen impliziter Messung und Geschichtenmessung sowie zwischen Geschichtenmessung und Selbstbericht und verhältnismäßig schwächere indirekte Pfade zwischen impliziter Messung und Selbstbericht zu erwarten (vgl. Abb. 6-45). Aufgrund der theoretischen Zusammenhangshypothesen werden keine Kreuzpfade zwischen beiden Einstellungsobjekten angenommen. Der EAST wurde zunächst nicht in die Analyse einbezogen, weil Reliabilität und Validität der EAST-Maße nicht hinreichend gegeben sind und da die Modellierung zweier zusätzlicher latenter Variablen für EAST-Übergewicht und -Homosexualität hohe Komplexitätskosten bei vergleichsweise geringem zu erwartenden Nutzen bedeutet. Aus Gründen der Vollständigkeit ist jedoch auch das komplette Modell zur Vorhersage expliziter Einstellungen mit Geschichtenmessung, IAT und EAST überprüft worden. Mit der Modellierung wird der erfahrungsbedingten Annahme entsprochen, dass der EAST im Vergleich zum IAT noch weniger durchschaubar und kontrollierbar ist. Aus diesem Grund beschreibt der EAST im vollständigen Vorhersagemodell den lokalen impliziten Pol des theoretischen Explizit-Implizit-Kontinuums (siehe Abb. 6-44).



**Abbildung 6-44**, Vorhersagemodell gemäß theoretischer Dimensionalität der Daten (siehe Abb. 6-39, 6-46)

Die Vorhersagemodelle legen keine lineare Explizit-Implizit-Dimension nahe. Substanzielle positive Beziehungen bestehen allein zwischen den expliziten Einstellungen und der Geschichtenmessung je Einstellungsobjekt. Für Homosexualität sind zusätzlich signifikante direkte Zusammenhänge zwischen den IAT und der Fragebogenmessung zu dokumentieren, jedoch besteht für keines der beiden Einstellungsobjekte ein sinnvoller, substanzieller Zusammenhang zwischen der Geschichtenmessung und den impliziten Maßen. Die Einbeziehung des EAST in die Vorhersage expliziter Einstellungen resultiert in einer Verdopplung der Varianzaufklärung expliziter Einstellungen zu Homosexualität, aber auch in einer Steigerung der Varianzaufklärung für die Homosexualitätsgeschichten. Dies ist jedoch auf einen negativ gerichteten Vorhersagepfad ausgehend vom EAST zurückzuführen. Prinzipiell werden für das Einstellungsobjekt Homosexualität entgegen der Erwartungen jeweils negative Pfadgewichte zwischen beiden impliziten Maßen und der Geschichtenmessung ermittelt, während die Zusammenhänge für Übergewicht zum Einen stärker und zum Anderen hypthesenkonform gerichtet sind. Dies deutet erneut auf die differentielle Validität der Geschichtenmessung in Abhängigkeit des Einstellungsobjektes hin (siehe Gesamtdiskussion, Abschnitt 2.3).

Für das Gesamtmodell sind zudem positive Pfade zwischen latenten EAST und IAT-Faktoren zu verzeichnen, die zwar nicht auf einen engen, jedoch hypthesenkonform gerichteten Zusammenhang zwischen beiden impliziten Einstellungsmaßen hinweisen. Die Residuen der Vorhersage innerhalb einer Messung korrelieren jeweils positiv miteinander, was wiederum die Bedeutung methodenspezifischer Varianz der verschiedenen Messungen unterstreicht. Dies gilt allerdings nicht für die latenten EAST-Faktoren, die in keiner statistisch bedeutsamen Beziehung zueinander stehen. Als ausschlaggebend dafür muss der reliabilitätsbedingt geringe Varianzanteil der latenten EAST-Faktoren angeführt werden. Insofern sind auch keine varianzstarken Vorhersagepfade seitens des EAST zu erwarten.

Die Tatsache, dass für beide Einstellungsobjekte lediglich Geschichtenmessung und Selbstbericht in Beziehung stehen und für Homosexualität der IAT keinen Zusammenhang zum Vorurteilsgehalt der Geschichten aufweist, jedoch bei der Aufklärung von expliziten Einstellungen zu Homosexualität inkrementelle Validität besitzt, deutet darauf hin, dass die Annahme eines zweidimensionalen Kontinuums expliziter und impliziter Einstellungen nicht zulässig ist. Es scheint vielmehr nahe liegend, dass die Anzahl der hier eingesetzten Messungen die Dimensionalität des Beziehungsraumes definiert. Hierin verbirgt sich jedoch die Gefahr, bei jeder beobachteten Dissoziation zwischen Messungen einer Disposition, einen neuen inkrementellen Einstellungsaspekt zu postulieren. Olson & Fazio (2004, 655) geben entsprechend zu bedenken:

*According to this view, the attitude becomes a product of the method of measurement – and the perceiver is forced to find a place for as many attitude components as dissociated measures of them. Clearly, some theoretical reason for supposing that different measures will show how different patterns of results is needed before all of the measures are thrown into the matrix. “*

Theoretische Überlegungen zur Bedeutung verschiedener Einstellungsmaße betreffen beispielsweise die unterschiedlichen Varianzquellen der Messungen. Der unsystematischen *Messfehlervarianz*, die insbesondere für die impliziten Maße eine entscheidende Rolle spielt, wird in vorliegender Untersuchung durch die latente Strukturgleichungsmodellierung Rechnung getragen. Der Konstruktvalidierung von Einstellungen zufolge (Fragestellung 1), überlagert die *methodenspezifische Varianz* der Messungen die *inhaltliche Varianz* der Einstellungsdispositionen, wobei Selbstbericht und Geschichtenmessung ein gewisses Maß an methodischer Varianz teilen. Dies äußert sich in relativ varianzstarken und korrelierenden Vorhersageresiduen (siehe Abb. 7-1). Zudem ist der Anteil spezifischer *situativer Varianz* am IAT Übergewicht und dem entsprechenden Selbstbericht zu berücksichtigen, da zwischen beiden Messungen eine Woche Abstand liegt. Verfahren, die sehr sensibel für situative Einflüsse ist und insofern vorrangig State-Varianz erfassen, sollten entsprechende Divergenzen zu Maßen aufweisen, die zu anderen Messzeitpunkten erhoben werden. Verschiedene Arbeiten konnten zeigen, dass der IAT anfällig für situationsspezifische Varianz ist (z.B. Cunningham, Preacher & Banaji, 2001; Steffens & Buchner, 2003). Da aus der Erhebungsprozedur in dieser Untersuchung der einzige offenkundige Unterschied zwischen beiden Einstellungsobjekten resultiert, ist anzunehmen, dass die fehlende Konvergenz zwischen impliziten und expliziten Einstellungen zu Übergewicht auf die unterschiedlichen Messzeitpunkte zurückzuführen ist. Zwar kann nicht ausgeschlossen werden, dass sich Einstellungen gegenüber Übergewicht und Homosexualität fundamental voneinander unterscheiden, was möglicherweise diese abweichenden Beziehungsmuster verursacht. Es wird jedoch zunächst eine Verallgemeinerbarkeit der Einstellungsinhalte angenommen, da eine Konstruktvalidierung von Einstellungen oder gar eine allgemeingültige Einstellungskonzeptionalisierung anders nicht statthaft wären (siehe Gesamtdiskussion, Abschnitt 7.4).

### 6.2.2.3 Fragestellung III: Die Rolle individueller & kultureller Einstellungen

Devine (1989) bezeichnet implizite Einstellungen als Resultate individueller Sozialisationserfahrungen, als frühzeitig gelernte kulturelle Stereotype, die eine relative Allgemeingültigkeit für die Mitglieder einer Gesellschaft haben (vgl. Abschnitt 2.1.1.2). Explizite Einstellungen resultieren dagegen aus den individuellen Erfahrungen von Personen und können bei Abweichungen von kulturellen Stereotypen zu einer Divergenz expliziter und impliziter Einstellungen führen. Bei der Entwicklung impliziter und expliziter Einstellungen spielen also Umwelteinfluss und spezifische Erfahrungen eine entscheidende Rolle. Nichtsdestotrotz sind implizite Einstellungen als interindividuelle variierende Dispositionen zu betrachten, da verschiedene Personen trotz einer relativ gemeinschaftlichen Sozialisation unterschiedliche implizite Assoziationen aufweisen.

Nun konnte jedoch gezeigt werden, dass der IAT nicht nur individuelle implizite Einstellungen erfasst, sondern auch extrapersonale Umweltassoziationen. Karpinski & Hilton (2001) finden beispielsweise keinen Zusammenhang zwischen der expliziten und impliziten Präferenz von Äpfeln versus Schokoriegel. Mit Obst und Süßigkeiten betrachten die Autoren Einstellungsobjekte, die verhältnismäßig unabhängig von sozialer Erwünschtheit sind und vermutlich auch relativ bewusst verarbeitet werden, zu denen jedoch stark divergierende Einstellungen in der Gesellschaft existieren. Äpfel gelten als gesund, wogegen Schokoriegel nicht als gesundes Lebensmittel bewertet werden. Sie zeigen, dass eine Beziehung zwischen expliziter Präferenz von Äpfeln oder Schokoriegeln und entsprechendem Verhalten besteht, nicht jedoch zwischen IAT und Verhalten. Sie postulieren in diesem Zusammenhang die Beeinflussung des IAT-Effekts durch extrapersonale Umweltassoziationen (Äpfel = gesund; Schokoriegel = ungesund), die kulturell geteilt werden, jedoch nicht unbedingt die individuelle Haltung von Personen widerspiegeln müssen. Dieses Ergebnis wird darauf zurückgeführt, dass im IAT vor allem Assoziationen auf dem globalen Niveau der Kategorienlabel gemessen werden, die in der Regel normativ besetzt sind z.B. „negativ“ vs. „positiv“, dass zumeist universell positive oder negative Stimuli verwendet werden und dass die standardmäßig erfolgenden Fehlermeldungen eine normativ korrekte Antwort suggerieren (Olson & Fazio, 2004). Diese Eigenschaften des Standard-IAT könnten Ursache für eine stärkere Aktivierung impliziter kultureller, statt persönlicher Einstellungen im IAT sein.

Weitere empirische Hinweise auf eine Beeinflussung des IAT-Effekts durch kulturelle Einstellungen liefert der Vergleich von Untersuchungen, denen zufolge 80% der weißen Probanden einen Rassenbias im IAT aufzeigen (Nosek, Banaji & Greenwald, 2002a) mit Untersuchungen, in denen nur 50% der Probanden Vorurteile gegenüber Farbigen laut Affektiven Priming haben (Fazio et al., 1995). Hierbei sind jedoch auch Veränderungen der Einstellungen in der Gesellschaft über die Zeit zu berücksichtigen (z.B. Yang, 1997). Dennoch weisen die IAT-spezifischen Befunde einer fehlenden Bevorzugung der Eigengruppe unter Farbigen darauf hin, dass allgemeingültige, sozial geteilte, kulturelle Einstellungen ihren Niederschlag im IAT finden (Nosek, Banaji & Greenwald, 2002a; vgl. Fazio et al., 1995).

Olson & Fazio (2004) nehmen verschiedene Modifikationen zur Personalisierung des IAT vor, indem sie die Label „I like“ vs. „I don't like“ (statt „pleasant“ vs. „unpleasant“) vorgeben, nicht normativ bewertete Stimuli für das evaluative Material verwenden und keine Fehlerrückmeldung geben. Die Autoren zeigen in vier Experimenten, dass ein dieser Weise personalisierter IAT im Gegensatz zu einem klassischen IAT deutlich weniger Negativität gegenüber Fremdgruppen reflektiert und höher zu expliziten Einstellungen, Verhaltensintentionen und selbstberichtetem vergangenem Verhalten korreliert. Hinsichtlich politischer Einstellungen stellen die Autoren Korrelationsbeträge von .42 und .77 für klassischen und modifizierten IAT und eine Gefühlsthermometermessung oder respektive .37 und .80 für ein semantisches Differential. Diese hohen (manifesten) Korrelationen des personalisierten IATs mit expliziten Einstellungsmaßen legen eine Konvergenz in einem Konstrukt nahe. Die außerdem be-

richtete positive Korrelation zwischen traditionellem und modifiziertem IAT (.37) zeigt einerseits, dass die Modifikationen zu zwei unterschiedlichen IAT-Messungen geführt haben, die andererseits substantielle Gemeinsamkeiten aufweisen. Da Olson & Fazio (2004) die drei Modifikationsmöglichkeiten für den IAT nicht durchgängig gemeinsam gegen den traditionellen IAT testen, ist eine allgemeingültige Interpretation der vier Experimente, aus der eine klare Entscheidung abzuleiten wäre, welcher IAT der validere ist, nicht zu treffen.

Anstelle von personalisierenden Modifikationen des IAT im Sinne von Olson & Fazio wird in dieser Untersuchung der klassische Ablauf beibehalten und neben dem expliziten Selbstbericht ein expliziter „Fremdbericht“ über Einstellungen und vorurteilsrelevante Motivationen in der Gesellschaft erhoben. Sollte der klassische IAT tatsächlich extrapersonale Assoziationen erfassen, wären engere Beziehungen zwischen IAT und den antizipierten Einstellungen zu erwarten. Zudem ist somit ein Vergleich der beiden impliziten Maße hinsichtlich ihrer Relevanz für persönliche versus kulturell reflektierte Einstellungen möglich. Dessen ungeachtet aktiviert die forcierte Trennung zwischen persönlichen und gesellschaftlichen Einstellungen eine stärkere Introspektion der eignen Kognitionen, wovon prinzipiell eine zusammenhangssteigernde Wirkung für expliziten und impliziten Einstellungen zu erwarten ist (vgl. Hofmann, Gschwendner & Schmitt, 2005). Als Indikator einer Selbst-Andere-Abweichung der Einstellungen kann die Differenz zwischen antizipierten und individuellen Einstellungen auf eine moderierende Wirkung der Explizit-Implizit-Beziehung untersucht werden (Nosek, 2004). Die dritte Fragestellung lautete:

- (3) *Lassen sich, explizit erfasst, individuelle von gesellschaftlich antizipierten Vorurteilen trennen und welche Bedeutung hat eine solche Differenzierung für implizite Maße bzw. den Zusammenhang zwischen expliziten und impliziten Einstellungen?*

Die Trennung zwischen persönlichen und kulturellen Einstellungen mag von begrenztem Sinn sein, da manche Autoren der Ansicht sind, jede sozialisierte Person sei mehr oder weniger Produkt ihrer Kultur (z.B. Banaji, 2001). Die Mittelwerts- und Korrelationsbefunde von Studie III zeigen jedoch deutlich, dass es sich bei individuellen und antizipierten Einstellungen um unterscheidbare Konstrukte handelt (vgl. Tab. 6-19). Ein Vergleich der Skalenmittelwerte für die individuellen und antizipierten Einstellungsmessungen weisen ausnahmslos und durchweg statistisch bedeutsam der Gesellschaft negativere Einstellungen zu Übergewicht und Homosexualität zu. Dagegen werden bis auf die externe Motivation zur Vorurteilsfreiheit, für alle Motivationsmaße signifikant höhere Mittelwerte auf den selbstberichteten im Vergleich zu den antizipierten Skalen verzeichnet. Hierin spiegeln sich deutlich die soziale Brisanz und ergo die Erwünschtheitsrelevanz für eine Messung von Einstellungen zu gesellschaftlich diskriminierten Minderheiten wider. Die Korrelationen zwischen persönlichen und kulturellen Einstellungs- und Motivationsmaßen sind zwar positiv und in sechs von neun Fällen statis-

tisch bedeutsam, jedoch sind die Zusammenhangsbeträge sehr gering – sie rangieren von  $-.03$  (Motivation zu Vorurteilsfreiem Verhalten) bis  $.28$  (Externale Motivation zur Vorurteilsfreiheit). Dies zeigt, dass es sich nicht allein um Mittelwertsdifferenzen für individuelle und antizipierte Maße handelt.

Der korrelativen Zusammenhangsanalyse für persönliche bzw. kulturelle Einstellungen und die impliziten Maße IAT und EAST zufolge, bestehen keine gravierenden Unterschiede in den Beziehungsmustern für die beiden explizite Einstellungsmaße. Für den IAT-Übergewicht sind weder Korrelationen zu individuellen, noch zu antizipierten expliziten Einstellungen zu verzeichnen, wobei hier – wie bereits angeführt – die situationsspezifische Varianz der Maße möglicherweise einen wahren Zusammenhang überlagert. Für den IAT-Homosexualität bestehen bedeutsame positive Zusammenhänge ausschließlich zu den *individuellen* kognitiven und affektiven Einstellungen zu Homosexualität. Der positive Zusammenhang zwischen dem IAT und den *antizipierten* affektiven Einstellungen ist tendenziell bedeutsam. Der Zusammenhang könnte jedoch auch auf der Überlappung zwischen individuellen und antizipierten affektiven Einstellungen zu Homosexualität ( $r=.22$ ) beruhen. Die Befunde deuten eher auf die hauptsächliche Erfassung selbstrelevanter impliziter Einstellungen seitens des IAT-Homosexualität hin, wobei wahrgenommene Gesellschaftseinstellungen eine eher untergeordnete Rolle spielen.

Für den EAST werden tendenziell eher Zusammenhänge zu antizipierten Einstellungen verzeichnet, wobei hier stets die Reliabilitätsproblematik der EAST-Maße und die wie für den IAT-Übergewicht geltenden situativen Variabilitäten berücksichtigt werden müssen. Aufgrund der mangelnden internen Konsistenz des Gesamteffekts unterhalten insbesondere die EAST-Effekte auf Ebene der Messblöcke statistisch bedeutsame Beziehungen zu expliziten Einstellungen, wobei sich aufgrund der Inkonsistenz der jeweils zwei Blockeffekten pro EAST kein einheitliches Zusammenhangsmuster ergibt. Eindeutige Aussagen über einen verhältnismäßig „valideren“ Teileffekt des EAST, in Bezug auf den Messblock (1. versus 2. Block) bzw. das Datenformat (Reaktionszeiten versus Fehlerraten) lassen sich entsprechend der Zusammenhangsanalyse nicht ableiten. Zudem sind unerwünschte signifikante Korrelationen *zwischen* den Einstellungsobjekten bei fehlenden Zusammenhängen *innerhalb* eines Einstellungsobjektes zu berücksichtigen. Der Fehlereffekt gemäß ersten Block des EAST-Homosexualität unterhält eine signifikant positive Korrelation zu antizipierten affektiven Einstellungen zu Homosexualität, nicht jedoch zu den individuellen Einstellungen, außerdem bestehen negative Beziehung zwischen dem EAST-Homosexualität und den individuellen kognitiven und affektiven Einstellungen zu Übergewicht, die teilweise auch für den Gesamteffekt signifikant sind. Der EAST-Übergewicht unterhält verhältnismäßig mehr und zudem engere Beziehungen zu expliziten Einstellungen, hier jedoch speziell für das Einstellungsobjekt Homosexualität. Eine leicht positive Korrelation ist zwischen dem Reaktionszeiteffekt des zweiten Blocks und wiederum den *antizipierten* Einstellungen zu Übergewicht zu vermerken. Im Gegensatz zum EAST-Homosexualität werden ausschließlich signifikant positive Zu-

sammenhänge zu den expliziten antizipierten Einstellungen für das andere Einstellungsobjekt verzeichnet.

Für den IAT werden dagegen keinerlei Kreuzbeziehungen zwischen den Einstellungsobjekten beobachtet, was für diskriminante Validität des IAT spricht (vgl. Gawronski, 2002). Obwohl der Anteil signifikanter Zusammenhänge für den EAST nicht allein auf die festgelegte Zufallswahrscheinlichkeit von 5% zurückgeführt werden kann, ist angesichts der psychometrischen Unzulänglichkeiten des EAST und der daraus resultierenden Unterschiedlichkeit der Blockeffekte fragwürdig, inwieweit eine inhaltliche Interpretation der Korrelationsbefunde für die EAST-Daten vorgenommen werden kann. Der Vergleich beider impliziter Maße hinsichtlich ihrer Bedeutung für individuelle und antizipierte Einstellungen liefert Hinweise auf eine vorrangige Relevanz des IAT für *individuelle* Einstellungen und – unter Berücksichtigung des inkonsistenten Beziehungsmusters – eine relative Bedeutsamkeit des EAST für *antizipierte* Einstellungen.

Die Differenzen zwischen persönlichen und gesellschaftlich antizipierten Einstellungen kann als Indikator einer interindividuell variierenden Selbst-Andere-Diskrepanz der Einstellungen betrachtet werden. Um die Moderatorwirkung der Selbst-Andere-Abweichung für den Zusammenhang zwischen expliziten und impliziten Einstellungen zu untersuchen, betrachtet Nosek (2004) ein doppeltes Differenzmaß. Der Autor berechnet zunächst die Differenzen zwischen den selbstberichteten Gefühlsthermometermaßen der Ziel- und Kontrastkategorie (z.B. heterosexuelle – homosexuelle Personen) und entsprechend den gesellschaftlich antizipierten Gefühlsthermometermaßen. Die Abweichung zwischen beiden Gefühlsthermometerdifferenzen bezeichnet der Autor als Diskrepanz zwischen den persönlichen und den wahrgenommenen Einstellungen in der Gesellschaft. In ähnlicher Weise werden in vorliegender Arbeit die Zusammenhänge zwischen impliziten und expliziten (individuellen) Einstellungen getrennt für Hoch- und Niedrigabweichler in individuellen und antizipierten Einstellungen betrachtet. Dabei handelt es sich um einfache Mittelwertsunterschiede zwischen den eigenen Einstellungen und den wahrgenommenen Einstellungen der Bezugsgruppe, die ebenso wenig wie bei Nosek einen *objektiven* Referenzpunkt für die Abweichung zwischen eigenen und antizipierten Einstellungen bietet, sondern lediglich interindividuelle Unterschiede in der wahrgenommenen Abweichung zwischen persönlichen und allgemeinen Einstellungen indiziert.

Um die Bedeutung des Übereinstimmungsgrades zwischen persönlichen und wahrgenommenen gesellschaftlichen Einstellungen auf den Zusammenhang zwischen expliziten und impliziten Einstellungen zu überprüfen, wird die Gesamtstichprobe anhand der Selbst-Andere-Diskrepanz in zwei Gruppen geteilt (Mediansplit) und ein Vergleich der Korrelationsmuster expliziter und impliziter Einstellungen vorgenommen. Innerhalb des Einstellungsobjektes Übergewicht bestehen keine bedeutsamen Unterschiede in den Beziehungsmustern für implizite und explizite Einstellungen in Abhängigkeit vom Abweichungsgrad zwischen individuellen und antizipierten expliziten Einstellungen. Dagegen

sind innerhalb des Einstellungsobjekts Homosexualität deutliche Unterschiede zwischen Probanden mit hohen und geringen Selbst-Andere-Differenzen insbesondere für die Korrelationen zwischen IAT und expliziten Einstellungen festzustellen. Hierbei zeichnen sich Probanden mit *geringen* Einstellungsdivergenzen durch höhere Korrelationen zwischen IAT und expliziten Einstellungen zu Homosexualität aus, als Personen, die hohe Abweichungen zwischen den individuellen und antizipierten Einstellungen aufweisen. Dies steht den Überlegungen und Befunden von Nosek (2004) entgegen, der eine engere Implizit-Explizit-Beziehung bei hoher Selbst-Andere-Diskrepanz aufgrund der verbesserten Elaboriertheit stark individualisierter Einstellungsinhalte erwartet und aufzeigt. Der Autor findet dementsprechend ein positives Regressionsgewicht für einen Interaktionsterm der Selbst-Andere-Diskrepanz, weist jedoch auch nach, dass der Moderatorterm ebenso wie der Moderator *Einstellungsstärke* keine inkrementelle Validität neben der *Selbstpräsentation* und der *Einstellungsdimensionalität* besitzen.

Der in Studie III aufgezeigte zusammenhangsmindernde Einfluss einer hohen Selbst-Andere-Diskrepanz deutet darauf hin, dass Probanden, die sich in ihren selbstberichteten, individuellen Einstellungen deutlich von ihren wahrgenommenen Bezugsgruppeneinstellungen unterscheiden, ihre expliziten Einstellungen zum Zwecke einer positiveren Selbstdarstellung systematisch verzerren. Insofern spiegelt eine hohe Selbst-Andere-Diskrepanz wahrscheinlich nichts anderes als soziale Erwünschtheit oder eine Motivation zur Vorurteilsvermeidung wider. Beide Dispositionen können eine Moderatorwirkung hinsichtlich des Zusammenhanges zwischen expliziten und impliziten Einstellungen ausüben. Für die Soziale Erwünschtheit als Moderator der Explizit-Implizit-Beziehung ist eine insgesamt inkonsistente Befundlage mit sowohl positiver, als auch negativer, als auch fehlender Moderatorwirkung dokumentiert (Egloff & Schmukle, 2003; Hofmann, Gschwendner & Schmitt, 2005; Nosek & Banaji, 2002; Nosek, 2004). Die Motivation zur Vorurteilsfreiheit stellt dagegen einen relativ stabilen und konsistent gerichteten Moderator der Explizit-Implizit-Korrespondenz dar, wobei zwischen extrinsisch und intrinsisch orientierter Motivation zu unterscheiden ist (vgl. Banse & Gawronski, 2003; Banse, Seise & Zerbes, 2001; Devine et al., 2002; Dunton & Fazio, 1997). Eine Diskussion von Moderatortreffekten der Motivation zur Vorurteilsfreiheit erfolgt im Rahmen der vierten Fragestellung (siehe folgender Abschnitt).

Kritisch zu bedenken ist die für die dritte Fragestellung verwandte Methode zur Gruppenbildung, um die Bedeutung der Selbst-Andere-Diskrepanz für die Zusammenhänge zwischen expliziten und impliziten Einstellungen untersuchen zu können. Mittels Mediansplit erfolgt eine relativ grobe Teilung der Stichprobe in Hoch- und Niedrigabweichler, wobei graduelle Abstufungen in den Differenzmaßen vernachlässigt werden. Zudem ist prinzipiell fraglich, inwiefern der rein statistisch bedingte Mediansplit das beste Kriterium zu Gruppenbildung darstellt, da es ebenso angebrachter erscheint, Personen mit positiv gerichteter- von Personen mit negativ gerichteter Selbst-Andere-Diskrepanz vonein-



ander zu unterscheiden. Personen, die eine negative Differenz zwischen antizipierten und individuellen Einstellungen aufweisen, also negativere Einstellungen für die eigene Person berichten, als für die Gesellschaft, sind jedoch nur vereinzelt vertreten (im Mittel 2.9%).

Abschließend zur dritten Fragestellung kann festgehalten werden, dass der IAT in der aktuellen Untersuchung eine vorrangige Bedeutung für individuelle explizite Einstellungen aufweist, während der EAST eine – jedoch mit aller Vorsicht zu interpretierende – tendenzielle Bedeutung für antizipierte Einstellungen hat und dass eine höhere Beziehung zwischen IAT-Homosexualität und expliziten Einstellungen zu Homosexualität für Probanden besteht, die geringe Abweichungen zwischen individuellen und wahrgenommenen Gesellschaftseinstellungen aufweisen.

#### 6.2.2.4 *Fragestellung IV: Die Moderatorwirkung von Motivation zur Vorurteilsfreiheit*

Allgemein betrachtet ist die Motivation zur Vorurteilsfreiheit ein *spezifischer* Moderator der Explizit-Implizit-Beziehung, der ebenso wie soziale Erwünschtheit als Adjustierungsmoderator bezeichnet werden kann (Hofmann, Gschwendner & Schmitt, 2005). Für eine Untersuchung der Bedeutung vorurteilsrelevanter Motivationen für die Zusammenhänge zwischen expliziten und impliziten Einstellungen können verschiedene Konstrukte berücksichtigt werden, die mehr oder weniger zwei Kategorien zuzuordnen sind: extrinsisch versus intrinsisch orientierter Motivationen zur Vorurteilsfreiheit. Die meisten Motivationsmessungen, wie z.B. die *Motivation zu vorurteilsfreiem Verhalten* (Banse & Gawronski, 2003), zielen auf extrinsische Faktoren wie Political-Correctness-Standards oder soziale Erwünschtheit ab und haben insbesondere einen Einfluss auf bewusste, steuerbare Einstellungsmaßen – also explizite Einstellungen. Hierbei übt die Motivation zu vorurteilsfreiem Verhalten eine negative Moderatorwirkung auf die Korrespondenz zwischen expliziten und impliziten Einstellungen aus, da hoch motivierte Personen eine vorurteilsfreie Selbstpräsentation in expliziten, nicht aber in impliziten Maßen erzielen. Eine *internale Motivation zur Vorurteilsfreiheit* dagegen, ist am persönlichen Wertsystem orientiert und vermag auch die unbewussten, weniger kontrollierbaren impliziten Einstellungsmessungen zu beeinflussen (vgl. Blair, 2001, 2002; Blair, Ma & Lenton, 2001; Devine 1998; Devine et al., 2002). Von intrinsisch orientierter Motivation können somit auch positive Moderatorwirkungen erwartet werden. In Studie III wurden verschiedene Motivationsmaße einbezogen, um ihre Bedeutung für die Beziehung zwischen expliziten und impliziten Einstellungen zu Übergewicht und Homosexualität gemäß untersuchen zu können:

- (4) *Können die in der Literatur berichteten Moderatorwirkungen der Motivation zur Vorurteilsfreiheit in der Beziehung impliziter und expliziter Einstellungsmaßen auf vorliegende Fragestellung appliziert werden*

Zur Messung der Motivation zur Vorurteilsfreiheit wurden die Skalen *Internal* und *External Motivation to Respond without Prejudice* (IMS, EMS, Plant & Devine, 1998), *Motivation zu vorurteilsfreiem Verhalten* (MVV, Banse & Gawronski, 2003) und die *Vorurteilstrollskala* (VKS, Gschwendner-Lukas, Hofmann & Schmitt, 2004) eingesetzt. Für diese verschiedenen Motivationsmaße, die analog zu den expliziten Einstellungskriterien als individuelle und antizipierte Dispositionen erfasst wurden, kann zunächst eine dreidimensionale Faktorenstruktur angenommen werden, wobei der varianzstärkste Faktor die *antizipierte* vorurteilsrelevante Motivation repräsentiert und der zweite Faktor die entsprechenden *individuellen* Motivationen. Den dritten Motivationsfaktor bildet maßgeblich die individuelle und antizipierte Externale Motivation to Respond without Prejudice, was zugleich eine Validierung der Skala darstellt, da sich Selbst- und Fremdberticht gleichsam an den „Anderen“ orientieren. Entsprechend besteht für dieses Konstrukt die vergleichsweise größte Korrelation zwischen individuellen und antizipierten Dispositionen.

Für die Untersuchung der Moderatorwirkung von vorurteilsrelevanter Motivation werden die *individuellen* expliziten Einstellungen zu Übergewicht bzw. Homosexualität und dementsprechend der Faktor der *individuellen Motivation zur Vorurteilsfreiheit* (IVMF) betrachtet. Die hierarchische Regressionsanalyse liefert lediglich für die Beziehung zwischen expliziten affektiven Einstellungen zu Homosexualität und impliziten Einstellungen gemäß IAT-Homosexualität eine statistisch bedeutsame Moderation durch den Motivationsfaktor IVMF. Die inkrementelle Varianzaufklärung des Moderationsterms beträgt 3%. Im Gesamtmodell wird für den Moderator ein negatives und für den Moderationsterm ein positives Regressionsgewicht geschätzt. Das bedeutet, je höher die individuelle Motivation zur Vorurteilsfreiheit bei einem Probanden ausfällt, desto stärker ist der Zusammenhang zwischen IAT und affektiven Einstellungen zu Homosexualität, obwohl der Motivationsfaktor IVMF für sich genommen in negativer Beziehung zu expliziten affektiven Einstellungen steht. Insofern liegen nicht nur Unterschiede in den expliziten, sondern auch in den *impliziten* Einstellungen in Abhängigkeit von der individuellen Motivation zur Vorurteilsfreiheit vor. In vorliegender Untersuchung zeichnen sich hoch motivierte Personen (hohe Ausprägung in IVMF) durch positivere explizite Einstellungen und einen signifikant geringeren IAT-Effekt aus, wie es auch Devine (1998) und Devine et al. (2002) für Personen zeigen, die eine hohe interne Motivation zur Vorurteilsfreiheit aufweisen. Individuell zur Vorurteilsfreiheit motivierte Probanden unterscheiden sich laut Blair (2001) oder Blair, Ma & Lenton (2001) hinsichtlich verschiedener Kontrollstrategien wie Unterdrückung, Vermeidung stereotypischer Signale oder die aktive Suche von Informationen, die im Gegensatz zum Stereotypen stehen, von niedrig motivierten Probanden. Diese intrinsisch wirksamen Kontrollstrategien haben ein nachweisliches Potential zur tief greifenden Einstellungsveränderung bzw. –kontrolle, die auch implizite Einstellungsmessung beeinflussen können. Die im Selbstbericht in Studie III erforderte Trennung zwischen individuellen und antizipierten vorurteilsrelevanten Motivationen, hat möglicherweise eine stärkere interne Ausrichtung der

selbstberichteten individuellen Motivationen zur Folge. Dies erklärt die positive Wirkrichtung des Motivationsfaktors IVMF, der auch Motivationsmaße repräsentiert, für die einzeln eine *negative* Moderatorwirkung berichtet wird (z.B. MVV, Banse & Gawronski, 2002). Für den EAST sind im Gegensatz zum IAT *negative*, jedoch durchweg nicht signifikante Moderatorwirkungen der individuellen Motivation zur Vorurteilsfreiheit festzuhalten.

Moderatoreffekte werden an dieser Stelle also nur für den bereits signifikanten Zusammenhang zwischen expliziten (affektiven) Einstellungen zu Homosexualität und dem IAT Homosexualität registriert. Hierbei zeichnet sich nicht allein eine systematische Verzerrung der *expliziten* Einstellungsmessungen in Richtung Vorurteilsfreiheit ab, sondern eine tatsächliche Korrespondenz zwischen *geringen* IAT-Effekten und *geringen* expliziten Vorurteilen bei hoch motivierten Personen.

Gemäß der Unterscheidung zwischen internaler und externaler Motivation zur Vorurteilsfreiheit (IMS und EMS, Plant & Devine, 1998; Devine et al., 2002) werden die Auswirkungen hoher vs. niedriger internaler bzw. externaler Motivation mittels Varianzanalysen geprüft, wobei sich für die expliziten Einstellungen lediglich eine negative Wirkung der internalen Motivation abzeichnet. Die offene Assoziativmessung spricht dagegen auf Unterschiede in der externalen Motivation an, wobei hier eine hohe EMS-Ausprägung mit geringerem Vorurteilsgehalt der Geschichten, analog zum Selbstbericht, einhergeht. Für die impliziten Maße wird dagegen keine inhaltlich interpretierbare Bedeutung beider Motivationen zur Vorurteilsfreiheit festgestellt. Die im Gegensatz zu Devine et al. (2002) durchweg unbedeutenden Interaktionen zwischen externaler und internaler Motivation zur Vorurteilsfreiheit und die fehlende Bedeutung der internalen Motivation für implizite Einstellungen könnte eine Folge der starken Schiefverteilung beider Variablen und des Informationsverlustes bei der Bildung von jeweils zwei Motivationsgruppen anhand von Mediansplits sein.

An dieser Stelle ist ein allgemeiner statistischer Aspekt zur Untersuchung von Moderatoreffekten mit hierarchischer Regression, die nach wie vor standardmäßig für Moderationsanalysen eingesetzt werden, anzuführen. Klein (2000) führt gravierende Unzulänglichkeiten des Allgemeinen linearen Modells zur Moderatoranalyse auf. Da eine regressionsanalytische Moderatoranalyse weder Messfehler noch die Verteilungsvoraussetzungen, die für multiplikative Moderatorterme anzunehmen sind, berücksichtigt, bezeichnet Klein (2000) die Regressionsanalyse zur Untersuchung von Moderatoreffekten als nicht adäquat und schlägt eine auf latenter Strukturgleichungsmodellierung und alternativen Schätzalgorithmen basierende Auswertungsstrategie vor. Da die Reliabilität des Produktterms, also des Moderatorterms, immer *unter* der Reliabilität der beiden Ausgangsvariablen liegt, ist vor allem im vorliegenden Fall mit einer Verminderung der inkrementellen Validität des Moderatorterms, also einer erhöhten Betafehler-Wahrscheinlichkeit, zu rechnen (vgl. Busemeyer & Jones, 1983).

Neben den allgemeinen statistischen Problemen einer Moderatoranalyse mit hierarchischen Regressionen ist der spezifisch-methodische Aspekt der *Motivmessung* zu bedenken. Die empirische Untersu-

chung der Moderatorwirkung von Motivation zur Vorurteilsfreiheit hinsichtlich der Explizit-Implizit-Korrespondenz basiert bislang noch hauptsächlich auf *explizit* im Selbstbericht erfassten Motivationsmaßen. Dabei wird jedoch die, mit der Unterscheidung expliziter und impliziter Einstellungen vergleichbare Differenzierung zwischen expliziten und impliziten Motiven außer Acht gelassen. Im Hinblick auf die geteilte methodische Varianz des Selbstberichts und mögliche inhaltliche Überlappung von expliziter Einstellungs- und Motivationsmessung sollte die vorrangige Bedeutung der meisten vorurteilsrelevanten Motivationsmaße für die *explizite* Einstellungsmessung als messmethodisches Artefakt hinterfragt werden. Darüber hinaus sollten implizite einstellungsrelevante Motiven aus Validitätsgründen einbezogen werden, da für implizite Motive diskriminante und inkrementelle Validität nachgewiesen werden konnte (vgl. Schmalt & Sokolowski, 2000).

Für die Messung impliziter Motive, insbesondere der von Macht-, Leistungs- und Zugehörigkeitsmotiv, setzen verschiedene Autoren erfolgreich den Thematischen Apperzeptionstest ein (z.B. Kwon, Campbell & Williams, 2001; McClelland, 1980; McClelland, Koesters & Weinberger, 1989; Schultheiss & Brunstein, 2001; Spangler, 1992). In der aktuellen Arbeit wurde die Methodik des Verfahrens zwar vorrangig für eine Messung spontaner, unkontrollierter *und* reflektierter, kontrollierter Einstellungsinhalte adaptiert und eingesetzt, dennoch können den systematischen Zusammenhänge zwischen motivationsrelevanten Variablen der offenen Assoziativmessung und der expliziten Motivationsmessung wertvolle Hinweise in Bezug auf die Bedeutung impliziter und expliziter Motive entnommen werden. Die Korrelationen zwischen motivationalen Variablen der offenen Assoziativmessung (z.B. Geschichtenlänge) und den expliziten Motivationsmaßen, zeigen leichte, jedoch konsistente Überlappungen an. Die selbstberichtete individuelle Motivation zur Vorurteilsfreiheit (IVMF) unterhält positive Beziehungen zur Wortanzahl und negative Beziehungen zum individuellen Vorurteilsgehalt der Geschichten. Es sind zudem die expliziten Motivationsmaße, die sich für beide Einstellungsobjekte als varianzstärkste Prädiktoren bei der Vorhersage des Vorurteilsgehaltes der Geschichten auszeichnen. Das heißt, obwohl die Geschichtenlänge zunächst eine reine Kontrollvariable darstellte (vgl. Studie II), von der keine Beziehungen zu expliziten Einstellungen oder einstellungsrelevanten Motivationen erwartet wurden, erweist sich die Wortanzahl der Geschichten und somit die Geschichtenmessung an sich als hochgradig relevant für einstellungs- bzw. vorurteilsbezogene Motivationen.

Die nachweisliche Bedeutung der expliziten Motivation zur Vorurteilsfreiheit auch für die Geschichtenmessung könnte Ursache für die konsistenten positiven Zusammenhänge zwischen dem Vorurteilsgehalt der Geschichten und den expliziten Einstellungen sein. Wenn die Geschichten ebenso wie der Selbstbericht lediglich eine Art soziale Erwünschtheit reflektieren, die beispielsweise für die impliziten Messungen aufgrund der reduzierten Beeinflussbarkeit des Testergebnisses erwiesenermaßen weniger Bedeutung besitzt, ist ein vergleichsweise engerer Zusammenhang zwischen Geschichtenmessung und expliziten Einstellungen zu erwarten. Ein derart vermittelter Zusammenhang zwischen Ge-

schichtenmessung und expliziten Einstellungen konnte jedoch größtenteils ausgeschlossen werden, da die signifikanten Korrelationen durch die statistische Kontrolle von Motivation zur Vorurteilsfreiheit und sozialer Erwünschtheit keine dramatischen Einbußen erfahren. Lediglich für den Zusammenhang zwischen dem Vorurteilsgehalt der Homosexualitätsgeschichten und den expliziten Einstellungen zu Homosexualität ist ein Absinken der Korrelationen unter das statistisch bedeutsame Signifikanzniveau zu verzeichnen. Hier besteht jedoch auch ausgangs ein geringerer Zusammenhang, als für das Einstellungsobjekt Übergewicht. Es kann also davon ausgegangen werden, dass die konsistenten Beziehungen zwischen der offenen Assoziativmessung und dem expliziten Selbstbericht nicht auf der Vermittlung motivationaler Drittvariablen beruhen, sondern tatsächlich eine gemeinsame Einstellungsdisposition reflektieren.

#### 6.2.2.5 Fragestellung V: Nomologisches Netzwerk expliziter & impliziter Einstellungen

Zur Bedeutung expliziter sozialer Einstellungen und Vorurteile sind eine Vielzahl an Untersuchungen durchgeführt und demgemäß eine Reihe an stabilen Korrelaten etabliert worden. An dieser Stelle sind insbesondere interindividuell differierende Einstellungskorrelate von Relevanz. Als womöglich prominenteste Vorurteilsprädiktoren konnten globalere Konstrukte sozialer Einstellungen wie Autoritarismus, Soziale Dominanzorientierung oder Konservatismus durch eine positive Beziehung zu selbstberichteten sozialen Einstellungen ausgezeichnet werden (z.B. Adorno et al., 1950; Altemeyer, 1998; Duckitt et al., 2002; Heaven & St. Quintin, 2003; Lippa & Arad, 1999; Sidanius & Pratto, 1999; Zick, 1997). Darüber hinaus können kognitive Konstrukte wie Ambiguitätsintoleranz, Need for Closure oder Personal Need for Structure zur Nutzung von Stereotypen und zu individuellen Vorurteilen gegenüber sozialen Minderheiten in Verbindung gebracht werden (z.B. Hassan & Khaliq; 1987; Leone, 1999; Sarma, 1973; Sinha & Hassan, 1975). Auch demografische Variablen wie Geschlecht, Alter, sozioökonomischer Status und Bildungsniveau weisen enge, zum Teil jedoch vom spezifischen Einstellungsinhalt abhängige, Beziehungen zu sozialen Einstellungen auf (vgl. Kapitel 3.2.2).

Dagegen ist die Befundlage zum nomologischen Netzwerk impliziter Einstellungen noch relativ dürftig. Den bislang einzigen relevanten Studien von Rowatt & Franklin (2004) und Rowatt, Franklin & Cotton (2005) zufolge, ist das Beziehungsmuster für etablierte Prädiktoren expliziter Einstellungen und implizite Einstellungsmessung inkonsistent. In einer Untersuchung wurden keine signifikanten Zusammenhänge zwischen den bekanntesten Konstrukten im nomologischen Netzwerk expliziter Einstellungen – Autoritarismus und Soziale Dominanzorientierung – für implizite Einstellungen via IAT festgestellt (Rowatt, Franklin & Cotton, 2005). In einer anderen Untersuchung erwies sich RWA dagegen als bester Prädiktor für einen Rassen-IAT (Rowatt & Franklin, 2004).

Zur detaillierten Untersuchung der Bedeutung expliziter und impliziter Einstellungen wurden in Studie III eine Reihe an Konstrukten, die aus dem nomologischen Netzwerk expliziter Einstellungen bekannt sind, im Selbstbericht erfasst und hinsichtlich ihrer Zusammenhänge und des Vorhersagepotentials für die verschiedenen Einstellungsmaße analysiert. Die einbezogenen Konstrukte sind Right-Wing-Authoritarianism, Soziale Dominanzorientierung, Konservatismus, Ambiguitätsintoleranz, Need for Cognition, Soziale Erwünschtheit, Selbst- und Fremdtäuschung (sozial erwünschtes Verhalten), Modern Racism, Ekelempfindlichkeit, Intelligenz gemäß Mehrfachwortwahltest, einstellungsrelevante Selbstidentifikationsmaße wie der Body-Mass-Index, das Körperschema, die Geschlechtsidentität, die sexuelle Orientierung sowie einstellungsrelevante Kontaktmaße, z.B. die Anzahl übergewichtiger Freunde und Bekannte. Die entsprechende Fragestellung lautete:

(5) *Welche Bedeutung haben explizite und implizite soziale Einstellungen (Nomologische Netzwerk)*

Gemäß der drei methodischen Gruppen an Einstellungsmaßen (explizit, implizit, offen assoziativ) sind drei Korrelationsmatrizen zu den selbstberichteten Einstellungsprädiktoren zunächst gegenüberzustellen. Ein erster Augenscheinvergleich der resultierenden Beziehungsmuster zeigt die überragende Bedeutung der einbezogenen Prädiktorkonstrukte für *explizite* Einstellungen. Die Zusammenhänge zwischen den Geschichtenvariablen und den Einstellungsprädiktoren sind ebenfalls substantiell, jedoch bestehen zwischen den impliziten Einstellungsmaßen und den selbstberichteten Prädiktorfaktoren nicht deutlich überzufällige Beziehungen. Erst bei Betrachtung der unfaktorierten Primärkonstrukte ist ein statistisch bedeutsamer Anteil signifikanter Ergebnisse zu verzeichnen.

Hauptursächlich für die herausragende Bedeutung der Prädiktoren für explizite Einstellungen ist sicherlich die von expliziten Einstellungsmaßen und den Einstellungsprädiktoren geteilte methodenspezifische Varianz des Selbstberichts. Wie die latente Konstruktvalidierung offen legt, überlagern zum Teil methodische Varianzeffekte die inhaltliche Varianz der Maße deutlich. Auf der Ebene der manifesten Produkt-Moment-Korrelationen sind diese methodischen Varianzeffekte dagegen nicht kontrollierbar und führen bei entsprechender Bedeutung zu systematischen Zusammenhangssteigerungen für messmethodische ähnliche Maße.

Die trotz gravierender methodischer Differenzen substantiellen Beziehungen zwischen den Einstellungsmaßen der Geschichtenmessung und des Selbstberichts spiegeln sich in Überlappungen der nomologischen Netzwerke beider Messungen wider. Sowohl für explizite Einstellungen, als auch den Vorurteilsgehalt der Geschichten besteht eine konsistente positive Beziehung zu dem Einstellungsfaktor RWA und SDO für beide Einstellungsobjekte. Dieser Befund bestätigt erneut die bedeutsame Rolle von Autoritarismus und Soziale Dominanzorientierung für selbstberichtete Vorurteile (vgl. Altemeyer, 1996, 1998; Sidanius & Pratto, 1999) und zeigt darüber hinaus die Verallgemeinerbarkeit der Bedeutung beider Prädiktoren für eine methodisch abweichende Einstellungsmessung wie die ent-

sprechend ausgewerteten Geschichten. Ähnliches gilt für die individuelle vorurteilsbezogene Motivation, die ebenfalls konsistent für Selbstbericht und Geschichtenmessung in negativem Zusammenhang mit Einstellungen zu Übergewicht und Homosexualität steht. Dass diese gleichermaßen für explizite Einstellungen und den Vorurteilsgehalt der Geschichten bestehende Beziehung zu Motivation zur Vorurteilsfreiheit jedoch *nicht* den engen Zusammenhang zwischen beiden Einstellungsmessungen vermittelt, wurde im Zuge der Moderatoranalysen nachgewiesen und im vorangehenden Abschnitt diskutiert.

Da in den Korrelationsanalysen einerseits faktorenanalytisch gewonnene Einstellungskonstrukte betrachtet werden, was neben der erwünschten Komplexitätsreduktion einen unerwünschten Informationsverlust bedeutet, und andererseits die Multikolarität der Prädiktoren unberücksichtigt bleibt, wurden schrittweise multiple Regressionen berechnet. Diese Untersuchungen zur relativen Bedeutung der Primärkonstrukte für die expliziten Einstellungsmaße und den Vorurteilsgehalt der Geschichten stellen Autoritarismus als stabilen und positiv gerichtete Prädiktor kognitiver Einstellungen zu Übergewicht und Homosexualität sowie den Vorurteilsgehalt der Übergewichtsgeschichten heraus. Demzufolge nach repräsentiert Autoritarismus im Gegensatz zu Sozialer Dominanzorientierung inkrementelle Validität an expliziten und offen assoziierten Einstellungen, wobei sich die eindrucksvollste Bedeutung von Autoritarismus für die expliziten Einstellungen zu Homosexualität herauskristallisiert. Da autoritäre Personen häufig auch christlich konservative Werthaltungen verinnerlicht haben, zeichnen sie sich insbesondere auch durch eine Ablehnung von gleichgeschlechtlicher Liebe aus (z.B. Altemeyer & Hunsberger, 1992; Laythe, Finkel, Bringle & Kirkpatrick, 2002; Leak & Randal, 1995; Rowatt & Franklin, 2004).

Die Maße der Motivation zur Vorurteilsfreiheit erweisen sich als durchweg relevant für die Selbstbericht- und Geschichtenmaße beider Einstellungsobjekte. In aller Regel sind hierfür negative Zusammenhänge zu verzeichnen, die auf geringere Vorurteilsausprägungen bei Personen mit höherer Motivation zur Vorurteilsfreiheit deuten. Für die Vorhersage der Homosexualitätsgeschichten, die ausschließlich über motivationale Prädiktoren erfolgt, wird jedoch auch ein positives Regressionsgewicht für die internale Motivation zur Vorurteilsfreiheit ermittelt. Das heißt, je geringer die Motivation zu vorurteilsfreiem Verhalten und die Fremdtäuschung ausfallen und je höher die internale Motivation zur Vorurteilsfreiheit, desto stärker ist der Vorurteilgehalt der Homosexualitätsgeschichten. Dies ist mit hoher Wahrscheinlichkeit auf die positiven Zusammenhänge zwischen den zur Vorhersage aufgenommenen Prädiktoren zurückzuführen, denn auf der Ebene bivariater Korrelationen besteht ein negativer Zusammenhang zwischen dem Kriterium und der internalen Motivation zur Vorurteilsfreiheit. Neben den Ähnlichkeiten der Beziehungsmuster sind auch Unterschiede zwischen Selbstbericht und Geschichtenmessung festzustellen. So spielt für die expliziten kognitiven und affektiven Einstellungen zu Übergewicht die *Selbsttäuschung* eine zentrale Rolle, während sich für beide Geschichtenmessun-

gen ein Einfluss der *Fremdtäuschung* abzeichnet. Dies mag vorrangig an der Trennung zwischen individuellen und antizipierten Einstellungen im Selbstbericht liegen, mit der unter anderem auf eine stärkere Fokussierung der persönlichen Einstellungsdispositionen abgezielt wurde, weshalb die Selbsttäuschung hierfür von größerer Relevanz ist. In der Instruktion der Geschichtenmessung wurde zwar, wie im traditionellen Selbstbericht standardmäßig auch, darauf hingewiesen, dass die Geschichten dem entsprechen sollten, was die Probanden persönlich mit dem betreffenden Bildmotiv in Verbindung bringen. Nichtsdestotrotz scheint dieser Hinweis das sozial orientierte Motiv der optimalen Selbstpräsentation kaum abzuschwächen.

Außerdem ist Modern Racism ein positiver Prädiktor für explizite Einstellungen, insbesondere zu Homosexualität, während diese Variable keine Bedeutung für die Geschichtenmessung hat. Modern Racism misst allgemeine Ablehnung bzw. Diskriminierungstendenzen gegenüber sozialen Minderheiten und ist insofern als quasi inhaltsunspezifisches Vorurteilsmaß zu bezeichnen. Möglicherweise bedingt weniger der inhaltliche Aspekt, als vielmehr die gemeinsame Erfassung von expliziten Einstellungen und Modern Racism in einem Messinstrument (Laborset) diese stärkeren Zusammenhänge.

Für implizite Einstellungen sind zum Einen deutlich weniger statistisch bedeutsame Zusammenhänge zu den Prädiktorkonstrukten zu verzeichnen. Zum anderen weisen die Richtungen der Korrelationen nahezu durchweg in die entgegen gesetzte Richtung. Während Soziale Dominanzorientierung ein durchgehend positives Korrelat expliziter und offen assoziierter Einstellungen ist, bestehen signifikant negative Beziehungen zwischen SDO und beiden Fehlereffekten des IAT. Diese Zusammenhänge erscheinen schwer interpretierbar, da entgegen der Erwartung geringere implizite Assoziationsunterschiede für Zielkategorie und Kontrastkategorie bei Personen festgestellt werden, die aufgrund höherer sozialer Dominanzorientierung dazu tendieren, die Gesellschaft auf der Basis von Gruppenunterschieden als hierarchisch strukturiert wahrzunehmen. In der Untersuchung von Rowatt, Franklin & Cotton (2005) besteht zwar ein Zusammenhang nahe Null zwischen einem Christen-Muslime-IAT und SDO, dieser weist jedoch ein negatives Vorzeichen auf ( $r = -.03$ ), während RWA wie auch in vorliegender Untersuchung positiv mit dem IAT korreliert. Da für beide IATs negative Korrelationen zu SDO bestehen und dies trotz der unterschiedlichen Messmethoden und der geringen Reliabilitäten der Fehlereffekte, sind diese Zusammenhänge nicht als statistische Artefakte zu betrachten. Es zeigt sich, dass die geringeren IAT-Fehlereffekte hoch sozial dominanter Probanden auf signifikant geringere Fehlerraten in den jeweiligen inkompatiblen Messblöcken, und zwar speziell *direkt* nach dem Aufgabenwechsel, also in den inkompatiblen Übungsblöcken, zurückzuführen sind. Zwischen den Fehlerraten der kompatiblen Blöcke und SDO bestehen jeweils nicht signifikante, positive Zusammenhänge. Jedoch kehrt sich das Vorzeichen der Korrelationen zu SDO für die inkompatiblen Blöcke um, hier sind signifikant geringere Fehlerraten zu verzeichnen, je sozial dominanter ein Proband ist. Infolgedessen fällt die Fehlerratendifferenz aus inkompatiblen und kompatiblen Block kleiner aus, als bei



geringerer SDO-Ausprägung, was schließlich den negativen Zusammenhang zwischen SDO und IAT erklärt. Es bliebe zu untersuchen, wodurch der Befund geringerer Fehlerraten in den inkompatiblen Blöcken bei sozial dominanteren Personen erklärt werden kann. Ein möglicher Interpretationsansatz könnte aus dem Befund hoher Leistungsorientierung sozial dominanter Personen abgeleitet werden (Duriez & van Hiel, 2002; Heaven & Connors, 2001). Demzufolge nach sollten sozial dominante Probanden den Aufgabenwechsel im IAT als Herausforderung wahrnehmen, trotz der erschwerenden Umstände möglichst fehlerfrei zu reagieren. Da im jeweiligen inkompatiblen Übungsblock noch Fehlermeldungen gegeben werden, wird eine Leistungsmotivation insbesondere hier angesprochen, was die vergleichsweise engeren Zusammenhänge zwischen den Fehlerraten der Übungsblöcke beider IATs und SDO erklärt

Ebenso erwartungskonträr gerichtet ist die positive Korrelation zwischen Sozialer Erwünschtheit und dem Fehlereffekt des IAT Übergewicht. Je sozial erwünschter eine Person ist, umso höhere ist der IAT-Effekt, der an sich eine sozial unerwünschte Ablehnung übergewichtiger Personen bedeutet. Dieser Zusammenhang basiert, in ähnlicher Weise wie die beschriebene Beziehung zwischen Fehler-IAT und SDO, hauptsächlich auf signifikanten Unterschieden in den Fehlerraten des inkompatiblen Übungsblocks in Abhängigkeit von der Ausprägung sozialer Erwünschtheit. Hier bestehen für die Fehlerraten der kompatiblen Blöcke positive Korrelationen zu sozialer Erwünschtheit, während nach dem Aufgabenwechsel eine Umkehrung des Vorzeichens festzustellen ist und der negative Zusammenhang für den Übungsblock statistische Bedeutsamkeit erlangt. Je sozial erwünschter eine Versuchsperson angibt zu sein, desto mehr Fehler macht sie bzw. er dort, wo es aufgrund der Inkompatibilität auch zu erwarten ist, jedoch ungeachtet ihrer expliziten Einstellungen. Ein Vergleich mit den Ergebnissen von Rowatt, Franklin & Cotton (2005) erbringt zumindest ebenfalls positiv gerichtete, wenn auch nicht signifikante Korrelationen zwischen sozialer Erwünschtheit (Balanced Inventory of Desirable Responding, BIDR; Paulhus & Reid, 1991) und dem IAT; für die Skala Selbsttäuschung wird eine Korrelation von  $r=.09$  berichtet, für Fremdtäuschung von  $r=.04$ .

Die reaktionszeitbasierten IAT-Effekte unterhalten im Gegensatz zu den Fehlereffekten keine statistisch bedeutsamen Beziehungen zu den Prädiktorkonstrukten. Lediglich die Korrelation zwischen IAT Übergewicht und Autoritarismus ist tendenziell bedeutsam. Die aufgezeigten systematischen Unterschiede in den Fehlerraten des IAT nach dem Aufgabenwechsel bei gleichzeitiger Unabhängigkeit der Reaktionszeiten zeugen möglicherweise von einer gewissen Steuerbarkeit des IAT-Fehlereffekts. In Abhängigkeit von anderen Merkmalen (z.B. Soziale Erwünschtheit, Dominanzorientierung), verfolgen Probanden unter Umständen eine verstärkt auf Fehlervermeidung oder -zulassung basierende Antwortstrategie, wenn das Risiko von Fehlern, z.B. nach einem Aufgabenwechsel, steigt. Dies erregt Zweifel an der Validität des IAT als unkontrollierbares, unbewusstes implizites Einstellungsmaß und unterstützt die theoretischen und empirischen Hinweise für alternative Funktionsprinzipien bzw. Aufga-

benstrategien des IATs (vgl. Brendl, Markman & Messner, 2001; McFarland & Crouch, 2002; Mierke & Klauer, 2001, 2003; Rothermund & Wentura, 2001, 2004).

Für den EAST ist die Betrachtung manifester Korrelationen aufgrund der mangelnden Reliabilität der Effekte noch kritischer zu bewerten, als für den IAT. Hier sind statistisch bedeutsame Zusammenhänge, im Gegensatz zur Befundlage für den IAT, lediglich für die reaktionszeitbasierten Effekte zu diskutieren. Der Aufgabenwechsel im EAST erfolgt auch nicht derart berechenbar wie im IAT, da im EAST für jeden dargebotenen Stimulus die Kompatibilität bzw. Inkompatibilität manipuliert werden kann. Relativ konsistent sind die positiven Zusammenhänge zwischen den beiden Skalen des BIDR und dem EAST Übergewicht sowie Homosexualität, mit Ausnahme der nicht signifikanten Beziehung zwischen der Selbsttäuschung und dem EAST Homosexualität. Das heißt, die sich zur Selbst- und Fremdtäuschung erwünscht verhaltenden Probanden weisen größere Differenzen zwischen inkompatiblen und kompatiblen Stimuli im EAST Übergewicht auf, während dies im EAST Homosexualität nur für Fremdtäuschung gilt. Diese Zusammenhänge sind nicht auf entweder die kompatiblen oder die inkompatiblen Stimuli zurückzuführen, sondern ergeben sich zumindest für einen der jeweils gemittelten Blockeffekte. Soziale Erwünschtheit steht zwar nicht signifikant, jedoch durchweg positiv zu den reaktionszeitbasierten EAST-Effekten in Beziehung (jeweils  $r=.12$ ).

Die negative Korrelation zwischen dem EAST Übergewicht und Konservatismus beruht auf den signifikant kürzeren Reaktionszeiten bei inkompatiblen Stimuli im ersten Block, da die Reaktionszeitdifferenz (EAST-Effekt) zwischen inkompatiblen und kompatiblen Stimuli umso kleiner ist, je geringer die Reaktionszeiten bei inkompatiblen Stimuli sind (bzw. je höher die Reaktionszeiten der kompatiblen Stimuli ausfallen).

Bedeutsame Geschlechtseffekte sind hauptsächlich für die expliziten Einstellungen zu verzeichnen. Wie bereits vielfach dokumentiert, weisen die männlichen Probanden in allen individuellen expliziten Maßen signifikant negativere Einstellungen zu Homosexualität auf, als die weiblichen (z.B. Hayes, 1995; Herek, 1994; Steffens & Wagner, 2004). Auch die Valenz der Homosexualitätsgeschichten fällt bei Männern bedeutend negativer aus, im Vergleich zu Frauen. Für die impliziten Maße sind dagegen keine statistisch bedeutsamen Geschlechtsunterschiede festzustellen. Für Einstellungen zu Übergewicht erscheint das geschlechtsabhängige Zusammenhangsmuster etwas weniger konsistent. Hier hat das Geschlecht unterschiedliche Bedeutungen für die selbstberichteten Einstellungen. In den affektiven individuellen Einstellungen weisen die Männer deutlich negativere Haltung übergewichtigen Personen gegenüber auf, während die Frauen nahezu durchweg negativere Einstellungen für die Gesellschaft *antizipieren*, jedoch nicht zwangsläufig andere individuelle Einstellungen aufzeigen, als die Männer. Für die Geschichtenmessung und die impliziten Maße zum Einstellungsobjekt Übergewicht zeichnen sich keine Geschlechtseffekte ab.

Das Alter der Probanden spielt für die expliziten Einstellungen eine untergeordnete Rolle. Leicht negativere affektive Einstellungen zu Homosexualität weisen wie erwartet ältere Probanden auf, während die jüngeren Probanden weniger positive Empfindungen gleichermaßen gegenüber homo- und heterosexuellen Personen im Gefühlsthermometer zum Ausdruck bringen. Ältere Probanden schreiben zudem negativere Homosexualitätsgeschichten, als jüngere, sie weisen jedoch etwas geringere EAST-Effekte für die Reaktionszeiten und Fehlerraten des ersten Blocks im EAST Homosexualität auf.

Für Übergewicht ist in Abhängigkeit vom Alter lediglich eine die negativere Bewertung untergewichtiger Personen im Gefühlsthermometer bei jüngeren Probanden relevant. In den Geschichtenvariablen für Übergewicht spiegeln sich keine einstellungsrelevanten Alterseffekte wider, ähnliches gilt für die impliziten Einstellungen. Hier weisen ältere Probanden einen höheren EAST-Effekt auf.

Das nomologische Netzwerk impliziter Einstellungen weicht also gravierend von denen der selbstberichteten und via Geschichtenmessung erfassten Einstellungen ab, während Selbstbericht und Geschichtenmessung deutliche Gemeinsamkeiten in den hypothesenkonformen Zusammenhangsmustern aufweisen. Für die impliziten Einstellungen sind häufig erwartungskonträr gerichtete Zusammenhänge zu verzeichnen, wobei sich diese „Erwartungen“ mehr oder weniger von dem bereits gut untersuchten nomologischen Netzwerk expliziter Einstellungen ableiten müssen, da Informationen zur Verortung impliziter Einstellungen hinsichtlich anderer Korrelate nicht ausreichend vorliegen. Die expliziten Einstellungskorrelate als Referenz für die impliziten Einstellungen zu betrachten, ist gemäß vorliegender Befunde offensichtlich nicht zulässig. Nichtsdestotrotz sind eine Reihe an signifikanten Zusammenhängen zwischen IAT bzw. EAST und selbstberichteten Prädiktorkonstrukten zu dokumentieren. Konsistenz weisen diese Korrelationen insofern auf, als dass SDO beständig negativ mit impliziten Fehlereffekten gemäß IAT korreliert und verschiedenen Maße sozialer Erwünschtheit durchweg positiv mit IAT- und EAST-Effekten in Beziehung stehen.

Es wäre prinzipiell denkbar, dass beispielsweise die positiven Zusammenhänge zwischen impliziten Einstellungsmaßen und sozialer Erwünschtheit gleichermaßen auf einer Art impliziten Verständnisses des Experiments im Sinne der fokussierte Zielkonzeptkategorie (Übergewichtige bzw. Homosexuelle) basieren. Um diese Hypothese zu überprüfen, müssen jedoch noch eine Reihe kontrollierter Experimentalstudien durchgeführt werden. Es wird jedoch auch deutlich, dass die Zusammenhänge zwischen impliziten Maßen und anderen Konstrukten häufig auf der einseitigen Abhängigkeit einer der voneinander abzuziehenden Ausgangsvariablen beruhen. An dieser Stelle seien nicht dokumentierte, jedoch eventuell zukunftsweisende Ergebnisse erwähnt, die ein besseres Verständnis der Bedeutung impliziter Einstellungsmessung ermöglichen. Einen kurzen Ausblick dazu liefert folgender Abschnitt.

### 6.2.3 Weitere Befunde

Wie in Studie III gezeigt wurde, stehen die impliziten Differenzmaßen nicht durchweg mit den expliziten Einstellungen in Beziehungen, jedoch werden signifikant positive Korrelationen von überwältigender Stabilität zwischen den Reaktionszeitmittelwerten (blockweise und gesamt) des IAT und den individuellen *kognitiven* Einstellungen zu Übergewicht und Homosexualität festgestellt. Negative explizite Einstellungen gehen demnach systematisch mit höheren Reaktionszeiten im IAT einher. Hier bestehen Zusammenhänge auch über die Einstellungsobjekte hinweg. Für die Anzahl von Fehlern (blockweise und gesamt) im IAT sind zunächst keine signifikanten Zusammenhänge zu den zentralen expliziten Einstellungsmaßen zu verzeichnen, jedoch bestehen mit vergleichbarer Stabilität eine Reihe negativer Korrelationen zwischen der Anzahl falscher Reaktionen im IAT und der Differenzmaße antizipierter und individueller Einstellungen. Das bedeutet, je geringer die Diskrepanz zwischen den selbstberichteten persönlichen und den gesellschaftlich antizipierten Einstellungen ausfällt, desto mehr Fehler macht eine Versuchsperson im Mittel. Für den EAST sind diese Zusammenhänge weniger konsistent. Jedoch zeichnen sich die signifikanten Korrelationen zwischen expliziten Einstellungen und Reaktionszeitmittelwerten durch beständig positives Vorzeichen aus, während die Korrelationen mit der Anzahl an Fehlern prinzipiell negativ gerichtet sind.

Zur Interpretation des zunächst konsistenten Befundes positiver Zusammenhänge zwischen expliziten Einstellungen und Reaktionszeiten können zwei unterschiedliche Ansätze verfolgt werden. Einerseits ist es nahe liegend, als Erklärungsansatz den Emotionalen Stroop-Effekt in Erwägung zu ziehen (vgl. Williams, Matthews & MacLeod, 1996). Als Emotionaler Stroop Effekt wird die momentane Aufmerksamkeitsbündelung durch emotional negative Reize bezeichnet, aus der eine Verlangsamung der Reaktionen auf diese Reize resultiert (z.B. Suslow, Ohrmann & Arolt, 2001). Pratto & John (1991) nehmen an, dass der Emotionale Stroop Effekt einen automatischen Mechanismus darstellt, während Bargh (1997) von einem vorbewussten Mechanismus spricht. Die hier aufgezeigte Beziehung zwischen negativen Einstellungen zu Übergewicht und Homosexualität und längeren gemittelten Reaktionszeiten in den Experimenten wäre demnach mit der unwillkürlichen Aufmerksamkeitsabsorption durch das emotional negativere Stimulusmaterial der Übergewichts- bzw. zu begründen. Zur Untermauerung dieses Ansatzes, der zwar nicht das IAT-Differenzmaß als valide ausweist, wohl aber einen testbaren Bezug zwischen reflektierten und spontanen Einstellungsanteilen darstellt, bedarf es noch einer Reihe an experimentellen differentialpsychologischen Untersuchungen.

Eine alternative Erklärung für den konsistenten Befund negativer Zusammenhänge zwischen expliziten Einstellungen und Reaktionszeiten könnten moderierende Drittvariablen wie z.B. das Alter oder die Intelligenz liefern. Höheres Alter geht zum Einen mit höheren Konservatismusausprägungen einher (z.B. Denz, 1988) und in diesem Zusammenhang auch mit negativeren Einstellungen zu normativ

bewerteten Minderheiten wie z.B. Homosexuellen (z.B. Bernat, Calhoun, Adams & Zeichner, 2001; Hayes, 1995; Herek, 1994; Morrison, Parriag & Morrison, 1999; Steffens & Wagner, 2004). Zum anderen ist mit steigendem Alter ein Abbau kognitiver Verarbeitungsgeschwindigkeit nachgewiesen worden (vgl. Meyer, Glass, Mueller, Seymour & Kieras, 2001). Die Intelligenz beschreibt einerseits die interindividuellen Unterschiede in der kognitiven Kapazität, die sich u.a. in der Verarbeitungsgeschwindigkeit niederschlagen, wobei Intelligenz positiv mit kognitiver Verarbeitungsgeschwindigkeit in Beziehung steht (z.B. Neubauer, 1997). Zum anderen sind Intelligenz bzw. Bildungsniveau ebenfalls stabile negative Vorurteilskorrelate (vgl. Zick, 1997). Sowohl das Alter, als auch die Intelligenz weisen in Studie III allerdings einen eingeschränkten Variationsbereich auf, was das Erklärungspotential eines Moderatoransatzes hinsichtlich der Beziehung zwischen expliziten Einstellungen und Reaktionszeitmittelwerten abschwächt. Das mittlere Alter der Versuchsteilnehmer beträgt etwa 22 Jahre mit einer Standardabweichung von weniger als 4 Jahren, somit besteht wenig Anlass zur Annahme gravierender Unterschiede in der kognitiven Verarbeitungsgeschwindigkeit der Probanden. Die Stichprobe studentischer Probanden kann zudem als *Bildungsstichprobe* bezeichnet werden, die sich gemäß Mehrfachwortwahltest durch eine hohe mittlere Intelligenz auszeichnet.

Die negativen Zusammenhänge zwischen den Differenzmaßen persönlicher bzw. antizipierter Einstellungen und der Höhe der Fehler in den impliziten Einstellungsmessungen treten weniger konsistent auf und sind zumeist auf entgegen gesetzt gerichtete Korrelationen zwischen den Ausgangsvariablen individueller bzw. antizipierter Einstellungen und den Fehlern zurückzuführen. Möglicherweise reflektieren diese Befunde schlichtweg motivationale oder kapazitive Differenzen zwischen den Probanden hinsichtlich der Untersuchungsteilnahme. Eine geringe Motivation (bzw. Kapazität) könnte sich entsprechend in minimalen Bemühungen hinsichtlich des Selbst- und Quasi-Fremdberichts zu gesellschaftlichen Einstellungen äußern, woraus eine geringe Diskrepanz zwischen individuellen und antizipierten resultieren sollte. Zudem wäre zu erwarten, dass unmotivierte bzw. überforderte Probanden in den impliziten Experimenten mehr Fehler produzieren. Es bleibt zu klären, inwieweit eine inhaltliche Interpretation der beschriebenen Zusammenhänge auf der Basis motivationaler bzw. kognitiver Kapazitätsfaktoren statthaft ist. Entsprechende Klarheit können weitere Untersuchungen erbringen.

#### 6.2.4 Zusammenfassung

Mit Studie III sind einige zentrale Fragestellungen vorliegender Arbeit untersucht worden. Hauptziel dabei stellte die Konstruktvalidierung expliziter und impliziter Einstellungen dar. Gemäß der empirischen Datenauswertung ist eher eine konzeptuelle Distinktheit expliziter und impliziter Einstellungen, im Sinne von Repräsentanten unabhängiger Systeme anzunehmen (z.B. Devine, 1989). Diese Ergebnis-

se stützen sich maßgeblich auf den Selbstbericht, die Geschichtenmessung und den IAT. Hierbei ist jedoch zu bedenken, dass einerseits die Validität des IAT zur Messung impliziter Einstellungen *prinzipiell* angezweifelt wird (vgl. Brendl, Markman & Messner, 2001; McFarland & Crouch, 2002; Mierke & Klauer, 2001, 2003; Rothermund & Wentura, 2001, 2004) und andererseits auch die Befunde vorliegende Arbeit die Zweifel an der Validität des Verfahrens nähren.

Neben den kritischen Befunde zum IAT (Karpinski & Hilton, 2001) konnten jedoch einige Studien die differentielle Bedeutung des IAT im Hinblick auf spontanes bzw. unkontrollierbares Verhalten untermauern (z.B. Asendorf, Banse & Mücke, 2002; Boson, Swann & Peenebaker, 2000; Egloff & Schmukle, 2004; Phelps et al., 2000). In zukünftigen Untersuchungen sollten daher neben der Betrachtung von mindestens 2 Einstellungsobjekten oder Dispositionen auch konkrete *Verhaltensmaße* zur Validierung expliziter und impliziter Einstellungsmaße herangezogen werden.

## 7 Gesamtdiskussion

Vorliegende Arbeit ist der Untersuchung expliziter und impliziter sozialer Einstellungen, ihrer Konstruktvalidität, ihrer direkten bzw. vermittelten Beziehungen untereinander sowie ihrer Bedeutung für andere einstellungsrelevante Konstrukte gewidmet. Zu diesem Zweck wurden Einstellungen zu zwei unterschiedlichen Inhalten, und zwar Übergewicht und Homosexualität, mittels verschiedener Verfahren gemessen. Zum Einsatz kamen in vorliegender Arbeit traditionelle Selbstberichte, für die zum Teil eine Trennung zwischen persönlichen und antizipierten Einstellungsinhalten erfordert wurde und zwei implizite Messverfahren, wovon der Implizite Assoziationstest als relativ etabliert und gut untersucht bezeichnet werden kann (vgl. Gawronski & Conrey 2004), wogegen der Extrinsic Affective Simon Task hauptsächlich aufgrund eines Potentials, zur Erweiterung des diesbezüglichen Kenntnisstandes und zum Vergleich mit dem IAT einbezogen wurde. Zusätzlich wurde im Rahmen der verfolgten Fragestellungen ein Messverfahren zur Erfassung reflektierter, bewusster *und* spontaner, unbewusster Einstellungsinhalte entwickelt und gemeinsam mit expliziten und impliziten Einstellungsmessungen eingesetzt.

Zwei erste Untersuchungen waren zur Entwicklung und Erprobung der, in der dritten Studie schließlich gemeinsam eingesetzten, Einstellungsmessungen angelegt. Mit Studie I wurde maßgeblich eine Untersuchung der psychometrischen Eigenschaften der beiden ausgewählten impliziten Einstellungsmaße am Beispiel von Einstellungen zu Homosexualität verfolgt. Hierbei wurden insbesondere die Auswirkungen einer festgelegten Stimulussequenz für die Reliabilität der impliziten Maße und schließlich die Analyse interindividueller Differenzen in expliziten und impliziten Einstellungen erörtert. Den Ergebnissen von Studie I zufolge, weist der reaktionszeitbasierte IAT-Effekt eine vergleichbar hohe interne Konsistenz auf und korreliert signifikant positiv mit expliziten Einstellungen zu Homosexualität. Der Fehlereffekt des IAT ist deutlich weniger reliabel, unterhält jedoch ebenso positive Zusammenhänge, allerdings nicht signifikante, Beziehungen zu den expliziten Kriterien. Dagegen ist für den EAST keine hinreichende interne Konsistenz gegeben. Es wird sehr deutlich, dass jeweils nur Teileffekte basierend auf einzelnen Messblöcken betrachtet werden können und hier insbesondere der erste Block erwartungskonforme Zusammenhänge zu expliziten Einstellungen aufzeigt (siehe Studie I).

Studie II diente der Entwicklung einer offenen Assoziativmessung, mit dem sowohl explizite, als auch implizite Aspekte von Einstellungen zu Übergewicht bzw. Homosexualität erfasst werden können. In Anlehnung an den Thematischen Apperzeptionstest der als projektive Verfahren insbesondere zur Motivmessung eingesetzt wird, wurden thematisch relevante Bildmotive erstellt, zu denen Probanden eine frei erfundene Geschichte schreiben sollten, die mittels adäquater Kodierschemata inhaltsanalytisch ausgewertet wurden. Zur Evaluation dieser Messung wurden expliziten Einstellungen zu Übergewicht und Homosexualität mit erhoben. Ein Hauptergebnis von Studie II ist das Potential der aus-

gewählten Messmethodik zur Erfassung von Einstellungen. Konsistente Zusammenhänge zu expliziten Einstellungen und einstellungsrelevanter Motivation werden für die Geschichtenlänge in Form der Wortanzahl verzeichnet. Darüber hinaus sind eine Reihe sinnvoller Zusammenhänge für die Valenz der Geschichten festzustellen. Ziel der Evaluation der Bildmotive war schließlich die Auswahl der jeweils zwei am besten geeigneten Motive je Einstellungsobjekt.

Den Kern vorliegender Arbeit bildet Studie III, in der die verschiedenen Einstellungsmessungen gemeinsam in einer Stichprobe von 165 Personen erhoben wurden, wodurch die Untersuchung einer Reihe zentraler Fragestellungen zur Bedeutung expliziter und impliziter Einstellungen zum Teil mit messfehlerbereinigten Variablen erfolgen konnte. Kurz zusammengefasst liefert Studie III Hinweise auf eine notwendige Unterscheidung zwischen expliziten und impliziten Einstellungen, für die keine konvergente Validität angenommen werden kann. Lediglich für Homosexualität korrespondieren IAT und Selbstbericht in statistisch bedeutsamem Ausmaß. Aufgrund fehlender substanzieller Korrespondenzen zwischen expliziten und impliziten Maßen resultiert die herausragende Bedeutung methodenspezifischer Varianz für die Variablen. Dagegen werden trotz unterschiedlicher Messmethoden stabile Korrespondenzen zwischen Selbstbericht und Geschichtenvariablen verzeichnet. Es ist jedoch nicht von einer Art Explizit-Implizit-Dimension der Einstellungen auszugehen bzw. zeugen die Beziehungsmuster nicht von einer relativen Zwischenstellung der offenen Assoziativmessung hinsichtlich expliziter und impliziter Einstellungsanteile. Weder die schwachen Zusammenhänge zwischen expliziten und impliziten Einstellungen, noch die engen Beziehungen zwischen expliziten und offen assoziierten Einstellungen lassen sich durch Motivation zur Vorurteilsfreiheit oder Soziale Erwünschtheit erklären. Lediglich für affektive Einstellungen zu Homosexualität und den IAT wird eine zusammenhangssteigernde Wirkung durch die individuelle Motivation zur Vorurteilsfreiheit festgestellt. Dagegen kann gezeigt werden, dass die Berücksichtigung sozialer Erwünschtheit im Fragebogen in Form einer forcierten Trennung zwischen individuellen und gesellschaftlich antizipierten Einstellungen zu zwei distinkten Einstellungsdispositionen führt, deren Abweichung insbesondere für die Beziehung expliziter und impliziter Einstellungen zu Homosexualität bedeutsam ist. Schließlich werden gravierende Unterschiede zwischen den nomologischen Netzwerken expliziter und impliziter Einstellungen aufgezeigt. Während selbstberichtete Einstellungen und implizite Assoziationsstärkemaße nur teilweise in Zusammenhang stehen (hier für Einstellungsobjekt Homosexualität), sind die konsistenten und nicht auf der Wirkung motivationaler Drittvariablen basierenden Korrespondenzen zwischen der Geschichtsmessung und den selbstberichteten expliziten Einstellungen im Fragebogen für beide Einstellungsobjekte hervorzuheben.



Zur Evaluation der Ergebnisse aktueller Arbeit wird im Folgenden eine resümierende Gesamtbetrachtung vorgenommen, der zufolge mögliche Generalisierungen und weiterführende Fragestellungen abgeleitet werden. Im zentralen Fokus der Gesamtdiskussion steht die Konstruktvalidierung sozialer Einstellungen, die Anhaltspunkte zur Konzeptualisierung von expliziten und impliziten Einstellungen liefert. Da adäquate Messverfahren eine Grundvoraussetzung zur Untersuchung inhaltlicher Fragestellungen darstellen, liegt auf einer Bewertung der eingesetzten Verfahren zur Einstellungsmessung ebenfalls großes Gewicht. In diesem Zusammenhang wird eine Evaluation der impliziten Einstellungsmessungen IAT und EAST hinsichtlich ihrer psychometrischen Eigenschaften und der daraus resultierenden Eignung zur Untersuchung interindividueller Differenzen in impliziten Einstellungen vorgenommen. Ein weiterer Diskussionspunkt zur Verfahrensevaluation betrifft die offene Assoziativmessung, die im Rahmen dieser Arbeit zur Erfassung impliziter und expliziter Einstellungsanteile entwickelt wurde. Schließlich werden mögliche Charakteristika der hier betrachteten Einstellungsobjekte Übergewicht und Homosexualität hinsichtlich ihrer Bedeutung für die Ergebnisse der empirischen Untersuchungen dieser Arbeit diskutiert.

## 7.1 Konzeptualisierung von Einstellungen

Mit der Arbeit von Banaji & Greenwald (1995) zur Unterscheidung zwischen expliziten und impliziten Einstellungen und der Publikation des Impliziten Assoziationstestes etwa 4 Jahre später (Greenwald, McGhee & Schwartz, 1998) begann eine geradezu euphorisierte Untersuchung impliziter und expliziter Anteile von *Einstellungen*, *Stereotypen*, *Vorurteilen* (vgl. Devine, 2001), von *Identität*, *Selbstkonzept* und *Selbstwert* (z.B. Aidman & Carroll, 2003; Asendorpf, Banse & Mücke, 2002; Greenwald & Farnham, 2000; Greenwald et al., 2002) oder von dysfunktionalen Überzeugungen bzw. Dispositionen zu *Angst* und *Furcht* (De Jong, 2002; De Jong, Pasma, Kindt, van den Hout, 2001; Egloff & Schmukle, 2002; Schmukle & Egloff, 2004), *Depression* (Gemar, Segal, Sagrati & Kennedy, 2001) oder *Suchtverhalten* (De Houwer et al., 2004; Wiers, van Worden, Smulders & de Jong, 2002).

Theoretische Konzeptualisierungsansätze für explizite und implizite Einstellungen stehen einer äußerst heterogenen Befundlage zur Beziehung zwischen beiden Einstellungsaspekten gegenüber. Im Mittel kann von schwach positiven Zusammenhängen zwischen zumindest IAT-basierten impliziten und selbstberichteten expliziten Einstellungen ausgegangen werden (Hofman et al., 2003), die zu einem beträchtlichen Teil von Drittvariablen moderiert werden. Nosek (2004) findet beispielsweise eine Varianzaufklärung von nahezu 40% an der Beziehung expliziter und impliziter Einstellungen durch die vier Moderatorvariablen *Selbstpräsentation*, *Einstellungsstärke*, *Einstellungsdimensionalität* und *wahrgenommene Abweichung der persönlichen Einstellungen von den Einstellungen der Bezugsgruppe* (vgl. Banse & Gaw-

ronski, 2003; Devine et al., 2002; Hofmann, Gschwendner & Schmitt, 2005; Plant & Devine, 1998). Vorliegende Arbeit reflektiert die inkonsistente Befundlage zur Beziehung impliziter und expliziter Einstellungen sehr anschaulich. Den Ergebnissen von Studie I zufolge, stehen impliziten Einstellungen zu Homosexualität gemäß IAT und EAST<sup>4</sup> in signifikant positivem Zusammenhang mit expliziten Einstellungen zu Homosexualität, wobei eine differentielle Validität für IAT und EAST in Bezug auf kognitive und affektive Einstellungen festzustellen ist (siehe aber Abschnitt 7.2.1). Während der IAT vor allem kognitive Einstellungen zu Homosexualität vorhersagt, erweist sich der EAST als Hauptprädiktor affektiver Einstellungen. Das heißt, für explizite und implizite Einstellungen zu Homosexualität ist eine erwartungskonforme Korrespondenz anzunehmen, die auf konvergente Validität der verschiedenen Einstellungsmessungen zu Homosexualität hindeutet. Statthafte Aussagen über konvergente und diskriminante Validität expliziter und impliziter Einstellungen sind an dieser Stelle jedoch nicht zu treffen, da hierzu mehr als ein Einstellungsinhalt betrachtet werden muss. Um Anhaltspunkte für eine theoretische Konzeptualisierung expliziter und impliziter Einstellungen im Sinne einer *Lügendetektorperspektive* oder einer *Perspektive unabhängiger Bewertung* (Nosek & Banaji, 2002; vgl. Kapitel 2.1.1) zu gewinnen, bedarf es entsprechend einer Konstruktvalidierung. Dies wurde in Studie III mit einer latenten Multitrait-Multimethod-Analyse expliziter und impliziter Einstellungen zu Übergewicht und Homosexualität vorgenommen.

Nach den klassischen Konventionen konvergenter und diskriminanter Validität zu urteilen, ist allein für offen assoziierte und explizite Einstellungen Konstruktvalidität gegeben. Implizite Einstellungen konvergieren nicht hinreichend in einem gemeinsamen Einstellungsfaktor mit expliziten oder offen assoziierten Einstellungen. Dafür können verschiedene Ursachen diskutiert werden, die jedoch zu meist darauf hinweisen, dass implizite Einstellungsmessverfahren wie der IAT oder EAST die psychometrischen Anforderungen interindividueller Differenzmaße nicht ausreichend erfüllen und für eine Einstellungsmessung im Sinne überdauernder Disposition nicht geeignet erscheinen (siehe dazu Abschnitt 2.1.1).

Die resultierenden Einstellungsmaße aus den hier einbezogenen impliziten Messungen, wovon der IAT ein ausgesprochen populäres und weitestgehend etabliertes Verfahren darstellt, werden als Indikatoren relativer kognitiver Assoziationsstärken interpretiert. Diese Maße sollten unbewusste, unreflektierte und unkontrollierbare Kognitionsinhalte reflektieren, die sich durch ein hohes Maß an Stabilität auszeichnen (z.B. Devine, 1989; Wilson, Lindsey & Schooler, 2000). Zum einen können die Annahmen zur Unbewusstheit, Unreflektiertheit und Unkontrollierbarkeit impliziter Maße soweit noch nicht als gesichert gelten (vgl. De Houwer, in press). Zum anderen zeichnen sich vor allem Messungen

---

<sup>4</sup> Hier jedoch wird lediglich der *erste* von vier EAST-Blöcken betrachtet, da der Gesamteffekt gemäß den vier Messblöcken keine befriedigende interne Konsistenz aufweist und der Effekt des ersten Blocks erwartungskongruente Zusammenhänge zu IAT und EAST aufzeigt. Diese *post-hoc*-Auswahl des ersten EAST-Blocks ist ungeachtet der externen Zusammenhänge sehr unbefriedigend und kann nicht zu einer allgemeingültigen, positiven Evaluation des Verfahrens reichen.

*expliziter* Einstellungen durch in aller Regel hohe Stabilitäten aus (siehe unten), so dass eher die Rolle des Bewusstseins als Stabilitätsfaktor für Einstellungsdispositionen diskutiert werden müsste.

Es ist davon auszugehen, dass allein von *einem* Fragebogenitem ausgehend, ein umfangreiches assoziatives Netzwerk aktiviert werden sollte, bevor eine Person dem Grad ihrer Zustimmung bzw. Ablehnung als Resultat eines Entscheidungsprozesses auf einer zweidimensionalen Antwortskala Ausdruck verleiht. Es ist insofern nahe liegend, anzunehmen, dass die Prozesse, die der Beantwortung eines Items zugrunde liegen, komplexer Natur sind und sowohl kognitive, als auch emotionale und motivationale Anteile reflektieren. Über die Items einer Skala zur Messung von Einstellungen hinweg werden somit unterschiedliche Aspekte zum betrachteten Einstellungsobjekt erfasst, die aufgrund des stark reduzierten Antwortformats in der Regel eine hohe interne Konsistenz aufweisen. Im Rahmen der Antwortskala hat der Proband die freie Reaktionswahl. Im Vergleich dazu werden in den hier eingesetzten impliziten Messverfahren bildhafte Reize in kurzer Abfolge dargeboten, wobei die Versuchspersonen nur zwei Antwortmöglichkeiten zur Verfügung haben, ohne jedoch eine wirkliche freie Wahl über die Reaktion, denn es ist genau festgelegt, welcher Stimulus, mit welcher Taste zu beantworten ist. Das bedeutet, eine Person ist dazu angehalten, anhand der normativen Antwortvorgaben und der Instruktion, so schnell und so korrekt wie möglich zu reagieren, eine Art Leistungstest abzulegen. Interpretiert wird das Messergebnis jedoch nicht hinsichtlich der tatsächlichen Leistung, z.B. niedrige Reaktionszeiten bei geringen Fehlerraten, sondern als indirekter Einstellungsindikator. Zudem werden die impliziten Maße in aller Regel durch Differenzbildung gewonnen. Ungeachtet der theoretischen Unterscheidung zwischen expliziten und impliziten Einstellungen (z.B. Banaji & Greenwald; 1995; Devine, 1989; Rudman, 2004; Wilson, Lindsey & Schooler, 2000) bestehen also bereits hinsichtlich der Prozesse, auf denen die notwendigen Urteilsbildungen, Antwortentscheidungen bzw. Reaktionen basieren, ebenso große Welten zwischen expliziten und impliziten Einstellungsmessungen, wie letztlich in Bezug auf die statistischen Grundlagen der betrachteten Einstellungsmaße.

Die dritte eingesetzte Messmethode in dieser Arbeit wird als offene Assoziativmessung bezeichnet und sollte Aspekte expliziter und impliziter Einstellungen erfassen. Es werden jedoch lediglich Beziehungen zwischen der offenen Assoziativmessung und dem Selbstbericht verzeichnet, weshalb keine Verortung der Geschichtenmessung zwischen Selbstbericht und IAT bzw. EAST nahe liegt, sondern eher eine Zwischenstellung des Selbstberichtes, der sowohl zu impliziten Messungen, hier hauptsächlich IAT, als auch zur offenen Assoziativmessung in Zusammenhang steht. Dies ist insofern nachvollziehbar, als dass Selbstbericht und IAT beiderseits eine explizite Thematisierung der Thematik vornehmen – der Selbstbericht notwendigerweise in den Itemformulierungen und der IAT in Form der Kategorienlabel. Wogegen Selbstbericht und Geschichtenmessung, bei der bis auf anleitende Fragen zum Strukturieren der Geschichten keine Vorgaben gemacht werden, das wahlfreie, stärker kontrollierbare, verbale Antwortformat gemein haben. Dies betrifft jedoch durchweg messmethodische As-

pekte, die relativ genau zu benennen sind, während die viel kritischeren inhaltlichen Übereinstimmungen bzw. Unterschiede der verschiedenen Einstellungsmessungen ungeklärt bleiben, so lange keine gemeinsame externe Validierung der drei unterschiedlichen Einstellungsmaße erfolgt.

Gemäß vorliegender Arbeit lassen sich über die externe Validität der Einstellungsmaße keine Schlüsse ziehen, es wird lediglich zunächst herausgestellt, dass Konstruktvalidität im konventionellen Sinne für explizite und implizite Einstellungen nicht gegeben ist. Stattdessen legen die Ergebnisse der MTMM-Analyse das Vorherrschen zweier Methodenfaktoren nahe. Das legt nahe, den Dissoziationsansatz von Devine (1989) bzw. Wilson, Lindsey & Schooler (2000) als adäquate Konzeptualisierung expliziter und impliziter Einstellungen anzunehmen. Jedoch kann ein statthafter Konzeptionsansatz nur auf der Basis reliabler und valider Messinstrumente erfolgen. Die Reliabilitätsproblematik erscheint, gemessen an der Häufigkeit bzw. Seltenheit publizierter Reliabilitätsbefunde, für viele psychologische Disziplinen weniger zentral, wobei es eine Grundvoraussetzung für die Untersuchung differentialpsychologischer Fragestellungen darstellt und deshalb in vorliegender Arbeit entsprechendes Augenmerk verdient. Die Validität impliziter Maße wie dem IAT ist nach wie vor nicht hinreichend bestätigt. Zwar liegen einige richtungweisende Befunde vor, dennoch sind immer noch alternative Erklärungsansätze für den IAT nicht ausgeräumt, wie die Hypothese der Reaktionsschwellenanpassung von Brendl, Markman & Messner (2001), die Hinweise auf den Einfluss exekutiver Kontrollprozesse beim IAT (Mierke & Klauer, 2001, 2003) oder die Rolle von Salienzunterschieden, die den IAT-Effekt über das Prinzip der Figur-Grund-Asymmetrie formen (Rothermund & Wentura, 2001, 2004)

Es bleibt demnach zu bedenken, dass Aussagen zur Konstruktvalidität expliziter und impliziter Einstellungen nur maximal so gültig sein können, inwieweit es Reliabilität und Validität der zugrunde liegenden Maße es erlauben. Zu den psychometrischen Eigenschaften der eingesetzten Einstellungsmessverfahren erfolgt nun eine resümierende Ergebnisdiskussion.

## **7.2 Evaluation der impliziten Einstellungsmessungen**

Einstellungen sind als stabile, dem Verhalten zugrunde liegend Konstrukte konzipiert worden (z.B. Allport, 1935; Beck, 1976). Die traditionelle Messung von Einstellungen im Form von Selbstberichten unterliegt zwar einer Reihe von Schwächen, wie einleitend aufgezeigt wurde (Abschnitt 1.5), dennoch ist für die Messungen von Einstellungen mit Fragebogen in der Regel eine hohe Zuverlässigkeit z.B. in Form interner Konsistenz oder zeitlicher Stabilität gegeben (z.B. Cunningham, Preacher & Banaji, 2001; Steffens & Buchner, 2003). Im Gegensatz dazu, zeichnen sich die implizite Einstellungsmessungen vielfach durch unbefriedigende Reliabilitäten aus (z.B. Banse, 2001; Hofmann et al., in press; Teige et al., 2004). Ungeachtet des empirisch untersuchten Gegenstandes sind valide und zuverlässige

Messinstrumente jedoch eine Grundvoraussetzung für die Erforschung psychologischer Fragestellungen. Insbesondere bei der Untersuchung interindividueller Unterschiede, kommt den psychometrischen Eigenschaften von Messinstrumenten eine entscheidende Bedeutung zu, da interindividuelle Differenzen in einem Merkmal wie z.B. sozialen Einstellungen auf tatsächlicher Merkmalsvariation beruhen sollten und nicht hauptsächlich auf Messfehlern. Als Messfehler eines Maßes wird alle unsystematische, unvorhersagbare Varianz bezeichnet, wobei häufig offen bleibt, wie unsystematische Messwertvariation inhaltlich zu interpretieren ist. Störquellen einer jeden Messung können sowohl personeninterne, als auch –externe Einflussfaktoren darstellen. Je höher die Messfehlerbehaftung eines Messwertes, desto eingeschränkter ist seine Interpretierbarkeit. Die Voraussetzung psychometrischer Qualität von Messverfahren muss besonders im Bereich impliziter Einstellungsmessung als kritisch bezeichnet werden.

In den letzten beiden Dekaden lag der Fokus der Einstellungsforschung vor allen Dingen auf der adäquaten Erfassung impliziter Einstellungen, von denen man annahm, die „wahren Einstellungen“ von Personen zu reflektieren. Mit der Entwicklung des Impliziten Assoziationstests durch Greenwald, McGhee & Schwartz (1998) verband sich zunächst für viele sicherlich die Hoffnung, implizite Einstellungsmessung würde den Zugang zu unbewussten und unkontrollierbaren Kognitionsinhalte eröffnen, nachdem bereits Generationen von Einstellungsforschern versuchten, mit diversen methodischen und statistischen Werkzeugen, wie z.B. der Bogus-Pipeline-Methode, unverfälschte Einstellungen untersuchen zu können (für eine Übersicht, siehe Webb, 2000). Dem anfänglich euphorisch betriebenen empirischen Einsatz des IAT für diverse psychologische Fragestellungen und der Neu- bzw. Weiterentwicklung impliziter Messverfahren folgte eine Phase des kritischen Hinterfragens und des resümierenden Überblickens der Befundlage. Doch nach wie vor stehen implizite Maße hochgradig im Interessenfokus, kaum ein psychologisches Forschungsfeld bedient sich nicht der Methodik des IAT, was diesem Verfahren schließlich eine beispielhafte Erfolgsgeschichte verleiht. Vor allem in den vergangenen beiden Jahren gewann jedoch auch die kritische Evaluation des IAT Aufmerksamkeit.

Die implizite Einstellungsmessung allgemein betreffen problematische psychometrische Eigenschaften der Messverfahren, da nicht zuletzt die Reliabilität impliziter Maße als Ursache für fehlende Korrespondenzen zwischen expliziten und impliziten Einstellungen, die mehr oder weniger als interindividuell variierende Dispositionen konzipiert sind, diskutiert wird (vgl. Fazio & Olson, 2003). Für die impliziten Messverfahren lassen sich interne Konsistenzen von zufrieden stellend (IAT) bis völlig unzureichend (EAST) aufführen (vgl. Cunningham, Preacher & Banaji, 2001; De Houwer, 2003a; De Houwer & De Bruycker, 2005; Degner & Wentura, 2005; Ellwart, Rinck & Becker, 2005; Teige et al., 2004). Bisherigen Befunden zur Stabilität impliziter Maße nach, ist von recht geringen Test-Retest-Reliabilitäten auszugehen (Hofman et al., 2003). Die allgemein mangelnde Zuverlässigkeit ist umso kritischer, als dass Reliabilität eine Voraussetzung für die Validierung von Messinstrumenten ist. Hierbei

ist jedoch zu beachten, dass zweifelhaft ist, inwiefern die Gütekriterien der klassischen Testtheorie auf Reaktionszeitparadigmen wie den EAST anzuwenden sind, da nicht davon auszugehen ist, dass sich die systematische Varianz impliziter Maße in der Kongruenz von Reaktionszeit- oder Fehlereffekten der ersten und zweiten Experimentalhälfte ausdrückt. Insofern ist jedoch allenfalls von einer Unterschätzung zumindest der internen Konsistenz auszugehen, was auf die externen Zusammenhänge der impliziten Maße keinen Einfluss hat.

In vorliegender Arbeit können hauptsächlich über die *interne Konsistenz* der Einstellungsmessungen Aussagen getroffen werden. Dabei erweist sich der IAT erneut als zufrieden stellend bis sehr gut reliabel. Der EAST dagegen erfüllt *nicht* die Konsistenzvoraussetzung zur Untersuchung interindividueller Differenzen (Abschnitt. 7.2.1). Die *Stabilität* als spezieller Reliabilitätsaspekt ist zwar für vorliegende Arbeit weniger zentral, jedoch ist der Gesichtspunkt situativer Variabilität aufgrund des Aufbaus und der Ergebnisse von Studie III ebenfalls zu berücksichtigen (Abschnitt. 7.2.2). Schließlich soll die allgemeine Problematik von Differenzmaßen hinsichtlich der Reliabilität der impliziten Einstellungsmessungen diskutiert werden (Abschnitt. 7.2.3).

### 7.2.1 Interne Konsistenz

Für die empirische Untersuchung interindividueller Unterschiede sind prozedurale Variationen im Versuchsaufbau zur „Ausmittlung“ von Reihenfolgeeffekten und Auslegung von Haupteffekten nicht angeraten, da dies eine Interpretation von Unterschieden zwischen Personen erschwert. Aus diesem Grund, wurden die Bedingungen hinsichtlich der Erhebungsprozedur für alle Probanden weitestgehend konstant gehalten. Die Probanden absolvierten entsprechend völlig identische Experimente im Rahmen einheitlicher Untersuchungsdurchführungen. Ziel war es dabei, eine Steigerung der Reliabilität insbesondere für den EAST zu erreichen, da hier gravierende Reliabilitätsmängel bestehen (z.B. Teige et al., 2004).

Der IAT ist – in Übereinstimmung mit der Metaanalyse von Hofmann et al. (2003), die für den IAT eine stark variierende, im Mittel .80 betragende interne Konsistenz ermitteln – als relativ zuverlässig zu bezeichnen. In vorliegender Arbeit werden Reliabilitäten von .87 (Studie I, IAT Homosexualität) bis zu .60 (Studie II, IAT Übergewicht) für die betrachteten IAT-Maße (Reaktionszeiten) verzeichnet.

Für den EAST wird dagegen in Studie I und III eine äußerst unzureichende Zuverlässigkeit festgestellt, weshalb die EAST-Maße als ungeeignet zur Untersuchung interindividueller Differenzen in impliziten Einstellungen bewertet werden müssen. Dieser Befund steht in Kongruenz zu den zitierten EAST-Studien und den bislang noch unpublizierten empirischen Untersuchungen zum EAST (z.B. De Houwer & De Bruycker, 2005; De Jong & Huijding, 2005; Degner & Wentura, 2005; Ellwart,

Rinck & Becker, 2005). Zwei von drei bislang veröffentlichten EAST-Studien berichten sehr unbefriedigende Reliabilitäten. Für ihre Studie zur Messung alkoholbezogener Kognitionen mit IAT und EAST dokumentieren De Houwer et al. (2004) keine Reliabilitäten, wohl aber argumentieren sie hinsichtlich der Nullkorrelationen zwischen IAT und EAST mit der wohl bekannten Reliabilitätsproblematik impliziter Maße bzw. den strukturellen Unterschieden zwischen verschiedenen impliziten Messungen. Zwar konnte in Studie I die Bedeutung eines Teileffekts des EAST (erster Block) herausgestellt werden, nichtsdestotrotz steht der Gesamteffekt weder in Zusammenhang zum IAT, noch zu expliziten Einstellungen bzw. Einstellungsprädiktoren. Eine aufgrund der korrelativen Ergebnisse im nachhinein getroffene Auswahl des ersten EAST-Blocks zur Gegenüberstellung mit dem IAT und Selbstbericht kann *nicht* die Ansprüche wissenschaftlich-empirischen Arbeitens erfüllen. Sofern keine sinnvollen und ad hoc festgelegten Kriterien zur Auswahl eines spezifischen Teileffekts vorliegen, muss der Gesamteffekt als EAST-Maß angenommen werden. Darüber hinaus konnte die in Studie I aufgezeigte Bedeutung des reaktionszeitbasierten Effektes gemäß erstem EAST-Block in Studie III nicht repliziert werden, was dafür spricht, dass der EAST noch kein stabiles Messverfahren darstellt. Eine mögliche Ursache für die geringe interne Konsistenz des EAST ist das niedrige Aggregationsniveau der Effektmaße (vgl. De Houwer, 2003a; Fazio & Olson, 2003; Teige et al., 2004). In den empirischen Studien vorliegender Arbeit basiert der IAT-Effekt beispielsweise auf insgesamt 96 Trials (jeweils 48 Trials in kompatibelem und inkompatiblen Block). Eine vergleichbare Steigerung des Aggregationsniveaus für den EAST würde die Darbietung von insgesamt 192 Trials in einem Block bedeuten. Auch wenn diese Stimulusanzahl auf zwei Blöcke aufgeteilt werden würde, müssten die Probanden ohne Unterbrechung 96 Stimuli beantwortet. Zudem zeichnen sich die EAST-Blöcke durch teilweise sogar gegensätzliche Haupteffekte und sehr inkonsistente Zusammenhänge aus, weswegen es nicht im Sinne einer hohen internen Konsistenz ist, die notwendige Trialanzahl auf zwei Blöcke zu verteilen. Möglicherweise sind die fehlenden oder entgegen gesetzt gerichteten Beziehungen zwischen den Blockeffekten auf interindividuelle Veränderungen der Bearbeitungsstrategie (z.B. schnell vs. korrekt) zurückzuführen. Dies sollte während der kontinuierlichen Stimulusdarbietung eines Blocks nicht möglich sein. Der durchweg festgestellte Befund interner *Inkonsistenz* der Blöcke eines EASTs mag jedoch auch für die Stimuli innerhalb eines Blocks gelten, wie die Korrelationsanalysen mit den zielkonzept-spezifischen EAST-Effekten (z.B. Homosexuell vs. Heterosexuell) nahe legen. Dies kann allerdings nur als Hinweis gelten, da stimulusspezifische Effekte nicht notwendigerweise in (gleich gerichteter) Beziehung stehen müssen. In der getrennten Betrachtung von Effekten auf der Basis der Zielkonzeptkategorien (z.B. EAST\_Homo vs. EAST\_Hetero) liegt deshalb ein Vorteil des EAST gegenüber dem IAT (De Houwer, 2003a; Teige et al., 2004). Ein tatsächlicher Nutzen der Analyse zielkonzeptspezifischer EAST-Effekte ist jedoch Post-Hoc-Analysen zufolge für keine der betreffenden Studien dieser Arbeit zu verzeichnen. Die Beziehungsmuster zwischen den spezifischen EAST-Effekten und zum

Beispiel expliziten Einstellungen unterscheiden sich vernachlässigbar geringfügig von denen der allgemeinen EAST-Effekte.

Es bleiben neben der prozeduralen Variation von beispielsweise Stimulusabfolge oder Tastenbelegung und dem geringen Aggregationsniveau zwei weitere, von Teige et al. (2004) angeführte Ursachen für die unbefriedigende Reliabilität des EAST. Diese sind zum einen eine geringe interindividuelle Varianz der Versuchspersonen in den Einstellungen zu den betrachteten Einstellungsobjekten und zum anderen die mangelnde Repräsentativität der Stimuli. In den beiden durchgeführten Untersuchungen (Studie I und II) werden, gemessen an den Mittelwerten sogar recht hohe Varianzen der impliziten Maße erhalten. Mangelnde interindividuelle Einstellungsvarianz zeichnet sich zum Vergleich auch nicht in den Fragebogenmessungen ab. Wie einleitend bereits angedeutet wurde, ist die Repräsentativität des Stimulusmaterials für den EAST kritischer, als für den IAT (vgl. De Houwer, 2003a), weil die Erfassung kognitiver Assoziationsstärken mit dem IAT eher auf der Ebene von Kategorienlabel stattfindet (Karpinski & Hilton, 2001; Nosek, Banaji & Greenwald, 2002). Im IAT ist prinzipiell jeder Stimulus in eine von zwei gegensätzlichen Kategorien zu ordnen, z.B. Homosexuell vs. Heterosexuell, positiv vs. negativ, wobei die Kompatibilität der gemeinsam eine Antworttaste belegenden Kategorien variiert wird (z.B. Homosexuell oder positiv). Deshalb werden hier komplette Blöcke miteinander verrechnet. Der EAST dagegen unterscheidet sich diesbezüglich in 2 Punkten vom IAT. Erstens werden für den EAST-Effekt nicht alle Stimuli eines Blocks betrachtet, sondern nur die Zielkonzeptstimuli. Zweitens sind die Zielkonzeptstimuli im EAST nicht aufgrund ihrer konzeptuellen Bedeutung, sondern auf der Basis unterschiedlicher perzeptuelle Eigenschaften (blau vs. grün), die extrinsisch valenzbesetzt sind, zu beantworten. Inwiefern es einem Probanden z.B. schwer fällt, ein homosexuelles Paarbild auf der Basis der extrinsisch positiven Farbe richtig mit der „guten Taste“ zu beantworten, könnte demnach sowohl an der individuellen Inkompatibilität der Kombination Homosexualität und positiv liegen, als auch daran, dass die dargestellten Personen vom Probanden als unsympathisch oder unattraktiv empfunden werden. Deshalb ist neben der Repräsentativität auch das Matchen der verwendeten Stimuli hinsichtlich verschiedener Merkmale wie Alter, Attraktivität, Sympathie u.a. von entscheidender Bedeutung. Bei gravierender und konsistent ausgerichteter Unterschiedlichkeit des Stimulusmaterials in Bezug auf wertigkeitsrelevante Aspekte, wie z.B. die Attraktivität, spiegelt, ganz gleich ob IAT oder EAST, das Differenzenmaß aus kompatiblen und inkompatiblen Trials keine einstellungsrelevanten Assoziationen wider, sondern schlichtweg den optischen Habitus der Stimuli (De Houwer, 2003a). Damit ist zudem eine schwerwiegende Validitätseinschränkung verbunden (Rothermund & Wentura, 2001).

Die hier verwendeten Stimuli für die Kategorien Homosexualität und Heterosexualität sowie Übergewicht und Normalgewicht wurden in Vorstudien auf Prototypikalität, Attraktivität und Sympathie getestet und entsprechend ausbalanciert. Dies sollte eigentlich sichern, dass die Stimuli nicht nur reprä-



sentativ für die Kategorien sind, sondern im Mittel auch ausgeglichen hinsichtlich anderer wertigkeitsrelevanter Aspekte. Gegeben, dass nichtsdestotrotz die Stimuli Ursache der geringen Reliabilität des EAST in vorliegender Arbeit sind, ist der EAST als Messung interindividueller Differenzen in kognitiven Assoziationsstärken mehr als anzuzweifeln: Ein Effektmaß, das derart sensibel auf solche interindividuelle Unterschiede in der Wahrnehmung inhaltlich irrelevanter Merkmale der, nach allen Regeln der Kunst ausgewählten, Stimuli reagiert, gewährleistet keine Objektivität und muss als sehr fragwürdig eingeschätzt werden.

### 7.2.2 Stabilität

Gemäß der Definition impliziter Kognitionen durch Banaji & Greenwald (1995; vgl. Abb. 1-1) und den Modellen von Devine (1989) oder Wilson, Lindsey & Schooler (2000) ist von einer vergleichsweise *hohen* Stabilität impliziter Einstellungen auszugehen. Implizite Einstellungen seien, wenn überhaupt, dann nur über einen längeren Zeitraum zu verändern. Entsprechend dieser Konzeptionalisierung sollten implizite Einstellungsmessungen eine hohe zeitliche Stabilität aufweisen. Dies ist jedoch für die derzeit verfügbaren impliziten Messverfahren nicht gegeben. Für die am häufigsten eingesetzten Instrumente wie den IAT oder das Affektive Priming werden alles andere als eine hohe Test-Retest-Reliabilitäten berichtet. Cunningham, Preacher & Banaji (2001) setzen mit zweiwöchigem Abstand insgesamt viermal einen Standard-IAT, einen Antwortfenster-IAT und ein Affektive Primings mit Antwortfenster zur Messung von impliziten Einstellungen gegenüber Farbigen relativ zu Weißen ein. Sie erhalten einen mittleren manifesten Stabilitätskoeffizienten von lediglich .27, können jedoch mit latenter Strukturgleichungsmodellierung für die entsprechenden messfehlerbereinigten impliziten Variablen Stabilitäten von .68 (Standard-IAT), .60 (Antwortfenster-IAT) und .83 (Antwortfenster-Priming) nachweisen. Steffens & Buchner (2003) gehen nach dem gleichen Schema wie Cunningham, Preacher & Banaji (2001) vor berechnen Stabilitäten für einen IAT Homosexualität von .56 (einwöchiges Intervall) und .61 (zehnminütiges Intervall). Für den Selbstbericht werden im Vergleich dazu Stabilitäten von .87 und .92 dokumentiert. Auch für die Messung von Ängstlichkeit zeigen Schmukle & Egloff (2004), dass der IAT mit einer Test-Retest-Reliabilität von .50 deutlich unter der Stabilität des Selbstberichtes von .85 liegt. Die Autoren verweisen auf andere eigene Untersuchungen, deren zufolge sie eine Stabilität von .58 für ein einwöchiges Intervall, von .62 für einen Monat und von .47 für einen Zeitraum von einem Jahr für den IAT ermittelten (Egloff, Schwerdtfeger & Schmukle, 2005).

Die Befundlage zusammenfassend ist für den IAT eine im Mittel eine geringere Stabilität bei hoher interner Konsistenz zu vermerken (Hofmann et al., 2003). Es scheint also nahe liegend, von einer relativ großen Anfälligkeit des IAT für Situationsvariabilität auszugehen, was angesichts der Definition impli-

ziter Einstellungen bereits den Validitätsanspruch dieses Verfahrens unmissverständlich hinterfragen lässt. Da die Messung der impliziten und expliziten Einstellungen in Studie III nicht zu einem Zeitpunkt erfolgten, sondern IAT und EAST Übergewicht einen einwöchigen Abstand zum IAT und EAST Homosexualität sowie zum Selbstbericht (Übergewicht und Homosexualität) aufweisen, müsste die vorgenommene Konstruktvalidierung expliziter und impliziter Einstellungen zusätzlich zu Trait-, Methoden- und Messfehlervarianz die Effekte situativer Einflüsse berücksichtigen. Dies ist jedoch aufgrund des Erhebungsdesigns leider nicht möglich. Zudem weist der IAT Übergewicht im Vergleich zum IAT Homosexualität in Studie I und III eine relativ geringe interne Konsistenz auf. Erklärungsansätze für die verhältnismäßig geringe Reliabilität des IAT Übergewicht müssen methodische, prozedurale und inhaltliche Fragen aufwerfen. Der methodische Aspekt des verwendeten Stimulusmaterials für die beiden Zielkonzeptkategorien Übergewicht und Homosexualität wurde bereits im vorangegangenen Abschnitt angeführt. Da alle verwendeten Bildstimuli vorgetestet und hinsichtlich wertigkeitsrelevanter Merkmale gematcht wurden und zudem der IAT-Effekt relativ robust gegen die präsentierten Einzelstimuli ist (vgl. Steffens, Lichau & Still, 2004), wird diesem möglichen Erklärungsansatz weniger Gewicht beigemessen. Zwar konnten Materialeffekte auf der Ebene der Stimuli nachgewiesen werden (z.B. De Houwer, 2001), jedoch wurden dazu Extrembeispiele von Stimuluskombinationen verwendet, was den Bedingungen der hier berichteten Untersuchung keineswegs nahe kommt. Ein prozeduraler Unterschied zwischen IAT Übergewicht und IAT Homosexualität ergibt sich aus der relativen Unerfahrenheit der Versuchspersonen bezüglich des IAT und der Laborsituation. Zumindest zum ersten Messzeitpunkt ist die Stichprobe hinsichtlich der Erfahrungen mit Laborsituationen im allgemeinen und dem IAT als speziellen Messverfahren möglicherweise sehr heterogen. So haben einige Versuchspersonen eventuell noch nie zuvor einen IAT absolviert, während einige aufgrund der Teilnahme an anderen Versuchen bereits vertraut mit dem Verfahren sind. Um Lerneffekte als Ursache für die Reliabilitätsschwankungen untersuchen zu können, hätte die Reihenfolge der Einstellungsobjekte bei der Erhebung ausbalanciert werden müssen, was wiederum für die messfehlerberücksichtigende Analyse interindividueller Differenzen eine Erhöhung der Probandenzahl erforderlich gemacht hätte. Eine eher inhaltlich ausgerichtete Erklärungsmöglichkeit zur geringen Reliabilität des IAT Übergewicht im Vergleich zum IAT Homosexualität betrifft die Einstellungsobjekte selbst. Es ist durchaus plausibel, Unterschiede zwischen den Einstellungen zu Homosexualität und Übergewicht anzunehmen, die sich mit der unterschiedlichen Selbstidentifikation im Hinblick auf beide Stereotypen begründen lassen. Aus unterschiedlich stabilen Selbstidentifikationen hinsichtlich sexueller Orientierung vs. Körperumfang sollten entsprechend unterschiedlich konsistente Einstellungen resultieren (siehe dazu Abschnitt 7.4).

Es konnte gezeigt werden, dass der IAT – obwohl relativ unanfällig gegenüber bewusster Verfälschungsversuche (z.B. Banse, Seise & Zerbes; Egloff & Schmukle, 2002; Steffens, 2004) – sehr sensi-

bel für einfache kontextuelle Variationen ist, wie z.B. die Aktivierung stereotypenkonträrer Kognitionen oder das „Primen“ der Versuchspersonen mit positiv bewerteten farbigen Personen (z.B. Blair, Ma & Lenton, 2001; Karpinski & Hilton, 2001; Wittenbrink, Judd & Park, 2001). Nichtsdestotrotz sind auch stabile Verhaltensvorhersagen für IAT-Maße erbracht worden (Hugenberg & Bodenhausen, 2004), die zeigen, dass der IAT nicht allein situativ bedingte Assoziationen erfasst.

Zusammenfassend bleibt festzustellen, dass, sollte die Situationsanfälligkeit des IAT Ursache für die fehlende Korrespondenz zwischen impliziten und expliziten Einstellungen sein, kann der IAT als Messinstrument interindividueller Differenzen in impliziten Einstellungshaltungen nicht empfohlen werden. Entweder ist die Konzeptualisierung von impliziten Einstellungen als die stabileren Dispositionen im Vergleich zu expliziten Einstellungen nicht zulässig, oder die zur Messung dieser impliziten Einstellungen eingesetzten Verfahren sind nicht adäquat.

### 7.2.3 Aspekte der Reliabilitätsberechnung

Abschließend ist der Diskussion zur Reliabilitätsproblematik impliziter Einstellungsmaße noch ein eher statistischer Aspekt beizufügen. Die hier betrachteten impliziten Maße sind Differenzmaße, weshalb allgemein nochmals auf die Problematik von Differenzmaßen hinsichtlich der Reliabilität hinzuweisen ist. Das prinzipielle und hauptsächlich aus der Veränderungsmessung bekannte Problem von Differenzwerten basiert auf dem Effekt der Regression zur Mitte und resultiert insbesondere aus der hohen Messfehlerbehaftung von Extremwerten (Trautner, 1992). Zusätzlich ist für Veränderungsmessung das Reliabilitäts-Validitäts-Dilemma zu berücksichtigen (Collins, 1991; Rost, 2004): Stehen beide Ausgangswerte in enger Beziehung, ergibt sich für das Differenzmaß eine geringe Reliabilität aufgrund der Varianzeinschränkung, je niedriger dagegen der Zusammenhang zwischen beiden Ausgangswerten ausfällt, desto geringer ist ihre Validität. Insbesondere ersteres ist für vorliegenden Fall von Bedeutung, da die Mittelwerte der Reaktionszeiten und Fehlerraten von inkompatiblen und kompatiblen Blöcken (IAT) bzw. Stimuli (EAST) sehr hoch miteinander korrelieren. Im Mittel beträgt die Korrelation der Reaktionsgeschwindigkeit über die impliziten Messungen und die Einstellungsobjekte hinweg  $.67$  ( $p < .001$ ). Die mittlere Korrelation der Fehlerraten beträgt entsprechend  $.50$  ( $r < .001$ ). In diesen Korrelationen spiegeln sich die interindividuellen Unterschiede in der kognitiven Verarbeitungsgeschwindigkeit und Aspekte fluider Intelligenz wider.

Die Kritik an Differenzmaßen hinsichtlich des statistischen Regressionsartefakts halten Rogosa, Brandt & Zimowski (1982) jedoch angesichts der Bedeutung und der Vorteile von Differenzmaßen für übertrieben. Auch zeigt Zimmerman (1996), dass die allgemein diskutierte geringe Reliabilität von Differenzen hauptsächlich auf methodisch-theoretischen Annahmen beruht und nicht notwendiger-

weise statthaft ist. Darüber hinaus handelt es sich bei impliziten Maßen nicht um Veränderungsmessung im klassischen Sinne, vielmehr *ist* die Differenz zwischen inkompatibler und kompatibler Kombination das implizite Maß. Die Reliabilität des IAT-Effekts, der immerhin auf der Differenz der Mittelwerte (Reaktionszeiten, Fehlerraten) zweier voneinander getrennter Messblöcke basiert, ist zufriedenstellend.

Der EAST-Effekt dagegen stellt das Ausmaß der Inkompatibilität von Trials innerhalb *eines* Blockes dar, wobei hier vor allem Task-Switch- und Response-Switch-Effekte zu tragen kommen können (Schmitz, Teige, Voss & Klauer, 2005). Neben der theoretischen Fundamentalisierung des EAST ist nach wie vor noch nicht geklärt, inwiefern solche, für die eigentliche Messintention irrelevanten Effekte zum letztendlichen EAST-Maß beitragen und wie eine adäquate Reliabilitätsberechnung vorzunehmen ist (vgl. Hoffmann, Kiesel & Sebald, 2003; Klauer, Mierke & Musch, 2003). In diesem Zusammenhang soll darauf hingewiesen werden, dass im Rahmen eines kürzlich stattgefundenen EAST-Workshops (08. – 10. April, 2005, Oppurg) mögliche Alternativen für die Berechnung der internen Konsistenz des EAST diskutiert worden sind. Da stellenweise höhere Test-Retest-Reliabilitäten, als interne Konsistenzen gemäß Cronbach's Alpha oder Split-Half-Reliabilität gefunden wurden (Ellwart, Rinck & Becker, 2005), erscheint eine Auseinandersetzung mit der adäquaten Bestimmung der psychometrischen Eigenschaften von reaktionszeit- bzw. fehlerbasierten Effektmaßen notwendig.

Möglicherweise sind multinomiale Schätzmodelle zur Unterscheidung kontrollierter, automatisierter und zufälliger Rateprozesse im EAST die adäquatere Methode zur Berechnung eines reliablen und validen EAST-Effekts (Stahl & Degner, 2005). Diese Auswertungsmethodik ist im Hinblick auf die impliziten Einstellungsmaße relativ unerforscht und fokussiert derzeit noch ausschließlich auf Fehlerdaten. Für den IAT wurde ein multinomiales Modell, das vier verschiedene Prozesse berücksichtigt, bereits erfolgreich getestet (Conrey, Sherman, Gawronski, Hugenberg & Groom, 2004, in press).

### **7.3 Evaluation der offenen Assoziativmessung**

Mit der offenen Assoziativmessung ist im Rahmen dieser Arbeit ein Verfahren zur Erfassung expliziter und impliziter Einstellungen entwickelt worden, was einer ersten Evaluation zufolge als überaus viel versprechend bewertet wird. Die Messung beruht auf der Methodik des Thematischen Apperzeptionstests, bei der Probanden instruiert werden, eine eigene Geschichte zu vorgegebenen Bildmotiven zu verfassen. Hinsichtlich der Bildmotive sowie der Kodierschemata zur Erfassung der Geschichteninformationen wurde eine Anpassung der Methodik für die Messung von Einstellungen, Vorurteilen zu Übergewicht und Homosexualität sowie relevanten Motiven vorgenommen. Für eine Beurteilung des Messverfahrens zur Erfassung interindividueller Einstellungsdifferenzen sollten psychometrische Gesichtspunkte der Kodierung, Validitätsbefunde hinsichtlich expliziter und impliziter Einstellungs-

und Motivationsanteile und Aspekte der Handhabung, wie z.B. die Ökonomie, berücksichtigt werden. Hinsichtlich der psychometrischen Eigenschaften zeichnen sich die Kodierschemata der Assoziativmessung durch eine sehr hohe Interraterreliabilität der entsprechenden Variablen aus. Das bedeutet, die fokussierten Geschichtenmerkmale wie Valenz oder Vorurteilsgehalt sind relativ objektiv erfassbar. Darüber hinaus werden zwischen den Geschichtenvariablen und den selbstberichteten Einstellungen zu Übergewicht bzw. Homosexualität stabile und sinnvolle Beziehungen registriert, trotz dass es sich um methodisch deutlich voneinander abweichende Verfahren handelt. Hier sind jedoch zunächst allgemeine Unterschiede in den Zusammenhangsmustern der Geschichtenmessung und der expliziten Einstellungen in Abhängigkeit vom Einstellungsobjekt anzunehmen. In Studie III sind insbesondere enge Beziehungen zwischen den Geschichtenvariablen und des selbstberichteten Einstellungen für Übergewicht zu berichten, während für das Einstellungsobjekt Homosexualität weniger signifikante Zusammenhänge zwischen beiden Messmethoden bestehen. Dennoch weisen die Korrelationen zwischen offener Assoziativmessung und dem Selbstbericht ungeachtet des Einstellungsobjektes auf einen substanziellen Zusammenhang hin.

Ein relativ konsistenter Befund in diesem Rahmen ist die negative Korrelation zwischen der Geschichtenlänge und den selbstberichteten Einstellungen, die in beiden hierfür relevanten Studien dieser Arbeit aufgezeigt werden konnten. Für beide Einstellungsobjekte gehen negativere Selbstberichte mit kürzeren Geschichten einher. Dieser Zusammenhang könnte zum einen interindividuelle Unterschiede in den Einstellungen zu Übergewicht und Homosexualität reflektieren, möglicherweise zeichnen sich hierin jedoch eher motivationale oder fähigkeitsbezogene Personenmerkmale ab, wie etwaige Selbstpräsentationsmotive oder (verbale) Intelligenz. Zum einen könnte sich eine motivierte Bereitschaft zur Auseinandersetzung mit den Einstellungsobjekten in einem positiveren Selbstbericht und in umfangreicheren Geschichten äußern. Zum anderen ist anzunehmen, dass sich (verbal) intelligente Personen durch ein besseres Verständnis/ eine höhere Kontrollierbarkeit von Fragebogen-Items und eine vergleichsweise komplexere und vielseitigere Ausdrucksweise auszeichnen. Entsprechend zeichnet sich Post-Hoc-Analysen<sup>5</sup> zufolge insbesondere für die Motivation zu Vorurteilsfreiheit eine Bedeutung für den Zusammenhang zwischen Wortanzahl und expliziten Einstellungen. Die statistische Kontrolle von sozialer Erwünschtheit oder der verbalen Intelligenz haben einen vergleichbar unbedeutenden Einfluss in diesem Zusammenhang. Der ebenfalls stabile positive Zusammenhang zwischen dem Vorurteilsgehalt der Geschichten und den expliziten Einstellungen (Studie III) konnte dagegen als relativ unabhängig von sozialer Erwünschtheit bzw. der Motivation zur Vorurteilsfreiheit herausgestellt werden. Sollten weitere Untersuchungen eine stabile Replizierbarkeit dieser Beziehung

---

<sup>5</sup> Aufgrund der Verfügbarkeit entsprechender Kontrollvariablen wie Intelligenz, soziale Erwünschtheit und Motivation zur Vorurteilsfreiheit wurde an dieser Stelle lediglich eine Post-Hoc-Analyse für die Daten aus Studie III durchgeführt. Dabei muss jedoch berücksichtigt werden, dass hier keine signifikanten Zusammenhänge zwischen Geschichtenlänge und expliziten Einstellungen für das Einstellungsobjekt Homosexualität bestehen. Eine Verallgemeinerung der Befunde ist also nicht möglich.

auch unter Kontrolle weiterer möglicher Drittvariablen erbringen, kann eine zugrunde liegende Einstellungsdisposition als das gemeinsame Substrat des Zusammenhangs zwischen Geschichtenmessung und Selbstbericht angenommen werden.

Die Validität der offenen Assoziativmessung hinsichtlich *impliziter* Einstellungen ist weniger eindeutig zu bewerten. Zwischen den Geschichtenvariablen und den impliziten Einstellungsmaßen bestehen zunächst augenfällig weniger substanzielle Zusammenhänge, als für den Selbstbericht. Dies kann einerseits als Hinweis auf die relative Unabhängigkeit von Geschichtenmessung und impliziten Einstellungen betrachtet werden, andererseits sind jedoch eine Reihe weiterer möglicher Einflussfaktoren zu bedenken. In dieser Arbeit sind es, wie bereits erwähnt, insbesondere die Übergewichtsgeschichten, die in verhältnismäßig engem Zusammenhang zu expliziten Einstellungen zu Übergewicht stehen. Da jedoch die impliziten Messungen zu Übergewicht relativ problematisch hinsichtlich ihrer psychometrischen Eigenschaften (geringe Reliabilität) bzw. ihrer Interpretierbarkeit (situative Variabilität vs. problematische Einstellungsinhalt) sind, ist die Divergenz von Geschichtenmessung und impliziten Einstellungen zu Übergewicht nur begrenzt inhaltlich interpretierbar.

Für das Einstellungsobjekt Homosexualität spielt vor allem die Variable der Stereotyperkennung bzw. -thematisierung eine entscheidende Rolle im Zusammenhang zu impliziten Einstellungen. Die entsprechenden Korrelationen zu IAT und EAST weisen in die Richtung höherer impliziter Ablehnung von Homosexualität bei Personen, die den Stereotyp Homosexualität in den Geschichten nicht erkannt bzw. nicht deutlich thematisiert haben. Bei der betreffenden Geschichtenvariablen ist fraglich, ob es sich um ein Nichterkennen oder ein Nichtansprechen im Sinne von Vermeidung handelt, aufgrund der signifikant negativen Zusammenhänge zu beiden impliziten Maßen liegt jedoch eine mehr oder weniger bewusste Vermeidung der Homosexualitätsthematik näher.

An dieser Stelle ist als kritischer Aspekt die Abhängigkeit der offenen Assoziativmessung von der Darstellbarkeit des Einstellungsobjekts in einem schematisch-mehrdeutigem Bildmotiv anzuführen. Bei grafisch schwer fassbaren Einstellungsobjekten besteht die Gefahr einer Fehlinterpretation des Bildmotivs und somit einer inhaltlich wertlosen Geschichtenmessung. Eine missglückte Stereotypaktivierung könnte demnach prinzipiell Ursache für fehlende Zusammenhänge zwischen Geschichtenmessung und anderen Einstellungsmaßen sein. Insofern würde die relative Unmissverständlichkeit von Bildmotiven zum Einstellungsobjekt Übergewicht eine maßgebliche Voraussetzung für die enge Beziehung zwischen Übergewichtsgeschichten und expliziten Einstellungen zu Übergewicht sein. Dagegen könnte für die Homosexualitätsmotive eine Identifikationsproblematik hinsichtlich des thematisierten Einstellungsobjektes vorliegen. Da die Bildmotive neben der Aktivierung des Einstellungsobjektes einen ausreichenden Interpretationsspielraum bieten sollten, bleibt es für die Homosexualitätsmotive der individuellen Auslegung überlassen, zwei enge Freunde, Brüder oder gar Vater und Sohn zu sehen. Hier zeichnet sich ein regelrechtes Dilemma zielsicherer Stereotypaktivierung bei maximaler

Mehrdeutigkeit der Darstellung ab. Eine projektive Methode zur Erfassung bewusster und unbewusster, reflektierter und unreflektierter, überlegter und spontaner Kognitionsinhalte muss jedoch eine entsprechende Projektionsfläche individueller Interpretationen bieten. Die absolut unmissverständliche Darstellung von Homosexualität, z.B. in Form sich küssender Männer, müsste eine starke inhaltliche Einschränkung der Geschichten und ergo eine Varianzeinschränkung der Geschichtenvariablen zur Folge haben. Entsprechende Geschichten könnten inhaltlich wenig mehr widerspiegeln als eine Kusszene. Da in dieser Arbeit jedoch interindividuelle Differenzen im Fokus der Betrachtung liegen, sollte die offene Assoziativmessung sensitiv für interindividuelle Unterschiede in der inhaltlichen Auslegung der Bildmotive sein und entsprechende Variabilität aufweisen. Für das Einstellungsobjekt Homosexualität könnte die um einiges variabelere Interpretierbarkeit der Bildmotive entsprechend Ursache für die geringeren Korrespondenzen zwischen Homosexualitätsgeschichten und expliziten Einstellungen zu Homosexualität sein.

Hinsichtlich Reliabilität und Validität der offenen Assoziativmessung ist soweit zu resümieren, dass die Methodik des Thematischen Apperzeptionstest grundsätzlich erfolgreich auf die Messung von Einstellungen, Vorurteilen und vorurteilsrelevanten Motivationen adaptiert werden konnte. Zwischen dem Vorurteilsgehalt der Geschichten und den expliziten Vorurteilen bestehen substantielle positive Beziehungen, wobei hier bereits erklärbare Unterschiede zwischen beiden Einstellungsobjekten zu verzeichnen sind. Diese werden bei der Betrachtung der Zusammenhänge zwischen den Geschichtenvariablen und den impliziten Assoziationsmaßen deutlicher und können u.a. auf Unterschiede im Grad der bewussten Themenaktivierung und -verarbeitung durch die Bildmotive beider Einstellungsobjekte zurückgeführt werden. Insofern reflektiert die offene Assoziativmessung die Besonderheiten von Einstellungen zu Übergewicht bzw. Homosexualität. Während übergewichtige Personen problemlos aufgrund augenfälliger Merkmale als übergewichtig zu identifizieren sind, müssen sich homosexuelle Menschen regelrecht „outen“, um von der Gesellschaft „erkannt“ zu werden. Dies macht eine Vermeidung der Homosexualitätsthematik für Außenstehende prinzipiell leichter.

Abschließend ist zur Evaluation der offenen Assoziativmessung der Aspekt der Verfahrensökonomie zu berücksichtigen. Zunächst ist die einzurechnende Bearbeitungsdauer der Geschichtenmessung im Vergleich zu Fragebogenmessungen und impliziten Messverfahren deutlich höher, wenn keine Zeitvorgaben festgelegt werden. Entscheidender ist jedoch die qualitative Datenbasis der Assoziativmessung. Für die Untersuchung empirischer Fragestellungen, sind quantitative Inhaltskodierungen des Geschichtenmaterials vorzunehmen, was zur Erhöhung der Objektivität und entsprechend auch der Validität von mindestens zwei unabhängigen Ratern vorgenommen werden sollte. In dieser Hinsicht ist die Auswertung der offenen Assoziativmessung deutlich aufwendiger, als für Fragebogenskalen und computergestützte Experimente. Sollte sich der Zusammenhang zwischen Geschichtenlänge und Einstellungen als stabil replizierbar erweisen, wäre jedoch eine objektive und zudem problemlos zu

kodierende Geschichtenvariable zur Untersuchung interindividueller Einstellungsdifferenzen verfügbar.

Die Aspekte, deren zufolge die offene Assoziativmessung als vergleichsweise weniger ökonomisch zu bewerten ist, bedingen gleichzeitig das erhebliche Potential des Verfahrens. Aufgrund des offenen Antwortformats und der qualitativen Datenbasis ist die Geschichtenmessung hinsichtlich Informationsgehalt und Auswertungsmöglichkeiten zweifellos vielfältiger, als jede Fragebogenmessung oder experimentelle Reaktionszeit- und Fehlerdaten. Die Möglichkeiten des Verfahrens sind im Rahmen vorliegender Arbeit notwendigerweise nur begrenzt ausgeschöpft worden. Speziell in Hinblick auf die Messung vorurteilsbezogener *impliziter* Motive zusätzlich zu selbstberichteten *expliziten* Motiven (z.B. Banse & Gawronski, 2003; Plant & Devine, 1998) erscheint eine weiterführende Untersuchung der offenen Assoziativmessung viel versprechend.

#### **7.4 Die Rolle der ausgewählten Einstellungsobjekte**

Für die Untersuchung der Konstruktvalidität von expliziten und impliziten Einstellungen wurden zwei soziale Minderheiten unserer Gesellschaft als Einstellungsobjekte ausgewählt, und zwar übergewichtige und homosexuelle Menschen. Diese beiden Personengruppen sind in vielen Kulturen nachweislich betroffen von Stereotypen, Vorurteilen und Diskriminierung seitens der sozialen Mehrheit (z.B. Crandall & Biernat, 1990; Crandall & Martinez, 1996; Crandall et al., 2001; Grover, Keel & Mitchell, 2003; Rothblum et al., 1990; Hayes, 1995; Hebl, et al., 2002; Herek, 1994; LaMar & Kite, 1998; Steffens & Wagner, 2004).

Ein diesbezüglich kritischer Aspekt aktueller Arbeit bezieht sich auf die Geschlechtsspezifität des Materials der Messinstrumente. Die Stimuli der impliziten Einstellungsmessungen stellen zum einen übergewichtige und normalgewichtige Frauen dar und zum anderen rein männliche und gemischtgeschlechtliche Paare. Während jedoch für das Einstellungsobjekt Homosexualität aus Repräsentativitätsgründen in allen Messungen männliche Homosexuelle fokussiert wurden, liegt eine, das Material betreffende Geschlechtsspezifität für das Einstellungsobjekt Übergewicht nur im Hinblick auf die impliziten Maße vor. Im Selbstbericht ist allgemein von „übergewichtigen Menschen“ bzw. „Dicken“ die Rede. Diese forcierte Selektivität des Stimulusmaterials diene der Vermeidung von ambivalenten Assoziationen zu männlichen und weiblichen Individuen, da sowohl übergewichtige *Frauen*, als auch homosexuelle *Männer* die jeweiligen Prototypen der ausgewählten sozialen Stereotype Homosexualität und Übergewicht darstellen (vgl. Abschnitt 3.2) und insbesondere die impliziten Maße auf Ambivalenzen im Stimulusmaterial sensibel reagieren sollten (z.B. de Liver, Wigboldus & van der Pligt, 2005). Dies ist umso bedeutsamer, derweil das Geschlecht an sich einen sozialen Stereotyp darstellt, zu dem divergente Einstellungen bzw. Assoziationen berichtet werden (z.B. Carpenter & Banaji, 2000; Eagly



& Mladinic, 1989; Sidanius, Pratto & Bobo, 1994; Skowronski & Lawrence, 2001). Eine Alternative zur weitestgehenden Vergleichbarkeit des verwendeten Materials hinsichtlich des Geschlechts wäre die ausdrückliche Fokussierung auf weibliche Übergewichtige im Selbstbericht, wobei dadurch Konfundierung des Stereotyps Übergewicht mit dem Stereotyp Frauen unauflösbar gewesen wäre. Fraglich ist, ob die Fragebogen-Items zu „übergewichtigen Menschen“ und das rein weibliche Stimulusmaterial für IAT- und EAST-Übergewicht derartige inhaltliche Abweichungen in den Einstellungen zur Folge haben können, die eine Erklärung für die mangelnden Korrespondenz zwischen expliziten und impliziten Einstellungen zu Übergewicht liefern. Dieser Frage kann über systematische empirische Untersuchungen nachgegangen werden.

Darüber hinaus ist im Rahmen vorliegender Arbeit nicht hinreichend zu klären, inwieweit Einstellungen gegenüber Übergewicht und gegenüber Homosexualität in ihrer *Konzeptualisierung* miteinander vergleichbar sind bzw. inwiefern fundamentale Unterschiede zwischen beiden Einstellungsrepräsentationen anzunehmen sind, weshalb eine gemeinsame Konstruktvalidierung nicht möglich wäre. Prinzipiell sollte eine Generalisierbarkeit der Konzeptionen über Einstellungsinhalte hinweg bestehen. Spezifischen Eigenschaften der Einstellungsobjekte selbst sind jedoch als Einflussfaktoren für die Beziehung zwischen expliziten und impliziten Einstellungen ausgewiesen worden (z.B. Gschwendner & Schmitt, 2005; Nosek, 2004), sie verdienen deshalb gesonderte Beachtung. Eine entscheidende Variable in vorliegender Arbeit ist entsprechend die *Selbstrelevanz* der Einstellungsobjekte für die Untersuchungsteilnehmer. Darüber hinaus sind an dieser Stelle Einstellungsunterschiede in Abhängigkeit von Personenmerkmalen, wie z.B. dem Geschlecht, zu berücksichtigen. So ist aufgrund der Körperschemamessung in Sinne von körperlichem Real- versus Idealbild die durchschnittliche Präferenz eines relativ schlanken Körpers für die Gesamtstichprobe in Studie III festzustellen, wobei Frauen eine höhere Abweichung zwischen Realität und Wunschvorstellung aufweisen als Männer. Diese allgemeine Tendenz verwundert ebenso wenig wie der spezifische Geschlechtsunterschied, da vor allem für Menschen aus westlichen Industriegesellschaften eine enorme Diskrepanz zwischen dem realen Gewichtsmittelwert der (erwachsenen) Bevölkerung und der medial verbreiteten Idealfigur besteht. Angesichts der konsensualen Präferenz eines extrem schlanken, ja häufig sogar abmagerten Körpers, sind Männer wie Frauen, jüngere wie ältere Menschen gleichsam mit einem zum Teil völlig übersteigerten Anspruch konfrontiert, möglichst dünn zu sein. Entsprechend berücksichtigen Ätiologiemodelle von Essstörungen wie *Anorexia Nervosa* oder *Bulimia Nervosa* den Einfluss des gesellschaftlichen Schlankheitsideals (vgl. Fichter & Quadflieg, 2000; Hänsel, 1995). Es liegt nahe, das Vorherrschen eines mehr oder weniger latenten „Dick-Seins-Gefühls“ bei vielen Menschen anzunehmen, auch wenn laut Body-Mass-Index keine Normabweichung in Richtung Übergewicht oder sogar bereits eine Normabweichung in Richtung Untergewicht vorliegt.

Eine solche irrealer Selbstidentifikation kann für Homosexualität dagegen nicht angenommen werden, was u.a. sicherlich darauf zurückzuführen ist, dass es sich bei der sexuellen Orientierung nicht um ein kontinuierlich verteiltes Merkmal wie das Körpergewicht bzw. der BMI handelt. Eine erwachsene Person kann im Normalfall vermutlich problemlos bestimmen, ob sie heterosexuell, homosexuell oder bisexuell orientiert ist, wobei die Kategorie „heterosexuell“ für die Mehrheit unserer Gesellschaft als Norm gilt, während homo- und bisexuell gleichermaßen deviant sind. Dagegen fällt es nicht schwer anzunehmen, dass viele Personen der Allgemeinbevölkerung der Meinung sind, hier und da nicht ganz dem figürlichen Wunschbild zu entsprechen. Darüber hinaus ist die Wahrscheinlichkeit eines sexuellen Orientierungswechsels relativ gering im Vergleich zu der Möglichkeit, Gewicht zuzunehmen und vielleicht selbst einmal als übergewichtig zu gelten, noch dazu, weil der BMI positiv mit dem Alter korreliert ist, was auch die Verschiebung der oberen und unteren Normschranken für den Normalgewichtsbereich um jeweils einen Punkt etwa alle 10 Jahre nach dem 25. Lebensjahr widerspiegelt. Empirische Hinweise auf die hier vermuteten unterschiedlichen Selbstidentifikationsphänomene der Merkmale Körpergewicht und sexuelle Orientierung liefert der Vergleich von Untersuchungen, denen zufolge homosexuelle Menschen eine Ingroup-Favorisierung aufweisen (z.B. Banse, Seise & Zerbes, 2001; Seise, Banse & Neyer, 2002), wogegen übergewichtige ebenso wie normalgewichtige Menschen negativere Einstellungen zu Übergewicht haben (z.B. Crandall, 1994).

Es kann also vermutet werden, dass Übergewicht für die Mehrzahl der Probanden von Studie III eine deutlich höhere Selbstrelevanz besitzt, als Homosexualität, noch dazu weil der Frauenanteil der Stichprobe relativ hoch ist, und insbesondere für die Frauen die Figur eine große Rolle spielt. Dies sollte berücksichtigt werden, da qualitative Unterschiede zwischen den Einstellungsobjekten Auswirkungen auf die Varianz der resultierenden Einstellungsmaße haben kann, wodurch die Vergleichbarkeit der beiden Einstellungsinhalte Übergewicht und Homosexualität gemindert werden kann. Um die Eigenschaften der Einstellungsobjekte näher bestimmen zu können, ist eine Untersuchung der kognitiven Repräsentationen der Einstellungsinhalte notwendig. Sollten jedoch die spezifischen Einstellungscharakteristika Hauptursache der divergenten Beziehungsmuster für Einstellungen zu Übergewicht im Vergleich zu Homosexualität in Studie III sein, wäre von einer allgemeingültigen, inhaltsunabhängigen Konzeptionalisierung von Einstellungen abzusehen. Nicht dokumentierten Zusammenhangsanalysen zufolge, resultieren aus einer statistischen Kontrolle des Geschlechts jedoch keine nennenswerten Veränderungen der Beziehungsmuster in Studie III.

## 7.5 Zusammenfassung und Ausblick

Mit vorliegender Arbeit wurden die Konzeptionalisierung und die Bedeutung expliziter und impliziter Einstellungen untersucht, woraus eine Reihe möglicher Anknüpfungspunkte für weiterführende Fragestellungen abgeleitet werden konnte. Es wird deutlich, dass die Einstellungsforschung mit der Unterscheidung zwischen bewusst verarbeiteten, kontrollierbaren und automatisierten, unkontrollierbaren Kognitionen einen komplexen Problemraum eröffnete, der nur durch gezielte Untersuchungen einzuengen ist. Der empirische Teil dieser Arbeit liefert Hinweise auf eine anzunehmende Divergenz expliziter und impliziter Einstellungen, indem gezeigt werden konnte, dass der Einfluss der Messmethode die inhaltliche Bedeutung überwiegt. Zudem wurden für explizite und implizite Einstellungen unterschiedliche Beziehungsmuster zu einstellungsrelevanten Konstrukten aufgedeckt. Der Rahmen für eine Generalisierung der Ergebnisse ist jedoch aufgrund verschiedener inhaltlicher und methodischer Aspekte begrenzt. Zum einen wurden in vorliegender Arbeit zwei spezifische Einstellungsobjekte fokussiert, die Relevanz für Vorurteile und soziale Diskriminierung haben. Notwendigerweise bleiben dadurch weitere Einstellungsinhalte unberücksichtigt; darüber hinaus müssen charakteristische inhaltliche Aspekte der ausgewählten Einstellungsobjekte Übergewicht und Homosexualität in Betracht gezogen werden (vgl. Abschnitt 7.4)

Die prozeduralen Aspekte in Studie III erschweren die Vergleichbarkeit der Messungen, zumindest für das Einstellungsobjekt Übergewicht, wobei es sehr fraglich erscheint, inwiefern die Erhebung expliziter und impliziter Maße für zwei oder gar mehr Einstellungsobjekte zu einem Messzeitpunkt realistisch ist bzw. andere unerwünschte Effekte wie Ermüdung oder Überforderung verursacht. Darüber hinaus muss bei der Verallgemeinerung der Befunde die Materialspezifik der impliziten Messungen kritisch berücksichtigt werden (vgl. Abschnitt 7.4).

Die psychometrischen Unzulänglichkeiten des EAST sind kein alleiniges Problem vorliegender Arbeit, wie die wenigen EAST-Publikationen in wissenschaftlichen Fachzeitschriften und die unveröffentlichten Ergebnissen der Teilnehmer am EAST-Workshop (08.-10. April, Oppurg) erkennen lassen. Den geringen Reliabilitäten des EAST zufolge liegen ebenso Einschränkung in der Generalisierbarkeit der Ergebnisse von Studie I und III vor (vgl. Abschnitt 7.2). Dies kann allerdings als eigenständiger Befund gewertet werden, dementsprechend es fraglich erscheint, inwieweit implizite (Einstellungs-) Messung hinreichend Gültigkeit besitzt.

## 8 Literaturverzeichnis

- Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*.
- Aidman, E. V. & Carroll, S. M. (2003). Implicit individual differences: relationships between implicit self-esteem, gender identity, and gender attitudes. *European Journal of Personality*, 17, 19-37.
- Allport, G. W. (1935). Attitudes. In C. Murchison, *A Handbook of Social Psychology* (798-844). Worcester, MA, US: Clark University Press.
- Altemeyer, B. (1981). *Right-Wing Authoritarianism*. Winnipeg: University of Manitoba Press.
- Altemeyer, B. (1988). *Enemies of freedom*. San Francisco: Jossey-Bass.
- Altemeyer, B. (1994). Reducing prejudice in right-wing authoritarians. In M. P. O.Zanna (Ed.), *The psychology of prejudice: The Ontario symposium, Vol. 7* (131-148).
- Altemeyer, B. (1996). *The authoritarian specter*. Cambridge: Harvard University Press.
- Altemeyer, B. (1998). The other "authoritarian personality". *Advances in Experimental Social Psychology*, 30, 47-92.
- Altemeyer and B. Hunsberger (1992). Authoritarianism, Religious Fundamentalism, Quest, and Prejudice, *International Journal for the Psychology of Religion* 2, 113-133.
- Amodio, D. M. H., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality & Social Psychology*, 84, 738-753.
- Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality & Social Psychology*, 84, 738-753.
- Anooshian, L. J. & Seibert, P. S. (1996). Conscious and unconscious retrieval in picture recognition: A framework for exploring gender differences. *Journal of Personality & Social Psychology*, 70(3), 637-645.
- Aronson, E., Wilson, T.D. & Akert, R.M. (2004). Sozialpsychologie und das Gesetz. In *Sozialpsychologie* (4. Aufl., S. 584-615). München: Pearson Studium.
- Asendorpf, J. (2004). *Psychologie der Persönlichkeit* (3. überarb. und aktual. Aufl.). Heidelberg: Springer-Verlag.
- 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 & Social Psychology*, 83, 380-393.
- Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality & Social Psychology*, 81, 789-799.
- Atkinson, J. W. & Feather, N. T. (1966): A theory of achievement motivation. New York: Wiley.
- Bagozzi, R.P. & Yi, Y. (1990). Assessing method variance in multitrait-multimethod matrices, the case of self-reported affect and perceptions at work. *Journal of Applied Psychology*, 75 (5), 547-560.
- Balk, D. E., Lampe, S., Sharpe, B., Schwinn, S., Holen, K., & Cook, L. D. R. I. (1998). TAT results in a longitudinal study of bereaved college students. *Death Studies*, 22, 3-21.
- Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion*, 15(4), 501-520.
- Banse, R. & Fischer, L. (2002). Implicit and explicit aggressiveness and the prediction of aggressive behavior. *Poster präsentiert auf der 11. Konferenz der European Association of Personality Psychology*. Jena.
- Banse, R. & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Psychometrische Eigenschaften und Validität. *Diagnostica*, 49, 4-13.

- Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift fuer Experimentelle Psychologie*, *48*, 145-160.
- Bargh, J. A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of Personality & Social Psychology*, *43*(3), 425-436.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh: *Unintended thought* (3-51). New York, NY, US: Guilford Press.
- Bargh, J. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer. & T. K. Srull: *Handbook of social cognition, Vol. 1: Basic processes, Vol. 2: Applications* (2nd ed. 1-40). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer: *Automaticity of everyday life: Advances in social cognition*, Vol. 10. Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Bargh, J. A., Chaiken, S., Gøvender, R. & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality & Social Psychology*, *62*(6), 893-912.
- Baron, R. A. & Byrne, D. (2003). *Social Psychology* (10 ed.). Boston MA: Allyn & Bacon.
- Beck, A. T (1976). *Cognitive therapy and the emotional disorders*. Oxford, England: International Universities Press.
- Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88* (3), 588-606.
- Bessenoff, G. R. & Sherman, J. W. (2000). Automatic and controlled components of prejudice toward fat people: Evaluation versus stereotype activation. *Social Cognition*, *18*, 329-353.
- Bierhoff, H.W. (1998). *Sozialpsychologie: Ein Lehrbuch* (2. Aufl.). Stuttgart: Kohlhammer.
- Blair, I. V. (2001). Implicit stereotypes and prejudice. In G.B.Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (359-374).
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality & Social Psychology*, *81*, 828-841.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality & Social Psychology Review*, *6*, 242-261.
- Block, J. & Block, J. (1951). An investigation of the relationship between intolerance of ambiguity and ethnocentrism. *Journal of Personality*, *19*, 303-311.
- Boniecki, K. A., Lacks, J., & Zuwerink, J. R. (2002). The elusive relationship between measures of implicit and explicit prejudice. *Representative Research in Social Psychology*, *26*, 1-14.
- Bortz, J. (2004). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.) Heidelberg: Springer.
- Boswell, D. A. (1980). Evaluation of transparencies for psychology instruction. *Teaching of Psychology*, *7*(3), 171-173.
- 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 & Social Psychology*, *79*, 631-643.
- Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology*, *4*, 79-101.
- 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 & Social Psychology*, *81*, 760-773.
- Brewer, M. B. (1988). A dual process model of impression formation. In, T. K. Srull & R.S. Wyer: *A dual process model of impression formation* (1-36). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Brockhaus Enzyklopädie (2005). Mannheim: Bibliographisches Institut & F. A. Brockhaus AG.
- Buchner, A. & Brandt, M. (2003). Further evidence for systematic reliability differences between explicit and implicit memory tests. *Quarterly Journal of Experimental Psychology: Section A*, *56*, 193-209.

- Buchner, A. & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology*, 40(3), 227-259.
- Budner, S. (1962). Intoleranz of ambiguity as a personality variable. *Journal of Personality*, 30, 29-50.
- Busemeyer, J. R. & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 88, 237-244.
- Byrne, B.M. (2001). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Busemeyer, J.R. & Jones, L.E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 93, 549-562.
- Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality & Social Psychology*, 42, 116-131.
- Campbell, J.D. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Carpenter, S. & Banaji, M. R. (2000). Implicit gender attitudes : Group membership, cultural construal, and malleability. unpublished manuscript, Yale University, New Haven.
- Chambliss, H. O., Finley, C. E. & Blair, S. N. (2004). Attitudes toward Obese Individuals among Exercise Science Students. *Medicine & Science in Sports & Exercise*, 36 (3), 468-474
- Cloetta, B. (1983). Der Fragebogen zur Erfassung von Machiavellismus und Konservatismus. *Schweizerische Zeitschrift für Psychologie*, 42 (2/3), 127-159.
- Collins, L. M. (1991). The measurement of dynamic latent variables in longitudinal aging research: Quantifying adult development. *Experimental Aging Research*, 17(1), 13-20.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg K. & Groom, C. (2004, in press). Beyond automaticity and control: The Quad-Model of behavioral responses. Manuscript submitted for publication.
- Corbin, W. R., Bernat, J. A. & Calhoun, K. S. (2001). The role of alcohol expectancies and alcohol consumption among sexually victimized and nonvictimized college women. *Journal of Interpersonal Violence*, 16(4), 297-311.
- Cramer, P. (2001). Identification and Its Relation To Identity Development. *Journal of Personality*, 69, 667-688.
- Cramer, P. (2002). Defense Mechanisms, Behavior, and Affect in Young Adulthood. *Journal of Personality*, 70, 103-126.
- Crandall, C. S. (1994). Prejudice against fat people: Ideology and self-interest. *Journal of Personality & Social Psychology*, 66, 882-894.
- Crandall, C. S. & Biernat, M. (1990). The ideology of anti-fat attitudes. *Journal of Applied Social Psychology*, 20, 227-243.
- Crandall, C. S., D'Anello, S., Sakalli, N., Lazarus, E., Wiczorkowska, G., & Feather, N. T. (2001). An Attribution-Value model of prejudice: Anti-fat attitudes in six nations. *Personality & Social Psychology Bulletin*, 27, 30-37.
- Crandall, C.S., Eshleman, A. & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalisation. *Journal of Personality and Social Psychology*, 82 (3), 359-378.
- Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin*, 87, 546-563.
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit Attitude Measures: Consistency, Stability, and Convergent Validity. *Psychological Science*, 12, 163-170.

- Czopp, A. M., Monteith, M. J., Zimmerman, R. S., & Lynam, D. R. (2004). Implicit Attitudes as Potential Protection From Risky Sex: Predicting Condom Use With the IAT. *Basic & Applied Social Psychology, 26*, 227-236.
- Dalbert, C. (1999). The world is more just for me than generally: About the personal belief in a just world scale's validity. *Social Justice Research, 12*(2), 79-98.
- 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 & Social Psychology, 81*, 800-814.
- Deci, E. L. & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*, 109-134.
- Deci, E. L. & Ryan, R. M. (2000). The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry, 11*, 227-268.
- DeCharms, R., Morrison, H. W., Reitman, W. R. & McClelland, D.C. (1955). Behavioral correlates of directly and indirectly measured achievement motivation. In D. C. McClelland (Ed.), *Studies in motivation* (414-423). New York: Appleton-Century-Crofts.
- De Houwer, J. (in press). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage Publishers.
- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443-451.
- De Houwer, J. (2002). The implicit association test as a tool for studying dysfunctional associations in psychopathology. *Behavior Therapy and Experimental Psychiatry, 53*, 115-133.
- De Houwer, J. (2003a). The extrinsic affective Simon task. *Experimental Psychology, 50*, 77-85.
- De Houwer, J. (2003b). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (219-244). Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers.
- De Houwer, J. (in press). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage Publishers.
- De Houwer, J., Crombez, G., Koster, E. H. W., & De Beul, N. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *Journal of Behavior Therapy & Experimental Psychiatry, 35*, 275-286.
- De Houwer, J. & De Bruycker, E. (2005). Do EAST scores provide a measure of interindividual differences in attitudes? *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- De Houwer, J. & Eelen, P. (1998). An Affective Variant of the Simon Paradigm. *Cognition & Emotion, 12*, 45-62.
- De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition & Emotion, 16*, 643-666.
- de Jong, W. (1980). The stigma of obesity: The consequences of naive assumptions concerning the causes of physical deviance. *Journal of Health & Social Behavior, 21*, 75-87.
- de Jong, P. J. (2002). Implicit self-esteem and social anxiety: differential self-favouring effects in high and low anxious individuals. *Behaviour Research & Therapy, 40*(5), 501-508.
- de Jong, P. J. & Huijding, J. (2005). Pictorial vs. verbal stimuli in the EAST: Sensitivity to generally affective and phobia-relevant stimuli in high and low spider fearful individual. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- de Jong, P. J., Pasma, W. & Kindt, M. (2001). A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research & Therapy, 39*(1), 101-113.
- Degner, J. (in prep.). Prejudice towards obese people. Dissertation an der Universität des Saarlandes in Saarbrücken.

- Degner, J. & Wentura, D. (2005). Assessing intergroup attitudes with the EAST. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- de Liver, Y., Wigboldus, D. & van der Pligt, J. (2005). Ambivalent attitudes and the EAST: the influence of evaluative context. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Denz, H. (1988). Strukturen des Konservatismus. *Angewandte Sozialforschung*, 15, 75-82.
- De Raedt, R., Schacht, R. & De Houwer, J. (2004). Depressed patients have positive implicit self-esteem, submitted for publications.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality & Social Psychology*, 56, 5-18.
- Devine, P.G. (2001). Implicit prejudice and stereotyping : How automatic are they? Introduction to the special section. *Journal of Personality and Social Psychology*, 81 (5), 757-759.
- Devine, P. G. & Baker, S. M. (1991). Measurement of racial stereotype subtyping. *Personality & Social Psychology Bulletin*, 17, 44-50.
- Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality & Social Psychology*, 60, 817-830.
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality & Social Psychology*, 82, 835-848.
- Dollinger, S. J. (2003). Need for uniqueness, need for cognition and creativity. *Journal of Creative Behavior*, 37, 99-116.
- Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. *Journal of Social Issues*, 57(4), 829-849.
- Dovidio, J. F. & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In J.M.Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (204-237).
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 33 (5), 510-540.
- Downing, L.L. & Monaco, N.R. (1986). In-Group/Out-Group Bias as a Function of Differential Contact and Authoritarian Personality. *The Journal of Social Psychology*, 126 (4), 445-452.
- Duckitt, J. H. (1992). *The social psychology of prejudice*. Westport, CT, US: Praeger Publishers/Greenwood Publishing Group, Inc.
- Duckitt, J. (2001). A dual process cognitive-motivational theory of ideology and prejudice. In M.P. Zanna (Hrsg.), *Advances in experimental social psychology*, Vol 33, (41-113). San Diego, CA: Academic Press.
- Duckitt, J., Wagner, C. & du Plessis, I. (2002). The psychological bases of ideology and prejudice: Testing a dual process model. *Journal of Personality & Social Psychology*, 83(1), 75-93.
- Duden - *Das Fremdwörterbuch* (2005). Bibliographisches Institut & F.A. Mannheim: Brockhaus AG.
- Dunton, B. C. & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality & Social Psychology Bulletin*, 23, 316-326.
- Duriez, B. & Van Hiel, A. (2002). The march of modern fascism. A comparison of social dominance orientation and authoritarianism. *Personality & Individual Differences*, 32, 1199-1213.
- Durrheim, K. (1998). The relationship between tolerance of ambiguity and attitudinal conservatism: A multidimensional analysis. *European Journal of Social Psychology*, 28(5), 731-753.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL, US: Harcourt Brace Jovanovich College Publishers.
- Eagly, A. H. & Mladinic, A. (1989). Gender stereotypes and attitudes towards women and men. *Personality and Social Psychological Bulletin*, 15, 543-558.



- Eckhardt, W. (1991). Authoritarianism. *Political Psychology*, 12, 97-124.
- Edwards, A.L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, 65 (2), 241-261.
- Eid, M., Lischetzke, T., Nussbeck, F. & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods*, 8(1), 38-60.
- Egloff, B. & Schmukle (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality & Social Psychology*, 83, 1441-1455.
- Egloff, B. & Schmukle, S. C. (2004). Gender differences in implicit and explicit anxiety measures. *Personality & Individual Differences*, 36, 1807-1815.
- Egloff, B., Schwerdtfeger, A. & Schmukle, S. C. (2005). Temporal Stability of the Implicit Association Test--Anxiety. *Journal of Personality Assessment*, 84(1), 82-88.
- Egloff, B., Wilhelm, F. H., & Neubauer, D. (2002). Implicit anxiety measure predicts cardiovascular reactivity to an evaluated speaking task. *Emotion*, 2, 3-11.
- Ekehammar, B., Akrami, N., & Araya, T. (2003). Gender differences in implicit prejudice. *Personality & Individual Differences*, 34, 1509-1523.
- Ellwart, T., Rinck, M. & Becker, E. (2005). Comparing indirect measures of associations: Application of STIAT and EAST to fear of spiders. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Eysenck, H.J. (1954). *The Psychology of Politics*. London: Routledge & Kegan Paul.
- Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In: C. Hendrick & M. S. Clark: *Research methods in personality and social psychology* ( 74-97). Thousand Oaks, CA, US: Sage Publications, Inc.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobstrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality & Social Psychology*, 69, 1013-1027.
- Fazio, R. H. & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology*, 54, 297-327.
- Fazio, R. H., Roskos-Ewoldsen, D. R., & Powell, M. C. (1994). *Attitudes, perception, and attention*. San Diego: Academic Press, Inc.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality & Social Psychology*, 50, 229-238.
- Fazio, R. H. & Towles-Schwen, T. (1990). The MODE model of attitude-behavior processes. In S. Chaiken, & Y. Trope: *Dual-process theories in social psychology* (97-116). New York, NY, US: Guilford Press.
- Feather, N.T. (1996). Reactions to Penalties for an Offense in Relation to Authoritarianism, Values: Perceived Responsibility, perceived Seriousness, and Deservingness. *Journal of Personality and Social Psychology*, 71 (3), 571-587.
- Fenigstein, A., Scheier, M. F. & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting & Clinical Psychology*, 43(4), 522-527.
- Fichter, M. M. & Quadflieg, N. (2000). Comparing self- and expert rating: A self-report screening version (SIAB-S) of the Structured Interview for Anorexic and Bulimic Syndromes for DSM-IV and ICD-10 (SIAB-EX). *European Archives of Psychiatry & Clinical Neuroscience*, 250(4), 175-185.
- Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *Journal of Personality*, 18, 108-143.
- Frindte, W., Funke, F. & Jacob, S. (1997). Autoritarismus, Wertorientierungen und jugendliche Identifikation. *Gruppendynamik*, 28, 273-289.

- Frindte, W., Funke, F. & Jacob, S. (1999). Neu-alte Mythen über Juden: Ein Forschungsbericht. In R. Dollase, T. Kliche & H. Moser (Hrsg.), *Politische Psychologie der Fremdenfeindlichkeit. Opfer-Täter-Mittäter* (119-130). Weinheim: Juventa.
- Frings, C. & Wentura, D. (2003). Who is watching Big Brother? TV consumption predicted by masked affective priming. *European Journal of Social Psychology, 33*, 779-791.
- Funke, F. (1996). *Autoritarismus und Ambiguitätsintoleranz. Die Bedeutung der politischen Orientierung für die Kovariation beider Konstrukte*. Unveröffentlichte Diplomarbeit. Friedrich-Schiller-Universität Jena.
- Gaertner, S. L. & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46*, 23-30.
- Garrow, J. (1974). *Energy balance and obesity in man*. New York: American Elsevier.
- Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology, 49*, 171-180.
- Gawronski, B. & Conrey, F. R. (2004). Der Implizite Assoziationstest als Maß automatisch aktivierter Assoziationen: Reichweite und Grenzen. *Psychologische Rundschau, 55(3)*, 118-126.
- Gemar, M. C., Segal, Z. V. & Sagrati, S. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology, 110(2)*, 282-289.
- Grayson, D. & Marsh, H.W. (1994). Identification with deficient rank loading matrices in confirmatory factor analysis: Multitrait-Multimethod-Models. *Psychometrika, 59* (1), 121-134.
- Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., Banaji, M. R., & Rudman, L. A. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.
- Greenwald, A. G. & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality & 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 & Social Psychology, 74*, 1464-1480.
- Grover, V. P., Keel, P. K., & Mitchell, J. P. (2003). Gender Differences in Implicit Weight Identity. *International Journal of Eating Disorders, 34*, 125-135.
- Gschwendner-Lukas T., Hofmann W. & Schmitt M. (2004). Motivation zur Vorurteilkontrolle und soziale Normen: Eine Validierungsstudie. *Poster, präsentiert auf dem 44. Kongress der Deutschen Gesellschaft für Psychologie*, Göttingen, September, 2004.
- Haddock, G, Zanna, M.P. & Esses, V.M. (1993). Assessing the structure of prejudicial attitudes. The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology, 65* (6), 1105-1118.
- Hänsel D. (1995): Eß-Störungen. In: V. Faust: *Psychiatrie - Ein Lehrbuch für Klinik, Praxis und Beratung* (655-664). Stuttgart: Gustav-Fischer-Verlag
- Hänze, M. (2002). Beduerfnis nach Struktur und Furcht vor Festlegung: Psychometrische Analysen einer deutschsprachigen Skala zur Erfassung der Konstrukte "Personal Need for Structure", "Personal Fear of Invalidity" und "Need for Closure". *Zeitschrift fuer Differentielle und Diagnostische Psychologie, 23*, 327-338.
- Harris, M. B., Walters, L. C., & Waschull, S. (1991). Gender and ethnic differences in obesity-related behaviors and attitudes in a college sample. *Journal of Applied Social Psychology, 21*, 1545-1566.
- Harris, M. B. & Foltz, S. (1999). Attitudes toward weight and eating in young women tennis players, their parents, and their coaches. *Eating Disorders: The Journal of Treatment & Prevention, 7*, 191-205.
- Hassan, M.K. & Khaliq, A. (1987). A study of prejudice in Hindu & Muslim college students. *Psychologia: Annual International Journal of Personality in the Orient, 30*, 80-84.
- Hayes, B. C. (1995). Religious identification and moral attitudes: The British case. *British Journal of Sociology, 46*, S. 457-474.

- Heaven, P. C. L. & Quintin, D. St. (2003). Personality factors predict racial prejudice. *Personality & Individual Differences, 34*, 625-634.
- Heaven, P. C. L. & Connors, J. R. (2001). TI- A note on the value correlates of social dominance orientation and right-wing authoritarianism. *Personality & Individual Differences, 31*, 925-930.
- Hebl, M. R., Foster, J. B., Mannix, L. & Dovidio, J. (2002). Formal and interpersonal discrimination: A field study of bias toward homosexual applicants. *Personality & Social Psychology Bulletin, 28(6)*, 815-825.
- Hebl, M. R. & Mannix, L. M. (2003). The weight of obesity in evaluating others: A mere proximity effect. *Personality & Social Psychology Bulletin, 29*, 28-38.
- Herek, G. M (1994). Assessing heterosexuals' attitudes toward lesbians and gay men: A review of empirical research with the ATLG scale. In: Greene, B. & Herek, G. M., *Lesbian and gay psychology: Theory, research, and clinical applications* (206-228). Thousand Oaks, CA, US: Sage Publications, Inc.
- Herek, G. M & Capitanio, J. P. (1996). "Some of my best friends": Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians. *Personality & Social Psychology Bulletin, 22(4)*, 412-424.
- Herek, G. M. (2000). The Psychology of Sexual Prejudice. *Current Directions in Psychological Science, 9*.
- Hess, E. H. (1965). Attitude and pupil size. *Scientific American, 212(4)*, 46-54.
- Hibbard, S., Farmer, L., & Wells, C. (1994). Validation of Cramer's Defense Mechanism Manual for the TAT. *Journal of Personality Assessment, 63*, 197.
- Hibbard, S., Mitchell, D., & Porcelli, J. (2001). Internal Consistency of the Object Relations and Social Cognition Scales for the Thematic Apperception Test. *Journal of Personality Assessment, 77*, 408-419.
- Hoffmann, J., Kiesel, A. & Sebald, A. (2003). Task switches under Go/NoGo conditions and the decomposition of switch costs. *European Journal of Cognitive Psychology, 15 (1)*, 101-128.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. (in press). A Meta-Analysis on the Correlation between the Implicit Association Test and Explicit Self-Report Measures. *Personality and Social Psychology Bulletin*.
- Hofmann, W., Gschwendner, T. & Schmitt, M. (2005). On implicit-explicit consistency: The role of individual differences in awareness and adjustment. *European Journal of Personality, 19*, 25-49.
- Holtgraves, T. (2004). Social Desirability and Self-Reports: Testing Models of Socially Desirable Responding. *Personality and Social Psychology Bulletin, 30 (2)*, 161-172.
- Hoyle, R.H. (1995). *Structural equation modelling: Concepts, issues and applications*. Thousand Oaks: Sage Publications.
- Hugenberg, K. & Bodenhausen, G. V. (2004). Ambiguity in Social Categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15*, 342-345.
- Huijding, J. (2005). The malleability of automatic spider-fear related affective associations over the course of an exposure treatment. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Huijding, J., De Jong, P. J. & ter Hart, L. (2003). Specific Predictive Power of Implicit Associations for Automatic Fear Behavior, submitted for publication.
- Hummert, M. L., Garstka, T. A., & O'Brien, L. T. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology & Aging, 17*, 482-495.
- Jacoby, L. L., Lindsay, D. S. & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist, 47*, 802-809.
- James W. P. T, Ferro-Luzzi A. & Waterlow J. C. (1988): Definition of chronic energy deficiency in adults. *European Journal of Clinical Nutrition, 42*, 969-981.
- Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and Explicit Measures of Sexual Orientation Attitudes: Ingroup Preferences and Related Behaviors and Beliefs Among Gay and Straight Men. *Personality & Social Psychology Bulletin, 30*, 629-642.

- Judd, C. M., Drake, R. A. & Downing, J. W. (1991). Some dynamic properties of attitude structures: Context-induced response facilitation and polarization. *Journal of Personality & Social Psychology*, 60(2), 193-202.
- Karpinski, A. & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality & Social Psychology*, 81, 774-788.
- Katz, I. & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality & Social Psychology*, 55, 893-905.
- Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality & Social Psychology Bulletin*, 24, 407-416.
- Kawakami, K. & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality & Social Psychology Bulletin*, 27, 212-225.
- Kawakami, K., Dovidio, J. F., & Dijksterhuis, A. (2003). Effect of social category priming on personal attitudes. *Psychological Science*, 14, 315.
- Kawakami, K., Spears, R., & Dovidio, J. F. (2002). Disinhibition of stereotyping: Context, prejudice, and target characteristics. *European Journal of Social Psychology*, 32, 517-530.
- Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activations. *Personality & Social Psychology Bulletin*, 28, 3-15.
- Kämpfe, N. (2002). *Persönlichkeit, soziale Einstellungen und Fremdenfeindlichkeit: Eine empirische Untersuchung mit Selbst- und Bekannteneinschätzung*.
- Kenny, D.A. & Kashy, D.A. (1992). Analysis of the Multitrait-Multimethod Matrix by confirmatory factor analysis. *Psychological Bulletin*, 112 (1), 165-172.
- Kirscht, J.P. & Dillehay, R.C. (1967). *Dimensions of authoritarianism*. Lexington, Ky: University of Kentucky Press.
- Kischkel, K.-H. (1983). Zum Effekt einiger Ausbildungsfaktoren auf die Ambiguitätstoleranz von Lehrern und die Beziehung zwischen Ambiguitätstoleranz und Konservatismus. *Zeitschrift für Empirische Pädagogik*, 7, 23-41.
- Kite, M. E. & Whitley, B. E. Do heterosexual women and men differ in their attitudes toward homosexuality? A conceptual and methodological analysis. In G. M. Herek: *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals* (39-61). Thousand Oaks, CA, US: Sage Publications, Inc.
- Klein, A. (2000). *Moderatormodelle: Verfahren zur Analyse von Moderator-effekten in Strukturgleichungsmodellen*. Hamburg: Kovac.
- Klein, O., Snyder, M. & Livingston, R. W. (2004). Prejudice on the stage: Self-monitoring and the public expression of group attitudes. *British Journal of Social Psychology*, 43(2), 299-314.
- Kline, R. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Koestner, R., Weinberger, J., & McClelland, D. C. (1991). Task-intrinsic and social-extrinsic sources of arousal for motives assessed in fantasy and self-report. *Journal of Personality*, 59, 57-82.
- Kraiger, K., Hakel, & Cornelius. Exploring Fantasies of TAT Reliability. *Journal of Personality Assessment* 48(4), 365. 1984. Lawrence Erlbaum Associates.
- Kruglanski, A. W., Atash, M. N., DeGrada, E., Manetti, L., Pierro, A., & Webster, D. M. (1997). Psychological theory testing versus psychometric nay-saying: Comment on Neuberg et al.'s (1997) critique of the Need for Closure Scale. *Journal of Personality & Social Psychology*, 73, 1005-1016.
- Kruglanski, A. W. & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review*, 103, 263-283.
- Kruglanski, A. W. & Webster, D. M. (2000). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review*, 103, 263-283.

- Kühnen, U., Schiessl, M., Bauer, N., Paulig, N., Poehlmann, C., & Schmidhals, K. (2001). How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. *Zeitschrift fuer Experimentelle Psychologie*, 48, 135-144.
- Kwon, P., Campbell, & Williams (2001). Sociotropy and Autonomy: Preliminary Evidence for Construct Validity Using TAT Narratives. *Journal of Personality Assessment*, 77, 128.
- LaMar, L. & Kite, M. (1998). Sex differences in attitudes toward gay men and lesbians: A multidimensional perspective. *Journal of Sex Research*, 35, 189-196.
- Lance, C.E., Noble, C.L. & Scullen, S.E. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. *Psychological Methods*, 7 (2), 228-244.
- Laythe, B., Finkel, D., Bringle, R., & Kirkpatrick, L. A. (2002). Religious fundamentalism as a predictor of prejudice: A two-component model. *Journal for the Scientific Study of Religion*, 41, 623-635.
- Leak, G. K. & Randall, B. A. (1995). Clarification of the link between right-wing authoritarianism and religiousness: The role of religious maturity. *Journal for the Scientific Study of Religion*, 34(2), 245-252.
- Lehrl, S., Merz, J., Burkard, G. & Fischer, B. (1991). Mehrfachwahl-Wortschatz-Intelligenztest (MWT-A). Erlangen: perimed Fachbuch-Verlagsgesellschaft mbH.
- Leigh, J., Westen, D., & Barends, A. (1992). The assessment of complexity of representations of people using TAT and interview data. *Journal of Personality*, 60, 809-837.
- Leone, C., Wallace, H. M., & Modglin, K. (1999). The need for closure and the need for structure: Interrelationships, correlates, and outcomes. *Journal of Psychology*, 133, 553-562.
- Lepore, L. & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality & Social Psychology*, 72, 275-287.
- Levin, S. & Sidanius, J. (1999). Social dominance and social identity in the United States and Israel: In-group favoritism or outgroup derogation? *Political Psychology*, 20(1), 99-126.
- Lippa, R. & Arad, S. (1999). Gender, Personality, and Prejudice: The Display of Authoritarianism and Social Dominance in Interviews with College Men and Women. *Journal of Research in Personality*, 33, 463-493.
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality & Social Psychology*, 81, 842-855.
- Mackenzie, M. (1984). *Fear of fat*. New York: Columbia University Press.
- Madigan, S. (1983). Picture memory, In J.C. Yuille (Ed.), *Imagery, memory and cognition: Essays in honor of Allan Paivio*. Hillsdale, NJ: Erlbaum.
- Maison, D., Greenwald, A. G., & Bruin, R. (2001). The Implicit Association Test as a measure of implicit consumer attitudes. *Polish Psychological Bulletin*, 32, 61-69.
- Marlowe, D. & Crowne, D.P (1960). *The approval motive-studies in evaluative dependence*. New York: Wiley.
- Marsh, H.W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, 13, 335-361.
- Marsh, H.W. & Craven, R.G. (1991). Self-other agreement on multiple dimensions of preadolescent self-concept inferences by teachers, mothers, and fathers. *Journal of Educational Psychology*, 83 (3), 393-404.
- May, K. & Hittner, J. B. (2003). On the relation between power and reliability of difference scores. *Perceptual & Motor Skills*, 97(3), 905-908.
- McClelland, D. C. (1985). How motives, skills, and values determine what people do. *American Psychologist*, 40, 812-825.
- McClelland, D. C. (1995). Achievement motivation in relation to achievement-related recall, performance, and urine flow, a marker associated with release of vasopressin. *Motivation & Emotion*, 19, 59-76.
- McClelland, D. C. & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied Psychology*, 67(6), 737-743.

- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690-702.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J.F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (91-125).
- McConnell, A. R. & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435-442.
- McFarland, S.G. & Adelson, S. (1996). *An omnibus study of personality, values, and prejudice. Vortrag präsentiert auf dem Kongress der International Society for Political Psychology*, Vancouver, British Columbia, July 1996.
- McFarland, S. G. & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, *20*(6), 483-510.
- McGivern, R. F., Mutter, K. L., Anderson, J., Wideman, G., Bodnar, M., & Huston, P. J. (1998). Gender differences in incidental learning and visual recognition memory: Support for a sex difference in unconscious environmental awareness. *Personality & Individual Differences*, *25*, 223-232.
- McGregor, I., Zanna, M.P., Holmes, J.G. & Spencer, S.J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology*, *80* (3), 472-488.
- Meier, B. & Perrig, W. J. (2000). Low Reliability of Perceptual Priming: Consequences for the Interpretation of Functional Dissociations Between Explicit and Implicit Memory. *Quarterly Journal of Experimental Psychology: Section A*, *53* (1), 211-233.
- Meyers Enzyklopädisches Lexikon in 25 Bänden (1971-1981, 9. Aufl.). Mannheim: Bibliographisches Institut.
- Meyer, D. E., Glass, J. M. & Mueller, S. T. (2001). Executive-process interactive control: A unified computational theory for answering 20 questions (and more) about cognitive ageing. *European Journal of Cognitive Psychology*, *13*(1-2), 123-164.
- Mierke, J., Musch, J. & Klauer, K. C. (2003). The positivity proportion effect: A list context effect in masked affective priming. *Memory & Cognition*, *31*(6), 953-967.
- Mierke, J. & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift fuer Experimentelle Psychologie*, *48* (2), 107-122.
- Mierke, J. & Klauer, K. C. (2003). Method-Specific Variance in the Implicit Association Test. *Journal of Personality & Social Psychology*, *85* (6), 1180-1192.
- Millman, J. (1980). Implications of minimum competence testing programs for achievement and achievement testing. *New Directions for Testing & Measurement*, *5*, 31-38.
- Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*, 455-469.
- Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (2000). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. In C. Stangor (Ed.), *Stereotypes and prejudice: Essential readings* (305-322).
- Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition*, *19*, 395-417.
- Morrison, T. G., McLeod, L. D., Morrison, M. A. (1997). Gender stereotyping, homonegativity, and misconceptions about sexually coercive behavior among adolescents: Erratum. *Youth & Society*, *29*(1), 134.
- Morrison, T. G., Parriag, A. V. & Morrison, M. A. (1999). The psychometric properties of the homonegativity scale. *Journal of Homosexuality*, *37*(4), 111-126.
- Muller, J. (1997). *Conservatism. An anthropology of social and political thought*. Princeton: University Press.
- Murray, H. A. (1943). *Thematic apperception test*. Cambridge, MA, US: Harvard University Press.

- Musch, J. & Klauer, K. C. (2003). *The psychology of evaluation: Affective processes in cognition and emotion*. Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Musch, J., Brockhaus, R., & Broeder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit. *Diagnostica, 48*, 121-129.
- Neubauer, A. C. (1997). The mental speed approach to the assessment of intelligence In: J. Kingma and W. Tomic, (Eds), *Advances in cognition and educational practice: Reflections on the concept of intelligence* (149-174). Greenwich, Connecticut: JAI Press.
- Neuberg, S. L., Judice, T. N., & West, S. G. (1997). What the Need for Closure Scale measures and what it does not: Toward differentiating among related epistemic motives. *Journal of Personality & Social Psychology, 72*, 1396-1412.
- Neuberg, S. L., West, S. G., Thompson, M. M., & Judice, T. N. (1997). On dimensionality, discriminant validity, and the role of psychometric analyses in personality theory and measurement: Reply to Kruglanski et al.'s (1997) defense of the Need for Closure Scale. *Journal of Personality & Social Psychology, 73*, 1017-1029.
- Neuberg, S. N. & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality & Social Psychology, 65*, 113-131.
- Neumann, R., Hülsenbeck, K. & Seibt, B. (2004). Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of Experimental Social Psychology, 40*(4), 543-550.
- Neumann, R. & Seibt, B. (2001). The structure of prejudice: Associative strength as a determinant of stereotype endorsement. *European Journal of Social Psychology, 31*, 609-620.
- Nosek, B. A. (2003). Moderators of the relationship between implicit and explicit attitudes. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 63*.
- Nosek, B. A. & Banaji, M. R. (2002). (At least) two factors moderate the relationship between implicit and explicit attitudes. In R.K.Ohme & M. Jarymowicz (Eds.), *Natura Automatyzmow* (49-56). Warszawa: WIP PAN & SWPS.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math not = me. *Journal of Personality & Social Psychology, 83*, 44-59.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics, 6*, 101-115.
- Nosek, B. A. & Banaji, M. R. (2001). The GO/NO-GO Association Task. *Social Cognition, 19*, 625-668.
- Nosek, B. A., Greenwald, A. G. & Banaji, M. R. (2005). Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity. *Personality & Social Psychology Bulletin, 31*(2), 166-180.
- Nosek, B. A., & Hansen, J. J. (2004). *The associations in our heads belong to us: Measuring the multifaceted attitude construct in implicit social cognition*. Unpublished manuscript. University of Virginia.
- Oesterreich, D. (1993). *Autoritäre Persönlichkeit und Gesellschaftsordnung. Der Stellenwert psychischer Faktoren für politische Einstellungen – eine empirische Untersuchung von Jugendlichen in Ost und West*. Weinheim: Juventa.
- Olson, M. A. & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14*, 636-639.
- Olson, M. A. & Fazio, R. H. (2004). Reducing the Influence of Extrapersonal Associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality & Social Psychology, 86*, 653-667.
- Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitudes and racism: Effects of word familiarity and frequency on the implicit association test. *Social Cognition, 19*, 97.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287.
- Paulhus, D. L. & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality & Social Psychology, 60*, 307-317.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York, NY, US: Cambridge University Press.

- Pelham, B. W., Koole, S. L., Hardin, C. D., Hetts, J. J., Seah, E. & DeHart, T. (2005). Gender moderates the relation between implicit and explicit self-esteem. *Journal of Experimental Social Psychology*, 41(1), 84-89.
- Pentony, J.F., Petersen, K.S.E., Philips, O., Leong, C., Harper, P., Bakowski, A., Steward, S. & Gonzales, R. (2000). A comparison of authoritarianism in the United States, England, and Hungary with selected non-random samples. *European Psychologist*, 5 (4), 259-268.
- Perugini, M., O'Gorman, R. & Connor, M. (2005). The EAST as an inter-individual difference measure: Promising results from three studies. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Petty, R. E. & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Wm. C. Brown Company Publishers.
- Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology*, 48, 609-647.
- Phelps, E. A.; O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C. & Bagnaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12 (5), 729-738.
- Plant, E. A. & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality & Social Psychology*, 75, 811-832.
- Pratto, F. (1994). Consciousness and automatic evaluation. , In P. M. Niedenthal & S. Kitayama: *Heart's eye: Emotional influences in perception and attention* (115-143). San Diego, CA, US: Academic Press
- Pratto, F. & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality & Social Psychology*, 61(3), 380-391.
- Pratto, F. & Shih, M. (2000). Social Dominance Orientation and Group Context in Implicit Group Prejudice. *Psychological Science*, 11, 515-518.
- Pratto, F., Sidanius, J., Stallworth, L.M. & Malle, B.F. (1994). Social dominance orientation, A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67 (4), 741-763.
- Prohansky, H. M. (1966). The development of intergroup attitudes. In L. W. Hoffman & M. L. Hoffman (Eds.), *Review of child development research*, Vol. 2, (311-371). New York: Russell Sage Foundation.
- Ray, J.J. & Lovejoy, F.H. (1986). The generality of racial prejudices. *Journal of Social Psychology*, 126, 563-564.
- Reis, J. (1997). *Ambiguitätstoleranz: Beiträge zur Entwicklung eines Persönlichkeitskonstruktes*. Heidelberg: A-sanger.
- Riemann, R., Grubich, C., Hempel, S., & Mergl, S. (1993). Personality and attitudes towards current political topics. *Personality & Individual Differences*, 15, 313-321.
- Rippl, S. & Seipel, C. (1998). Autoritarismus und Fremdenfeindlichkeit bei ost- und westdeutschen Jugendlichen. *Zeitschrift für Politische Psychologie*, 6 (3), 273-288.
- Roefs, A. & Jansen, A. (2002). Implicit and explicit attitudes toward high-fat foods in obesity. *Journal of Abnormal Psychology*, 111, 517-521.
- Rogosa, D., Brandt, D. & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, 92(3), 726-748.
- Ronan, G. F. (1990). Manual for the Personal Problem-Solving-System for Scoring Thematic Apperception Test responses Personal-Problem-Solving System for Scoring TAT Responses. Unpublished manuscript.
- Ronan, G. F., Colavito, V. A., & Hammontree, S. R. (1993). Personal Problem-Solving System for Scoring TAT Responses: Preliminary Validity and Reliability Data. *Journal of Personality Assessment*, 61, 28-40.
- Ronan, G. F., Date, A. L., & Weisbrod, M. (1995). Personal problem-solving scoring of the TAT: Sensitivity to training. *Journal of Personality Assessment*, 64, 119-131.



- Ronan, G. F., Senn, J., & Date, A. L. (1996). Personal Problem-Solving Scoring of TAT Responses: Known-Groups Validation. *Journal of Personality Assessment*, 67, 641-653.
- Rosenberg, S. D., Blatt, S. J., & Oxman, T. E. (1994). Assessment of Object Relatedness Through a Lexical Content Analysis of the TAT. *Journal of Personality Assessment*, 63, 345-362.
- Rosenberg, M. J., Hovland, C. I. & McGuire, W. J. (1960). *Attitude organization and change: An analysis of consistency among attitude components (Yales studies in attitude and communication. Vol. III.)*. Oxford, England: Yale University Press.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Göttingen: H. Huber.
- Rothblum, E. D., Brand, P. A., Miller, C. T. & Oetjen, H. A. (1990). The relationship between obesity, employment discrimination, and employment-related victimization. *Journal of Vocational Behavior*, 37(3), 251-266.
- Rothermund, K. & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift fuer Experimentelle Psychologie*, 48, 94-196.
- Rothermund, K. & Wentura, D. (2004). Underlying Processes in the Implicit Association Test: Disassociating Salience From Associations. *Journal of Experimental Psychology: General*, 133, 139-165.
- Rowatt, W. C. & Franklin, L. M. (2004). Christian orthodoxy, religious fundamentalism, and right-wing authoritarianism as predictors of implicit racial prejudice. *International Journal for the Psychology of Religion*, 14, 125-138
- Rowatt, W. C., Franklin, L. M., & Cotton, M. (2005). Patterns and Personality Correlates of Implicit and Explicit Attitudes Toward Christians and Muslims. *Journal for the Scientific Study of Religion*, 44, 29-43.
- Ruch, W. & Hehl, F.-J. (1983). Intolerance of Ambiguity as a factor in the appreciation of humour. *Personality & Individual Differences*, 4(5), 443-449.
- Rudman, L. A., Greenwald, A. G., & Mellott, D. S. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition*, 17, 437-465.
- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality & Social Psychology*, 81, 856-868.
- Rudman, L. A., Feinberg, J. M., & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition*, 20, 294-320.
- Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, 13, 79-82.
- Sanbonmatsu, D. M. & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality & Social Psychology*, 59(4), 614-622.
- Sarma, C. (1973). Prejudice and personality. In T.E. Stranumgam (Hrsg.). *Research in personality and social problems* (96-114). Madras: University of Madras Press.
- Saucier, G. (2000). Isms and the structure of social attitudes. *Journal of Personality and Social Psychology*, 78 (2), 366-385.
- Scheffer, D., Kuhl, J. & Eichstaedt, J. (2003). Der Operante Motiv-Test (OMT): Inhaltsklassen, Auswertung, psychometrische Kennwerte und Validierung. In F. Rheinberg & J. Stiensmeier-Pelster (Hrsg.): *Diagnostik von Motivation und Selbstkonzept* (S. 151-168). Göttingen; Hogrefe.
- Schenk, J. (1980). Clericalism and conservatism. *Psychologische Beiträge*, 22(1), 145-162.
- Schmalt, H.-D. & Sokolowski, K. (2000). Zum gegenwärtigen Stand der Motivdiagnostik. *Diagnostica*, 46, 115-123.
- Schmidt, T.E. (1995). *Straight & narrow? Compassion & clarity in the homosexual debate*. Aus Sonderdruck Verlag VBG 1997 (Hrsg: H. Gasser & W. Gasser).
- Schmitz, F., Teige, S., Voss A. & Klauer, C. (2005). Cognitive capacity, task shifts, and measurement issues in the EAST procedure. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Schmukle, S. C. & Egloff, B. (2004). Does the Implicit Association Test for Assessing Anxiety Measure Trait and State Variance? *European Journal of Personality*, 18(6), 483-494.

- Schmukle, S. C. & Egloff, B. (2005). A Latent State-Trait Analysis of Implicit and Explicit Personality Measures. *European Journal of Psychological Assessment, 21*(2), 100-107.
- Schuette, R. A. & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality & Social Psychology Bulletin, 21*(7), 704-710.
- Schultheiss, O. C. & Brunstein, J. C. (2001). Assessment of Implicit Motives With a Research Version of the TAT: Picture Profiles, Gender Differences, and Relations to Other Personality Measures. *Journal of Personality Assessment, 77*, 71-86.
- Seise, J., Banse, R., & Neyer, F. (2002). Individuelle Unterschiede in impliziten und expliziten Einstellungen zur Homosexualität: Eine empirische Studie. *Zeitschrift fuer Sexualforschung, 15*, 21-42.
- Sharkey, K. J. & Ritzler (1985). Comparing Diagnostic Validity of the TAT and a New Picture Projective Test. *Journal of Personality Assessment, 49*, 406-412.
- Sherman, S. J., Chassin, L., Presson, C., Rose, J., & Koch, K. (1999). The relationship between the implicit and explicit attitudes of smokers and nonsmokers toward cigarette smoking. *Vortrag präsentiert bei der Midwest Psychological Association, Chicago, April, 1999.*
- Sherrod, D. & Nardi, P. M. (1998). Homophobia in the courtroom: An assessment of biases against gay men and lesbians in a multiethnic sample of potential jurors. In G. M. Herek: *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals* (24-38). Thousand Oaks, CA, US: Sage Publications, Inc.
- Sidanius, J. (1978). Intolerance of ambiguity and socio-political ideology: A multidimensional analysis. *European Journal of Social Psychology, 8*(2), 215-235.
- Sidanius, J. & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression.*
- Sidanius, J., Pratto, F., & Bobo, L. (1994). Social dominance orientation and the political psychology of gender: A case of invariance? *Journal of Personality & Social Psychology, 67*, 998-1011.
- Sinha, R.R. & Hassan, M.K. (1975). Some personality correlates of social prejudice. *Journal of Social and Economic Studies, 3*, 225-231.
- Skowronski, J. J. & Lawrence, M. A. (2001). A Comparative Study of the Implicit and Explicit Gender Attitudes of Children and College Students. *Psychology of Women Quarterly, 25*, 155-165.
- Six, W. (1996). Generalisierte Einstellungen. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie. Differentielle und Persönlichkeitsforschung. Bd. 3: Temperaments- und Persönlichkeitsunterschiede* (1-50). Göttingen: Hogrefe.
- Six, B., Wolfradt, U. & Zick, A. (2001). Autoritarismus und Soziale Dominanzorientierung als generalisierte Einstellungen. *Zeitschrift für Politische Psychologie, 9* (2-3), 23-40.
- Snyder, M. (1992). Motivational foundations of behavioral confirmation. In M. P. Zanna: *Advances in experimental social psychology*, Vol. 25, (67-114). San Diego, CA, US: Academic Press.
- Son Hing, L. S., Li, W. & Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology, 38*(1), 71-78.
- Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*, 140-154.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 3*, 271-295.
- Stahl, C. & Degner, J. (2005). Automatic and controlled processes in the EAST. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop, Oppurg, 08.-10. April, 2005.*
- Stangor, C. (2000). *Stereotypes and prejudice: Essential readings.* New York, NY, US: Psychology Press.
- Steffens, M. C. (2004). Is the Implicit Association Test Immune to Faking? *Experimental Psychology, 51*(3), 165-179.
- Steffens, M. C. & Buchner, A. (2003). Implicit Association Test: Separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology, 50*, 33-48.

- Steffens, M. C., Lichau, J. & Still, Y. (2004). Individuum oder Gruppe, Exemplar oder Kategorie?: Ein Zweifaktorenmodell zur Erklärung der Reaktionszeitunterschiede im Implicit Association Test(IAT). *Zeitschrift für Psychologie*, 212(2), 57-65.
- Steffens, M. C. & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift fuer Experimentelle Psychologie*, 48, 123-134.
- Steffens, M. C. & Wagner, C. (2004). Attitudes Toward Lesbians, Gay Men, Bisexual Women, and Bisexual Men in Germany. *Journal of Sex Research*, 41(2), 137-149.
- Stoeber, J. (1999). Die Soziale-Erwuenschttheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität. *Diagnostica*, 45, 173-177.
- Strack, F. & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality & Social Psychology Review*, 8, 220-247.
- Stunkard, A. J., Sorensen, T. I., Hanis, C., & Teasdale, T. W. (1986). "An adoption study of human obesity": Reply. *New England Journal of Medicine*, 315, 130.
- Suslow, T., Ohrmann, P. & Arolt, V. (2001). Ausgestaltung so genannter affektiver Voraktivierungseffekte in der evaluativen Entscheidungsaufgabe: Hinweise auf automatische Vigilanz für negative Informationen. *Zeitschrift für Psychologie*, 209(2), 137-152.
- Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behaviour consistency for stigmatised behaviour. *Cognition & Emotion*, 15, 207-230.
- Tabachnik, B.G. & Fidell, L.S. (2001). *Using multivariate statistics* (4. Ausg.). Boston: Allyn & Bacon.
- Teachman, B. A., Gapinski, K. D., Brownell, K. D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology*, 22, 68-78.
- Teachman, B. A. & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology*, 112, 100-109.
- Tedeschi, J. T., Schlenker, B. R., & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, 26, 685-695.
- Teige, S., Schnabel, K. Banse, R. & Asendorpf, J. (2004). Assessment of Multiple Implicit Self-Concept Dimensions Using the Extrinsic Affective Simon Task (EAST). *European Journal of Personality*, 18(6), 495-520.
- Thrash, T. M. & Elliot, A. J. (2002). Implicit and self-attributed achievement motives: Concordance and predictive validity. *Journal of Personality*, 70, 729-755.
- Tiggemann, M. & Rothblum, E. D. (1988). Gender differences in social consequences of perceived overweight in the United States and Australia. *Sex Roles*, 18(1-2), 75-86.
- Tiggemann, M. & Rothblum, E. D. (1988). Gender differences in social consequences of perceived overweight in the United States and Australia. *Sex Roles*, 18(1-2), 75-86.
- Trautner, H. M. (1992). The development of sex-typing in children: A longitudinal analysis. *German Journal of Psychology*, 16(3), 183-199.
- van Hiel, A. & Mervielde, I. (2002). Explaining conservative beliefs and political preferences: A comparison of social dominance orientation and authoritarianism. *Journal of Applied Social Psychology*, 32 (5), 965-976.
- van Ijzendoorn, M.H. (1989). Moral Judgement, Authoritarianism, and Ethnocentrism. *The Journal of Social Psychology*, 129 (1), 37-45.
- Voss, A., Rothermund, K., & Wentura, D. (2003). Estimating the valence of single stimuli: A new variant of the affective Simon task. *Experimental Psychology*, 50, 86-96.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Oxford, England: Rand McNally.
- Webster, D. M. & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality & Social Psychology*, 67, 1049-1062.

- Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality & Social Psychology*, 39, 186-200.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- Widaman, K.F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied psychological measurement*, 9, (1), 1-26.
- Wiers, R. W., Van Woerden, N., Smulders, F. T. Y & de Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, 111(4), 648-658. Wikipedia, freie Enzyklopädie. <http://de.wikipedia.org/wiki>.
- Wigboldus, D., van Knippenberg, A. & Holland, R. (2001). The difference between relative and absolute implicit associations. *Vortrag präsentiert auf dem dritten Treffen des European Social Cognition Network*, Houfalize, Belgium.
- Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, 120, 3-24.
- Williams, R. H. & Zimmerman, D. W. (1996). Are simple gain scores obsolete: Commentary on the commentaries of Collins and Humphreys. *Applied Psychological Measurement*, 20(3), 295-297.
- Wilson, G.D. (1970). Is there a general factor in social attitudes? Evidence from a factor analysis of the Conservatism scale. *British Journal of Social and Clinical Psychology*, 9, 101-107.
- Wilson, G.D. (1973). *The Psychology of Conservatism*. London: Academic Press.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality & Social Psychology*, 72, 262-274.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. *Journal of Experimental Social Psychology*, 37, 244-252.
- Woike, B. & Polo, M. (2001). Motive-Related Memories: Content, Structure, and Affect. *Journal of Personality*, 69, 391-415.
- Wooley, S. C., Wooley, O. W. & Dyrenforth, S. R. (1979). Theoretical, practical, and social issues in behavioral treatments of obesity. *Journal of Applied Behavior Analysis*, 12(1), 3-25.
- Zick, A. (1997). *Vorurteile und Rassismus: eine Sozialpsychologische Analyse*. Berlin: Waxmann.
- Zimbardo, P. G. (1995). *Psychologie* (6. Aufl.). Berlin: Springer Verlag.
- Zimbardo, P. G. & Perrig, R. J. (2004). *Psychologie: Eine Einführung*. Pearson Studium.
- Zimmerman, D. W. (1996). An efficient alternative to the Wilcoxon signed-ranks test for paired nonnormal data. *Journal of General Psychology*, 123(1), 29-40.
- Zacker, J. (1973). Authoritarian avoidance of ambiguity. *Psychological Reports*, 33(3), 901-902.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151-175.

## 7 Gesamtdiskussion

Vorliegende Arbeit ist der Untersuchung expliziter und impliziter sozialer Einstellungen, ihrer Konstruktvalidität, ihrer direkten bzw. vermittelten Beziehungen untereinander sowie ihrer Bedeutung für andere einstellungsrelevante Konstrukte gewidmet. Zu diesem Zweck wurden Einstellungen zu zwei unterschiedlichen Inhalten, und zwar Übergewicht und Homosexualität, mittels verschiedener Verfahren gemessen. Zum Einsatz kamen in vorliegender Arbeit traditionelle Selbstberichte, für die zum Teil eine Trennung zwischen persönlichen und antizipierten Einstellungsinhalten erfordert wurde und zwei implizite Messverfahren, wovon der Implizite Assoziationstest als relativ etabliert und gut untersucht bezeichnet werden kann (vgl. Gawronski & Conrey 2004), wogegen der Extrinsic Affective Simon Task hauptsächlich aufgrund eines Potentials, zur Erweiterung des diesbezüglichen Kenntnisstandes und zum Vergleich mit dem IAT einbezogen wurde. Zusätzlich wurde im Rahmen der verfolgten Fragestellungen ein Messverfahren zur Erfassung reflektierter, bewusster *und* spontaner, unbewusster Einstellungsinhalte entwickelt und gemeinsam mit expliziten und impliziten Einstellungsmessungen eingesetzt.

Zwei erste Untersuchungen waren zur Entwicklung und Erprobung der, in der dritten Studie schließlich gemeinsam eingesetzten, Einstellungsmessungen angelegt. Mit Studie I wurde maßgeblich eine Untersuchung der psychometrischen Eigenschaften der beiden ausgewählten impliziten Einstellungsmaße am Beispiel von Einstellungen zu Homosexualität verfolgt. Hierbei wurden insbesondere die Auswirkungen einer festgelegten Stimulussequenz für die Reliabilität der impliziten Maße und schließlich die Analyse interindividueller Differenzen in expliziten und impliziten Einstellungen erörtert. Den Ergebnissen von Studie I zufolge, weist der reaktionszeitbasierte IAT-Effekt eine vergleichbar hohe interne Konsistenz auf und korreliert signifikant positiv mit expliziten Einstellungen zu Homosexualität. Der Fehlereffekt des IAT ist deutlich weniger reliabel, unterhält jedoch ebenso positive Zusammenhänge, allerdings nicht signifikante, Beziehungen zu den expliziten Kriterien. Dagegen ist für den EAST keine hinreichende interne Konsistenz gegeben. Es wird sehr deutlich, dass jeweils nur Teileffekte basierend auf einzelnen Messblöcken betrachtet werden können und hier insbesondere der erste Block erwartungskonforme Zusammenhänge zu expliziten Einstellungen aufzeigt (siehe Studie I).

Studie II diente der Entwicklung einer offenen Assoziativmessung, mit dem sowohl explizite, als auch implizite Aspekte von Einstellungen zu Übergewicht bzw. Homosexualität erfasst werden können. In Anlehnung an den Thematischen Apperzeptionstest der als projektive Verfahren insbesondere zur Motivmessung eingesetzt wird, wurden thematisch relevante Bildmotive erstellt, zu denen Probanden eine frei erfundene Geschichte schreiben sollten, die mittels adäquater Kodierschemata inhaltsanalytisch ausgewertet wurden. Zur Evaluation dieser Messung wurden expliziten Einstellungen zu Übergewicht und Homosexualität mit erhoben. Ein Hauptergebnis von Studie II ist das Potential der aus-

gewählten Messmethodik zur Erfassung von Einstellungen. Konsistente Zusammenhänge zu expliziten Einstellungen und einstellungsrelevanter Motivation werden für die Geschichtenlänge in Form der Wortanzahl verzeichnet. Darüber hinaus sind eine Reihe sinnvoller Zusammenhänge für die Valenz der Geschichten festzustellen. Ziel der Evaluation der Bildmotive war schließlich die Auswahl der jeweils zwei am besten geeigneten Motive je Einstellungsobjekt.

Den Kern vorliegender Arbeit bildet Studie III, in der die verschiedenen Einstellungsmessungen gemeinsam in einer Stichprobe von 165 Personen erhoben wurden, wodurch die Untersuchung einer Reihe zentraler Fragestellungen zur Bedeutung expliziter und impliziter Einstellungen zum Teil mit messfehlerbereinigten Variablen erfolgen konnte. Kurz zusammengefasst liefert Studie III Hinweise auf eine notwendige Unterscheidung zwischen expliziten und impliziten Einstellungen, für die keine konvergente Validität angenommen werden kann. Lediglich für Homosexualität korrespondieren IAT und Selbstbericht in statistisch bedeutsamem Ausmaß. Aufgrund fehlender substantieller Korrespondenzen zwischen expliziten und impliziten Maßen resultiert die herausragende Bedeutung methodenspezifischer Varianz für die Variablen. Dagegen werden trotz unterschiedlicher Messmethoden stabile Korrespondenzen zwischen Selbstbericht und Geschichtenvariablen verzeichnet. Es ist jedoch nicht von einer Art Explizit-Implizit-Dimension der Einstellungen auszugehen bzw. zeugen die Beziehungsmuster nicht von einer relativen Zwischenstellung der offenen Assoziativmessung hinsichtlich expliziter und impliziter Einstellungsanteile. Weder die schwachen Zusammenhänge zwischen expliziten und impliziten Einstellungen, noch die engen Beziehungen zwischen expliziten und offen assoziierten Einstellungen lassen sich durch Motivation zur Vorurteilsfreiheit oder Soziale Erwünschtheit erklären. Lediglich für affektive Einstellungen zu Homosexualität und den IAT wird eine zusammenhangssteigernde Wirkung durch die individuelle Motivation zur Vorurteilsfreiheit festgestellt. Dagegen kann gezeigt werden, dass die Berücksichtigung sozialer Erwünschtheit im Fragebogen in Form einer forcierten Trennung zwischen individuellen und gesellschaftlich antizipierten Einstellungen zu zwei distinkten Einstellungsdispositionen führt, deren Abweichung insbesondere für die Beziehung expliziter und impliziter Einstellungen zu Homosexualität bedeutsam ist. Schließlich werden gravierende Unterschiede zwischen den nomologischen Netzwerken expliziter und impliziter Einstellungen aufgezeigt. Während selbstberichtete Einstellungen und implizite Assoziationsstärkemaße nur teilweise in Zusammenhang stehen (hier für Einstellungsobjekt Homosexualität), sind die konsistenten und nicht auf der Wirkung motivationaler Drittvariablen basierenden Korrespondenzen zwischen der Geschichtsmessung und den selbstberichteten expliziten Einstellungen im Fragebogen für beide Einstellungsobjekte hervorzuheben.

Zur Evaluation der Ergebnisse aktueller Arbeit wird im Folgenden eine resümierende Gesamtbetrachtung vorgenommen, der zufolge mögliche Generalisierungen und weiterführende Fragestellungen abgeleitet werden. Im zentralen Fokus der Gesamtdiskussion steht die Konstruktvalidierung sozialer Einstellungen, die Anhaltspunkte zur Konzeptualisierung von expliziten und impliziten Einstellungen liefert. Da adäquate Messverfahren eine Grundvoraussetzung zur Untersuchung inhaltlicher Fragestellungen darstellen, liegt auf einer Bewertung der eingesetzten Verfahren zur Einstellungsmessung ebenfalls großes Gewicht. In diesem Zusammenhang wird eine Evaluation der impliziten Einstellungsmessungen IAT und EAST hinsichtlich ihrer psychometrischen Eigenschaften und der daraus resultierenden Eignung zur Untersuchung interindividueller Differenzen in impliziten Einstellungen vorgenommen. Ein weiterer Diskussionspunkt zur Verfahrensevaluation betrifft die offene Assoziativmessung, die im Rahmen dieser Arbeit zur Erfassung impliziter und expliziter Einstellungsanteile entwickelt wurde. Schließlich werden mögliche Charakteristika der hier betrachteten Einstellungsobjekte Übergewicht und Homosexualität hinsichtlich ihrer Bedeutung für die Ergebnisse der empirischen Untersuchungen dieser Arbeit diskutiert.

## 7.1 Konzeptualisierung von Einstellungen

Mit der Arbeit von Banaji & Greenwald (1995) zur Unterscheidung zwischen expliziten und impliziten Einstellungen und der Publikation des Impliziten Assoziationstestes etwa 4 Jahre später (Greenwald, McGhee & Schwartz, 1998) begann eine geradezu euphorisierte Untersuchung impliziter und expliziter Anteile von *Einstellungen*, *Stereotypen*, *Vorurteilen* (vgl. Devine, 2001), von *Identität*, *Selbstkonzept* und *Selbstwert* (z.B. Aidman & Carroll, 2003; Asendorpf, Banse & Mücke, 2002; Greenwald & Farnham, 2000; Greenwald et al., 2002) oder von dysfunktionalen Überzeugungen bzw. Dispositionen zu *Angst* und *Furcht* (De Jong, 2002; De Jong, Pasma, Kindt, van den Hout, 2001; Egloff & Schmukle, 2002; Schmukle & Egloff, 2004), *Depression* (Gamar, Segal, Sagrati & Kennedy, 2001) oder *Suchtverhalten* (De Houwer et al., 2004; Wiers, van Worden, Smulders & de Jong, 2002).

Theoretische Konzeptualisierungsansätze für explizite und implizite Einstellungen stehen einer äußerst heterogenen Befundlage zur Beziehung zwischen beiden Einstellungsaspekten gegenüber. Im Mittel kann von schwach positiven Zusammenhängen zwischen zumindest IAT-basierten impliziten und selbstberichteten expliziten Einstellungen ausgegangen werden (Hofman et al., 2003), die zu einem beträchtlichen Teil von Drittvariablen moderiert werden. Nosek (2004) findet beispielsweise eine Varianzaufklärung von nahezu 40% an der Beziehung expliziter und impliziter Einstellungen durch die vier Moderatorvariablen *Selbstpräsentation*, *Einstellungsstärke*, *Einstellungsdimensionalität* und *wahrgenommene Abweichung der persönlichen Einstellungen von den Einstellungen der Bezugsgruppe* (vgl. Banse & Gaw-

ronski, 2003; Devine et al., 2002; Hofmann, Gschwendner & Schmitt, 2005; Plant & Devine, 1998). Vorliegende Arbeit reflektiert die inkonsistente Befundlage zur Beziehung impliziter und expliziter Einstellungen sehr anschaulich. Den Ergebnissen von Studie I zufolge, stehen impliziten Einstellungen zu Homosexualität gemäß IAT und EAST<sup>4</sup> in signifikant positivem Zusammenhang mit expliziten Einstellungen zu Homosexualität, wobei eine differentielle Validität für IAT und EAST in Bezug auf kognitive und affektive Einstellungen festzustellen ist (siehe aber Abschnitt 7.2.1). Während der IAT vor allem kognitive Einstellungen zu Homosexualität vorhersagt, erweist sich der EAST als Hauptprädiktor affektiver Einstellungen. Das heißt, für explizite und implizite Einstellungen zu Homosexualität ist eine erwartungskonforme Korrespondenz anzunehmen, die auf konvergente Validität der verschiedenen Einstellungsmessungen zu Homosexualität hindeutet. Statthafte Aussagen über konvergente und diskriminante Validität expliziter und impliziter Einstellungen sind an dieser Stelle jedoch nicht zu treffen, da hierzu mehr als ein Einstellungsinhalt betrachtet werden muss. Um Anhaltspunkte für eine theoretische Konzeptualisierung expliziter und impliziter Einstellungen im Sinne einer *Lügendetektorperspektive* oder einer *Perspektive unabhängiger Bewertung* (Nosek & Banaji, 2002; vgl. Kapitel 2.1.1) zu gewinnen, bedarf es entsprechend einer Konstruktvalidierung. Dies wurde in Studie III mit einer latenten Multitrait-Multimethod-Analyse expliziter und impliziter Einstellungen zu Übergewicht und Homosexualität vorgenommen.

Nach den klassischen Konventionen konvergenter und diskriminanter Validität zu urteilen, ist allein für offen assoziierte und explizite Einstellungen Konstruktvalidität gegeben. Implizite Einstellungen konvergieren nicht hinreichend in einem gemeinsamen Einstellungsfaktor mit expliziten oder offen assoziierten Einstellungen. Dafür können verschiedene Ursachen diskutiert werden, die jedoch zu meist darauf hinweisen, dass implizite Einstellungsmessverfahren wie der IAT oder EAST die psychometrischen Anforderungen interindividueller Differenzmaße nicht ausreichend erfüllen und für eine Einstellungsmessung im Sinne überdauernder Disposition nicht geeignet erscheinen (siehe dazu Abschnitt 2.1.1).

Die resultierenden Einstellungsmaße aus den hier einbezogenen impliziten Messungen, wovon der IAT ein ausgesprochen populäres und weitestgehend etabliertes Verfahren darstellt, werden als Indikatoren relativer kognitiver Assoziationsstärken interpretiert. Diese Maße sollten unbewusste, unreflektierte und unkontrollierbare Kognitionsinhalte reflektieren, die sich durch ein hohes Maß an Stabilität auszeichnen (z.B. Devine, 1989; Wilson, Lindsey & Schooler, 2000). Zum einen können die Annahmen zur Unbewusstheit, Unreflektiertheit und Unkontrollierbarkeit impliziter Maße soweit noch nicht als gesichert gelten (vgl. De Houwer, in press). Zum anderen zeichnen sich vor allem Messungen

---

<sup>4</sup> Hier jedoch wird lediglich der *erste* von vier EAST-Blöcken betrachtet, da der Gesamteffekt gemäß den vier Messblöcken keine befriedigende interne Konsistenz aufweist und der Effekt des ersten Blocks erwartungskongruente Zusammenhänge zu IAT und EAST aufzeigt. Diese *post-hoc*-Auswahl des ersten EAST-Blocks ist ungeachtet der externen Zusammenhänge sehr unbefriedigend und kann nicht zu einer allgemeingültigen, positiven Evaluation des Verfahrens reichen.



*expliziter* Einstellungen durch in aller Regel hohe Stabilitäten aus (siehe unten), so dass eher die Rolle des Bewusstseins als Stabilitätsfaktor für Einstellungsdispositionen diskutiert werden müsste.

Es ist davon auszugehen, dass allein von *einem* Fragebogenitem ausgehend, ein umfangreiches assoziatives Netzwerk aktiviert werden sollte, bevor eine Person dem Grad ihrer Zustimmung bzw. Ablehnung als Resultat eines Entscheidungsprozesses auf einer zweidimensionalen Antwortskala Ausdruck verleiht. Es ist insofern nahe liegend, anzunehmen, dass die Prozesse, die der Beantwortung eines Items zugrunde liegen, komplexer Natur sind und sowohl kognitive, als auch emotionale und motivationale Anteile reflektieren. Über die Items einer Skala zur Messung von Einstellungen hinweg werden somit unterschiedliche Aspekte zum betrachteten Einstellungsobjekt erfasst, die aufgrund des stark reduzierten Antwortformats in der Regel eine hohe interne Konsistenz aufweisen. Im Rahmen der Antwortskala hat der Proband die freie Reaktionswahl. Im Vergleich dazu werden in den hier eingesetzten impliziten Messverfahren bildhafte Reize in kurzer Abfolge dargeboten, wobei die Versuchspersonen nur zwei Antwortmöglichkeiten zur Verfügung haben, ohne jedoch eine wirkliche freie Wahl über die Reaktion, denn es ist genau festgelegt, welcher Stimulus, mit welcher Taste zu beantworten ist. Das bedeutet, eine Person ist dazu angehalten, anhand der normativen Antwortvorgaben und der Instruktion, so schnell und so korrekt wie möglich zu reagieren, eine Art Leistungstest abzulegen. Interpretiert wird das Messergebnis jedoch nicht hinsichtlich der tatsächlichen Leistung, z.B. niedrige Reaktionszeiten bei geringen Fehlerraten, sondern als indirekter Einstellungsindikator. Zudem werden die impliziten Maße in aller Regel durch Differenzbildung gewonnen. Ungeachtet der theoretischen Unterscheidung zwischen expliziten und impliziten Einstellungen (z.B. Banaji & Greenwald; 1995; Devine, 1989; Rudman, 2004; Wilson, Lindsey & Schooler, 2000) bestehen also bereits hinsichtlich der Prozesse, auf denen die notwendigen Urteilsbildungen, Antwortentscheidungen bzw. Reaktionen basieren, ebenso große Welten zwischen expliziten und impliziten Einstellungsmessungen, wie letztlich in Bezug auf die statistischen Grundlagen der betrachteten Einstellungsmaße.

Die dritte eingesetzte Messmethode in dieser Arbeit wird als offene Assoziativmessung bezeichnet und sollte Aspekte expliziter und impliziter Einstellungen erfassen. Es werden jedoch lediglich Beziehungen zwischen der offenen Assoziativmessung und dem Selbstbericht verzeichnet, weshalb keine Verortung der Geschichtenmessung zwischen Selbstbericht und IAT bzw. EAST nahe liegt, sondern eher eine Zwischenstellung des Selbstberichtes, der sowohl zu impliziten Messungen, hier hauptsächlich IAT, als auch zur offenen Assoziativmessung in Zusammenhang steht. Dies ist insofern nachvollziehbar, als dass Selbstbericht und IAT beiderseits eine explizite Thematisierung der Thematik vornehmen – der Selbstbericht notwendigerweise in den Itemformulierungen und der IAT in Form der Kategorienlabel. Wogegen Selbstbericht und Geschichtenmessung, bei der bis auf anleitende Fragen zum Strukturieren der Geschichten keine Vorgaben gemacht werden, das wahlfreie, stärker kontrollierbare, verbale Antwortformat gemein haben. Dies betrifft jedoch durchweg messmethodische As-

pekte, die relativ genau zu benennen sind, während die viel kritischeren inhaltlichen Übereinstimmungen bzw. Unterschiede der verschiedenen Einstellungsmessungen ungeklärt bleiben, so lange keine gemeinsame externe Validierung der drei unterschiedlichen Einstellungsmaße erfolgt.

Gemäß vorliegender Arbeit lassen sich über die externe Validität der Einstellungsmaße keine Schlüsse ziehen, es wird lediglich zunächst herausgestellt, dass Konstruktvalidität im konventionellen Sinne für explizite und implizite Einstellungen nicht gegeben ist. Stattdessen legen die Ergebnisse der MTMM-Analyse das Vorherrschen zweier Methodenfaktoren nahe. Das legt nahe, den Dissoziationsansatz von Devine (1989) bzw. Wilson, Lindsey & Schooler (2000) als adäquate Konzeptualisierung expliziter und impliziter Einstellungen anzunehmen. Jedoch kann ein statthafter Konzeptionsansatz nur auf der Basis reliabler und valider Messinstrumente erfolgen. Die Reliabilitätsproblematik erscheint, gemessen an der Häufigkeit bzw. Seltenheit publizierter Reliabilitätsbefunde, für viele psychologische Disziplinen weniger zentral, wobei es eine Grundvoraussetzung für die Untersuchung differentialpsychologischer Fragestellungen darstellt und deshalb in vorliegender Arbeit entsprechendes Augenmerk verdient. Die Validität impliziter Maße wie dem IAT ist nach wie vor nicht hinreichend bestätigt. Zwar liegen einige richtungweisende Befunde vor, dennoch sind immer noch alternative Erklärungsansätze für den IAT nicht ausgeräumt, wie die Hypothese der Reaktionsschwellenanpassung von Brendl, Markman & Messner (2001), die Hinweise auf den Einfluss exekutiver Kontrollprozesse beim IAT (Mierke & Klauer, 2001, 2003) oder die Rolle von Salienzunterschieden, die den IAT-Effekt über das Prinzip der Figur-Grund-Asymmetrie formen (Rothermund & Wentura, 2001, 2004)

Es bleibt demnach zu bedenken, dass Aussagen zur Konstruktvalidität expliziter und impliziter Einstellungen nur maximal so gültig sein können, inwieweit es Reliabilität und Validität der zugrunde liegenden Maße es erlauben. Zu den psychometrischen Eigenschaften der eingesetzten Einstellungsmessverfahren erfolgt nun eine resümierende Ergebnisdiskussion.

## 7.2 Evaluation der impliziten Einstellungsmessungen

Einstellungen sind als stabile, dem Verhalten zugrunde liegend Konstrukte konzipiert worden (z.B. Allport, 1935; Beck, 1976). Die traditionelle Messung von Einstellungen im Form von Selbstberichten unterliegt zwar einer Reihe von Schwächen, wie einleitend aufgezeigt wurde (Abschnitt 1.5), dennoch ist für die Messungen von Einstellungen mit Fragebogen in der Regel eine hohe Zuverlässigkeit z.B. in Form interner Konsistenz oder zeitlicher Stabilität gegeben (z.B. Cunningham, Preacher & Banaji, 2001; Steffens & Buchner, 2003). Im Gegensatz dazu, zeichnen sich die implizite Einstellungsmessungen vielfach durch unbefriedigende Reliabilitäten aus (z.B. Banse, 2001; Hofmann et al., in press; Teige et al., 2004). Ungeachtet des empirisch untersuchten Gegenstandes sind valide und zuverlässige

Messinstrumente jedoch eine Grundvoraussetzung für die Erforschung psychologischer Fragestellungen. Insbesondere bei der Untersuchung interindividueller Unterschiede, kommt den psychometrischen Eigenschaften von Messinstrumenten eine entscheidende Bedeutung zu, da interindividuelle Differenzen in einem Merkmal wie z.B. sozialen Einstellungen auf tatsächlicher Merkmalsvariation beruhen sollten und nicht hauptsächlich auf Messfehlern. Als Messfehler eines Maßes wird alle unsystematische, unvorhersagbare Varianz bezeichnet, wobei häufig offen bleibt, wie unsystematische Messwertvariation inhaltlich zu interpretieren ist. Störquellen einer jeden Messung können sowohl personeninterne, als auch –externe Einflussfaktoren darstellen. Je höher die Messfehlerbehaftung eines Messwertes, desto eingeschränkter ist seine Interpretierbarkeit. Die Voraussetzung psychometrischer Qualität von Messverfahren muss besonders im Bereich impliziter Einstellungsmessung als kritisch bezeichnet werden.

In den letzten beiden Dekaden lag der Fokus der Einstellungsforschung vor allen Dingen auf der adäquaten Erfassung impliziter Einstellungen, von denen man annahm, die „wahren Einstellungen“ von Personen zu reflektieren. Mit der Entwicklung des Impliziten Assoziationstests durch Greenwald, McGhee & Schwartz (1998) verband sich zunächst für viele sicherlich die Hoffnung, implizite Einstellungsmessung würde den Zugang zu unbewussten und unkontrollierbaren Kognitionsinhalte eröffnen, nachdem bereits Generationen von Einstellungsforschern versuchten, mit diversen methodischen und statistischen Werkzeugen, wie z.B. der Bogus-Pipeline-Methode, unverfälschte Einstellungen untersuchen zu können (für eine Übersicht, siehe Webb, 2000). Dem anfänglich euphorisch betriebenen empirischen Einsatz des IAT für diverse psychologische Fragestellungen und der Neu- bzw. Weiterentwicklung impliziter Messverfahren folgte eine Phase des kritischen Hinterfragens und des resümierenden Überblickens der Befundlage. Doch nach wie vor stehen implizite Maße hochgradig im Interessenfokus, kaum ein psychologisches Forschungsfeld bedient sich nicht der Methodik des IAT, was diesem Verfahren schließlich eine beispielhafte Erfolgsgeschichte verleiht. Vor allem in den vergangenen beiden Jahren gewann jedoch auch die kritische Evaluation des IAT Aufmerksamkeit.

Die implizite Einstellungsmessung allgemein betreffen problematische psychometrische Eigenschaften der Messverfahren, da nicht zuletzt die Reliabilität impliziter Maße als Ursache für fehlende Korrespondenzen zwischen expliziten und impliziten Einstellungen, die mehr oder weniger als interindividuell variierende Dispositionen konzipiert sind, diskutiert wird (vgl. Fazio & Olson, 2003). Für die impliziten Messverfahren lassen sich interne Konsistenzen von zufrieden stellend (IAT) bis völlig unzureichend (EAST) aufführen (vgl. Cunningham, Preacher & Banaji, 2001; De Houwer, 2003a; De Houwer & De Bruycker, 2005; Degner & Wentura, 2005; Ellwart, Rinck & Becker, 2005; Teige et al., 2004). Bisherigen Befunden zur Stabilität impliziter Maße nach, ist von recht geringen Test-Retest-Reliabilitäten auszugehen (Hofman et al., 2003). Die allgemein mangelnde Zuverlässigkeit ist umso kritischer, als dass Reliabilität eine Voraussetzung für die Validierung von Messinstrumenten ist. Hierbei

ist jedoch zu beachten, dass zweifelhaft ist, inwiefern die Gütekriterien der klassischen Testtheorie auf Reaktionszeitparadigmen wie den EAST anzuwenden sind, da nicht davon auszugehen ist, dass sich die systematische Varianz impliziter Maße in der Kongruenz von Reaktionszeit- oder Fehlereffekten der ersten und zweiten Experimentalhälfte ausdrückt. Insofern ist jedoch allenfalls von einer Unterschätzung zumindest der internen Konsistenz auszugehen, was auf die externen Zusammenhänge der impliziten Maße keinen Einfluss hat.

In vorliegender Arbeit können hauptsächlich über die *interne Konsistenz* der Einstellungsmessungen Aussagen getroffen werden. Dabei erweist sich der IAT erneut als zufrieden stellend bis sehr gut reliabel. Der EAST dagegen erfüllt *nicht* die Konsistenzvoraussetzung zur Untersuchung interindividueller Differenzen (Abschnitt. 7.2.1). Die *Stabilität* als spezieller Reliabilitätsaspekt ist zwar für vorliegende Arbeit weniger zentral, jedoch ist der Gesichtspunkt situativer Variabilität aufgrund des Aufbaus und der Ergebnisse von Studie III ebenfalls zu berücksichtigen (Abschnitt. 7.2.2). Schließlich soll die allgemeine Problematik von Differenzmaßen hinsichtlich der Reliabilität der impliziten Einstellungsmessungen diskutiert werden (Abschnitt. 7.2.3).

### 7.2.1 Interne Konsistenz

Für die empirische Untersuchung interindividueller Unterschiede sind prozedurale Variationen im Versuchsaufbau zur „Ausmittlung“ von Reihenfolgeeffekten und Auslegung von Haupteffekten nicht angeraten, da dies eine Interpretation von Unterschieden zwischen Personen erschwert. Aus diesem Grund, wurden die Bedingungen hinsichtlich der Erhebungsprozedur für alle Probanden weitestgehend konstant gehalten. Die Probanden absolvierten entsprechend völlig identische Experimente im Rahmen einheitlicher Untersuchungsdurchführungen. Ziel war es dabei, eine Steigerung der Reliabilität insbesondere für den EAST zu erreichen, da hier gravierende Reliabilitätsmängel bestehen (z.B. Teige et al., 2004).

Der IAT ist – in Übereinstimmung mit der Metaanalyse von Hofmann et al. (2003), die für den IAT eine stark variierende, im Mittel .80 betragende interne Konsistenz ermitteln – als relativ zuverlässig zu bezeichnen. In vorliegender Arbeit werden Reliabilitäten von .87 (Studie I, IAT Homosexualität) bis zu .60 (Studie II, IAT Übergewicht) für die betrachteten IAT-Maße (Reaktionszeiten) verzeichnet.

Für den EAST wird dagegen in Studie I und III eine äußerst unzureichende Zuverlässigkeit festgestellt, weshalb die EAST-Maße als ungeeignet zur Untersuchung interindividueller Differenzen in impliziten Einstellungen bewertet werden müssen. Dieser Befund steht in Kongruenz zu den zitierten EAST-Studien und den bislang noch unpublizierten empirischen Untersuchungen zum EAST (z.B. De Houwer & De Bruycker, 2005; De Jong & Huijding, 2005; Degner & Wentura, 2005; Ellwart,

Rinck & Becker, 2005). Zwei von drei bislang veröffentlichten EAST-Studien berichten sehr unbefriedigende Reliabilitäten. Für ihre Studie zur Messung alkoholbezogener Kognitionen mit IAT und EAST dokumentieren De Houwer et al. (2004) keine Reliabilitäten, wohl aber argumentieren sie hinsichtlich der Nullkorrelationen zwischen IAT und EAST mit der wohl bekannten Reliabilitätsproblematik impliziter Maße bzw. den strukturellen Unterschieden zwischen verschiedenen impliziten Messungen. Zwar konnte in Studie I die Bedeutung eines Teileffekts des EAST (erster Block) herausgestellt werden, nichtsdestotrotz steht der Gesamteffekt weder in Zusammenhang zum IAT, noch zu expliziten Einstellungen bzw. Einstellungsprädiktoren. Eine aufgrund der korrelativen Ergebnisse im nachhinein getroffene Auswahl des ersten EAST-Blocks zur Gegenüberstellung mit dem IAT und Selbstbericht kann *nicht* die Ansprüche wissenschaftlich-empirischen Arbeitens erfüllen. Sofern keine sinnvollen und ad hoc festgelegten Kriterien zur Auswahl eines spezifischen Teileffekts vorliegen, muss der Gesamteffekt als EAST-Maß angenommen werden. Darüber hinaus konnte die in Studie I aufgezeigte Bedeutung des reaktionszeitbasierten Effektes gemäß erstem EAST-Block in Studie III nicht repliziert werden, was dafür spricht, dass der EAST noch kein stabiles Messverfahren darstellt. Eine mögliche Ursache für die geringe interne Konsistenz des EAST ist das niedrige Aggregationsniveau der Effektmaße (vgl. De Houwer, 2003a; Fazio & Olson, 2003; Teige et al., 2004). In den empirische Studien vorliegender Arbeit basiert der IAT-Effekt beispielsweise auf insgesamt 96 Trials (jeweils 48 Trials in kompatibelem und inkompatiblen Block). Eine vergleichbare Steigerung des Aggregationsniveaus für den EAST würde die Darbietung von insgesamt 192 Trials in einem Block bedeuten. Auch wenn diese Stimulusanzahl auf zwei Blöcke aufgeteilt werden würde, müssten die Probanden ohne Unterbrechung 96 Stimuli beantwortet. Zudem zeichnen sich die EAST-Blöcke durch teilweise sogar gegensätzliche Haupteffekte und sehr inkonsistente Zusammenhänge aus, weswegen es nicht im Sinne einer hohen internen Konsistenz ist, die notwendige Trialanzahl auf zwei Blöcke zu verteilen. Möglicherweise sind die fehlenden oder entgegen gesetzt gerichteten Beziehungen zwischen den Blockeffekten auf interindividuelle Veränderungen der Bearbeitungsstrategie (z.B. schnell vs. korrekt) zurückzuführen. Dies sollte während der kontinuierlichen Stimulusdarbietung eines Blocks nicht möglich sein. Der durchweg festgestellte Befund interner *Inkonsistenz* der Blöcke eines EASTs mag jedoch auch für die Stimuli innerhalb eines Blocks gelten, wie die Korrelationsanalysen mit den zielkonzept-spezifischen EAST-Effekten (z.B. Homosexuell vs. Heterosexuell) nahe legen. Dies kann allerdings nur als Hinweis gelten, da stimulusspezifische Effekte nicht notwendigerweise in (gleich gerichteter) Beziehung stehen müssen. In der getrennten Betrachtung von Effekten auf der Basis der Zielkonzeptkategorien (z.B. EAST\_Homo vs. EAST\_Hetero) liegt deshalb ein Vorteil des EAST gegenüber dem IAT (De Houwer, 2003a; Teige et al., 2004). Ein tatsächlicher Nutzen der Analyse zielkonzept-spezifischer EAST-Effekte ist jedoch Post-Hoc-Analysen zufolge für keine der betreffenden Studien dieser Arbeit zu verzeichnen. Die Beziehungsmuster zwischen den spezifischen EAST-Effekten und zum

Beispiel expliziten Einstellungen unterscheiden sich vernachlässigbar geringfügig von denen der allgemeinen EAST-Effekte.

Es bleiben neben der prozeduralen Variation von beispielsweise Stimulusabfolge oder Tastenbelegung und dem geringen Aggregationsniveau zwei weitere, von Teige et al. (2004) angeführte Ursachen für die unbefriedigende Reliabilität des EAST. Diese sind zum einen eine geringe interindividuelle Varianz der Versuchspersonen in den Einstellungen zu den betrachteten Einstellungsobjekten und zum anderen die mangelnde Repräsentativität der Stimuli. In den beiden durchgeführten Untersuchungen (Studie I und II) werden, gemessen an den Mittelwerten sogar recht hohe Varianzen der impliziten Maße erhalten. Mangelnde interindividuelle Einstellungsvarianz zeichnet sich zum Vergleich auch nicht in den Fragebogenmessungen ab. Wie einleitend bereits angedeutet wurde, ist die Repräsentativität des Stimulusmaterials für den EAST kritischer, als für den IAT (vgl. De Houwer, 2003a), weil die Erfassung kognitiver Assoziationsstärken mit dem IAT eher auf der Ebene von Kategorienlabel stattfindet (Karpinski & Hilton, 2001; Nosek, Banaji & Greenwald, 2002). Im IAT ist prinzipiell jeder Stimulus in eine von zwei gegensätzlichen Kategorien zu ordnen, z.B. Homosexuell vs. Heterosexuell, positiv vs. negativ, wobei die Kompatibilität der gemeinsam eine Antworttaste belegenden Kategorien variiert wird (z.B. Homosexuell oder positiv). Deshalb werden hier komplette Blöcke miteinander verrechnet. Der EAST dagegen unterscheidet sich diesbezüglich in 2 Punkten vom IAT. Erstens werden für den EAST-Effekt nicht alle Stimuli eines Blocks betrachtet, sondern nur die Zielkonzeptstimuli. Zweitens sind die Zielkonzeptstimuli im EAST nicht aufgrund ihrer konzeptuellen Bedeutung, sondern auf der Basis unterschiedlicher perzeptuelle Eigenschaften (blau vs. grün), die extrinsisch valenzbesetzt sind, zu beantworten. Inwiefern es einem Probanden z.B. schwer fällt, ein homosexuelles Paarbild auf der Basis der extrinsisch positiven Farbe richtig mit der „guten Taste“ zu beantworten, könnte demnach sowohl an der individuellen Inkompatibilität der Kombination Homosexualität und positiv liegen, als auch daran, dass die dargestellten Personen vom Probanden als unsympathisch oder unattraktiv empfunden werden. Deshalb ist neben der Repräsentativität auch das Matchen der verwendeten Stimuli hinsichtlich verschiedener Merkmale wie Alter, Attraktivität, Sympathie u.a. von entscheidender Bedeutung. Bei gravierender und konsistent ausgerichteter Unterschiedlichkeit des Stimulusmaterials in Bezug auf wertigkeitsrelevante Aspekte, wie z.B. die Attraktivität, spiegelt, ganz gleich ob IAT oder EAST, das Differenzenmaß aus kompatiblen und inkompatiblen Trials keine einstellungsrelevanten Assoziationen wider, sondern schlichtweg den optischen Habitus der Stimuli (De Houwer, 2003a). Damit ist zudem eine schwerwiegende Validitätseinschränkung verbunden (Rothermund & Wentura, 2001).

Die hier verwendeten Stimuli für die Kategorien Homosexualität und Heterosexualität sowie Übergewicht und Normalgewicht wurden in Vorstudien auf Prototypikalität, Attraktivität und Sympathie getestet und entsprechend ausbalanciert. Dies sollte eigentlich sichern, dass die Stimuli nicht nur reprä-

sentativ für die Kategorien sind, sondern im Mittel auch ausgeglichen hinsichtlich anderer wertigkeitsrelevanter Aspekte. Gegeben, dass nichtsdestotrotz die Stimuli Ursache der geringen Reliabilität des EAST in vorliegender Arbeit sind, ist der EAST als Messung interindividueller Differenzen in kognitiven Assoziationsstärken mehr als anzuzweifeln: Ein Effektmaß, das derart sensibel auf solche interindividuelle Unterschiede in der Wahrnehmung inhaltlich irrelevanter Merkmale der, nach allen Regeln der Kunst ausgewählten, Stimuli reagiert, gewährleistet keine Objektivität und muss als sehr fragwürdig eingeschätzt werden.

### 7.2.2 Stabilität

Gemäß der Definition impliziter Kognitionen durch Banaji & Greenwald (1995; vgl. Abb. 1-1) und den Modellen von Devine (1989) oder Wilson, Lindsey & Schooler (2000) ist von einer vergleichsweise *hohen* Stabilität impliziter Einstellungen auszugehen. Implizite Einstellungen seien, wenn überhaupt, dann nur über einen längeren Zeitraum zu verändern. Entsprechend dieser Konzeptionalisierung sollten implizite Einstellungsmessungen eine hohe zeitliche Stabilität aufweisen. Dies ist jedoch für die derzeit verfügbaren impliziten Messverfahren nicht gegeben. Für die am häufigsten eingesetzten Instrumente wie den IAT oder das Affektive Priming werden alles andere als eine hohe Test-Retest-Reliabilitäten berichtet. Cunningham, Preacher & Banaji (2001) setzen mit zweiwöchigem Abstand insgesamt viermal einen Standard-IAT, einen Antwortfenster-IAT und ein Affektive Primings mit Antwortfenster zur Messung von impliziten Einstellungen gegenüber Farbigen relativ zu Weißen ein. Sie erhalten einen mittleren manifesten Stabilitätskoeffizienten von lediglich .27, können jedoch mit latenter Strukturgleichungsmodellierung für die entsprechenden messfehlerbereinigten impliziten Variablen Stabilitäten von .68 (Standard-IAT), .60 (Antwortfenster-IAT) und .83 (Antwortfenster-Priming) nachweisen. Steffens & Buchner (2003) gehen nach dem gleichen Schema wie Cunningham, Preacher & Banaji (2001) vor berechnen Stabilitäten für einen IAT Homosexualität von .56 (einwöchiges Intervall) und .61 (zehnminütiges Intervall). Für den Selbstbericht werden im Vergleich dazu Stabilitäten von .87 und .92 dokumentiert. Auch für die Messung von Ängstlichkeit zeigen Schmukle & Egloff (2004), dass der IAT mit einer Test-Retest-Reliabilität von .50 deutlich unter der Stabilität des Selbstberichtes von .85 liegt. Die Autoren verweisen auf andere eigene Untersuchungen, deren zufolge sie eine Stabilität von .58 für ein einwöchiges Intervall, von .62 für einen Monat und von .47 für einen Zeitraum von einem Jahr für den IAT ermittelten (Egloff, Schwerdtfeger & Schmukle, 2005).

Die Befundlage zusammenfassend ist für den IAT eine im Mittel eine geringere Stabilität bei hoher interner Konsistenz zu vermerken (Hofmann et al., 2003). Es scheint also nahe liegend, von einer relativ großen Anfälligkeit des IAT für Situationsvariabilität auszugehen, was angesichts der Definition impli-

ziter Einstellungen bereits den Validitätsanspruch dieses Verfahrens unmissverständlich hinterfragen lässt. Da die Messung der impliziten und expliziten Einstellungen in Studie III nicht zu einem Zeitpunkt erfolgten, sondern IAT und EAST Übergewicht einen einwöchigen Abstand zum IAT und EAST Homosexualität sowie zum Selbstbericht (Übergewicht und Homosexualität) aufweisen, müsste die vorgenommene Konstruktvalidierung expliziter und impliziter Einstellungen zusätzlich zu Trait-, Methoden- und Messfehlervarianz die Effekte situativer Einflüsse berücksichtigen. Dies ist jedoch aufgrund des Erhebungsdesigns leider nicht möglich. Zudem weist der IAT Übergewicht im Vergleich zum IAT Homosexualität in Studie I und III eine relativ geringe interne Konsistenz auf. Erklärungsansätze für die verhältnismäßig geringe Reliabilität des IAT Übergewicht müssen methodische, prozedurale und inhaltliche Fragen aufwerfen. Der methodische Aspekt des verwendeten Stimulusmaterials für die beiden Zielkonzeptkategorien Übergewicht und Homosexualität wurde bereits im vorangegangenen Abschnitt angeführt. Da alle verwendeten Bildstimuli vorgetestet und hinsichtlich wertigkeitsrelevanter Merkmale gematcht wurden und zudem der IAT-Effekt relativ robust gegen die präsentierten Einzelstimuli ist (vgl. Steffens, Lichau & Still, 2004), wird diesem möglichen Erklärungsansatz weniger Gewicht beigemessen. Zwar konnten Materialeffekte auf der Ebene der Stimuli nachgewiesen werden (z.B. De Houwer, 2001), jedoch wurden dazu Extrembeispiele von Stimuluskombinationen verwendet, was den Bedingungen der hier berichteten Untersuchung keineswegs nahe kommt. Ein prozeduraler Unterschied zwischen IAT Übergewicht und IAT Homosexualität ergibt sich aus der relativen Unerfahrenheit der Versuchspersonen bezüglich des IAT und der Laborsituation. Zumindest zum ersten Messzeitpunkt ist die Stichprobe hinsichtlich der Erfahrungen mit Laborsituationen im allgemeinen und dem IAT als speziellen Messverfahren möglicherweise sehr heterogen. So haben einige Versuchspersonen eventuell noch nie zuvor einen IAT absolviert, während einige aufgrund der Teilnahme an anderen Versuchen bereits vertraut mit dem Verfahren sind. Um Lerneffekte als Ursache für die Reliabilitätsschwankungen untersuchen zu können, hätte die Reihenfolge der Einstellungsobjekte bei der Erhebung ausbalanciert werden müssen, was wiederum für die messfehlerberücksichtigende Analyse interindividueller Differenzen eine Erhöhung der Probandenzahl erforderlich gemacht hätte. Eine eher inhaltlich ausgerichtete Erklärungsmöglichkeit zur geringen Reliabilität des IAT Übergewicht im Vergleich zum IAT Homosexualität betrifft die Einstellungsobjekte selbst. Es ist durchaus plausibel, Unterschiede zwischen den Einstellungen zu Homosexualität und Übergewicht anzunehmen, die sich mit der unterschiedlichen Selbstidentifikation im Hinblick auf beide Stereotypen begründen lassen. Aus unterschiedlich stabilen Selbstidentifikationen hinsichtlich sexueller Orientierung vs. Körperumfang sollten entsprechend unterschiedlich konsistente Einstellungen resultieren (siehe dazu Abschnitt 7.4).

Es konnte gezeigt werden, dass der IAT – obwohl relativ unanfällig gegenüber bewusster Verfälschungsversuche (z.B. Banse, Seise & Zerbes; Egloff & Schmukle, 2002; Steffens, 2004) – sehr sensi-



bel für einfache kontextuelle Variationen ist, wie z.B. die Aktivierung stereotypenkonträrer Kognitionen oder das „Primen“ der Versuchspersonen mit positiv bewerteten farbigen Personen (z.B. Blair, Ma & Lenton, 2001; Karpinski & Hilton, 2001; Wittenbrink, Judd & Park, 2001). Nichtsdestotrotz sind auch stabile Verhaltensvorhersagen für IAT-Maße erbracht worden (Hugenberg & Bodenhausen, 2004), die zeigen, dass der IAT nicht allein situativ bedingte Assoziationen erfasst.

Zusammenfassend bleibt festzustellen, dass, sollte die Situationsanfälligkeit des IAT Ursache für die fehlende Korrespondenz zwischen impliziten und expliziten Einstellungen sein, kann der IAT als Messinstrument interindividueller Differenzen in impliziten Einstellungshaltungen nicht empfohlen werden. Entweder ist die Konzeptualisierung von impliziten Einstellungen als die stabileren Dispositionen im Vergleich zu expliziten Einstellungen nicht zulässig, oder die zur Messung dieser impliziten Einstellungen eingesetzten Verfahren sind nicht adäquat.

### 7.2.3 Aspekte der Reliabilitätsberechnung

Abschließend ist der Diskussion zur Reliabilitätsproblematik impliziter Einstellungsmaße noch ein eher statistischer Aspekt beizufügen. Die hier betrachteten impliziten Maße sind Differenzmaße, weshalb allgemein nochmals auf die Problematik von Differenzmaßen hinsichtlich der Reliabilität hinzuweisen ist. Das prinzipielle und hauptsächlich aus der Veränderungsmessung bekannte Problem von Differenzwerten basiert auf dem Effekt der Regression zur Mitte und resultiert insbesondere aus der hohen Messfehlerbehaftung von Extremwerten (Trautner, 1992). Zusätzlich ist für Veränderungsmessung das Reliabilitäts-Validitäts-Dilemma zu berücksichtigen (Collins, 1991; Rost, 2004): Stehen beide Ausgangswerte in enger Beziehung, ergibt sich für das Differenzmaß eine geringe Reliabilität aufgrund der Varianzeinschränkung, je niedriger dagegen der Zusammenhang zwischen beiden Ausgangswerten ausfällt, desto geringer ist ihre Validität. Insbesondere ersteres ist für vorliegenden Fall von Bedeutung, da die Mittelwerte der Reaktionszeiten und Fehlerraten von inkompatiblen und kompatiblen Blöcken (IAT) bzw. Stimuli (EAST) sehr hoch miteinander korrelieren. Im Mittel beträgt die Korrelation der Reaktionsgeschwindigkeit über die impliziten Messungen und die Einstellungsobjekte hinweg  $.67$  ( $p < .001$ ). Die mittlere Korrelation der Fehlerraten beträgt entsprechend  $.50$  ( $r < .001$ ). In diesen Korrelationen spiegeln sich die interindividuellen Unterschiede in der kognitiven Verarbeitungsgeschwindigkeit und Aspekte fluider Intelligenz wider.

Die Kritik an Differenzmaßen hinsichtlich des statistischen Regressionsartefakts halten Rogosa, Brandt & Zimowski (1982) jedoch angesichts der Bedeutung und der Vorteile von Differenzmaßen für übertrieben. Auch zeigt Zimmerman (1996), dass die allgemein diskutierte geringe Reliabilität von Differenzen hauptsächlich auf methodisch-theoretischen Annahmen beruht und nicht notwendiger-

weise statthaft ist. Darüber hinaus handelt es sich bei impliziten Maßen nicht um Veränderungsmessung im klassischen Sinne, vielmehr *ist* die Differenz zwischen inkompatibler und kompatibler Kombination das implizite Maß. Die Reliabilität des IAT-Effekts, der immerhin auf der Differenz der Mittelwerte (Reaktionszeiten, Fehlerraten) zweier voneinander getrennter Messblöcke basiert, ist zufriedenstellend.

Der EAST-Effekt dagegen stellt das Ausmaß der Inkompatibilität von Trials innerhalb *eines* Blockes dar, wobei hier vor allem Task-Switch- und Response-Switch-Effekte zu tragen kommen können (Schmitz, Teige, Voss & Klauer, 2005). Neben der theoretischen Fundamentalisierung des EAST ist nach wie vor noch nicht geklärt, inwiefern solche, für die eigentliche Messintention irrelevanten Effekte zum letztendlichen EAST-Maß beitragen und wie eine adäquate Reliabilitätsberechnung vorzunehmen ist (vgl. Hoffmann, Kiesel & Sebold, 2003; Klauer, Mierke & Musch, 2003). In diesem Zusammenhang soll darauf hingewiesen werden, dass im Rahmen eines kürzlich stattgefundenen EAST-Workshops (08. – 10. April, 2005, Oppurg) mögliche Alternativen für die Berechnung der internen Konsistenz des EAST diskutiert worden sind. Da stellenweise höhere Test-Retest-Reliabilitäten, als interne Konsistenzen gemäß Cronbach's Alpha oder Split-Half-Reliabilität gefunden wurden (Ellwart, Rinck & Becker, 2005), erscheint eine Auseinandersetzung mit der adäquaten Bestimmung der psychometrischen Eigenschaften von reaktionszeit- bzw. fehlerbasierten Effektmaßen notwendig.

Möglicherweise sind multinomiale Schätzmodelle zur Unterscheidung kontrollierter, automatisierter und zufälliger Rateprozesse im EAST die adäquatere Methode zur Berechnung eines reliablen und validen EAST-Effekts (Stahl & Degner, 2005). Diese Auswertungsmethodik ist im Hinblick auf die impliziten Einstellungsmaße relativ unerforscht und fokussiert derzeit noch ausschließlich auf Fehlerdaten. Für den IAT wurde ein multinomiales Modell, das vier verschiedene Prozesse berücksichtigt, bereits erfolgreich getestet (Conrey, Sherman, Gawronski, Hugenberg & Groom, 2004, in press).

### 7.3 Evaluation der offenen Assoziativmessung

Mit der offenen Assoziativmessung ist im Rahmen dieser Arbeit ein Verfahren zur Erfassung expliziter und impliziter Einstellungen entwickelt worden, was einer ersten Evaluation zufolge als überaus viel versprechend bewertet wird. Die Messung beruht auf der Methodik des Thematischen Apperzeptionstests, bei der Probanden instruiert werden, eine eigene Geschichte zu vorgegebenen Bildmotiven zu verfassen. Hinsichtlich der Bildmotive sowie der Kodierschemata zur Erfassung der Geschichteninformationen wurde eine Anpassung der Methodik für die Messung von Einstellungen, Vorurteilen zu Übergewicht und Homosexualität sowie relevanten Motiven vorgenommen. Für eine Beurteilung des Messverfahrens zur Erfassung interindividueller Einstellungsdifferenzen sollten psychometrische Gesichtspunkte der Kodierung, Validitätsbefunde hinsichtlich expliziter und impliziter Einstellungs-

und Motivationsanteile und Aspekte der Handhabung, wie z.B. die Ökonomie, berücksichtigt werden. Hinsichtlich der psychometrischen Eigenschaften zeichnen sich die Kodierschemata der Assoziativmessung durch eine sehr hohe Interraterreliabilität der entsprechenden Variablen aus. Das bedeutet, die fokussierten Geschichtenmerkmale wie Valenz oder Vorurteilsgehalt sind relativ objektiv erfassbar. Darüber hinaus werden zwischen den Geschichtenvariablen und den selbstberichteten Einstellungen zu Übergewicht bzw. Homosexualität stabile und sinnvolle Beziehungen registriert, trotz dass es sich um methodisch deutlich voneinander abweichende Verfahren handelt. Hier sind jedoch zunächst allgemeine Unterschiede in den Zusammenhangsmustern der Geschichtenmessung und der expliziten Einstellungen in Abhängigkeit vom Einstellungsobjekt anzunehmen. In Studie III sind insbesondere enge Beziehungen zwischen den Geschichtenvariablen und des selbstberichteten Einstellungen für Übergewicht zu berichten, während für das Einstellungsobjekt Homosexualität weniger signifikante Zusammenhänge zwischen beiden Messmethoden bestehen. Dennoch weisen die Korrelationen zwischen offener Assoziativmessung und dem Selbstbericht ungeachtet des Einstellungsobjektes auf einen substanziellen Zusammenhang hin.

Ein relativ konsistenter Befund in diesem Rahmen ist die negative Korrelation zwischen der Geschichtenlänge und den selbstberichteten Einstellungen, die in beiden hierfür relevanten Studien dieser Arbeit aufgezeigt werden konnten. Für beide Einstellungsobjekte gehen negativere Selbstberichte mit kürzeren Geschichten einher. Dieser Zusammenhang könnte zum einen interindividuelle Unterschiede in den Einstellungen zu Übergewicht und Homosexualität reflektieren, möglicherweise zeichnen sich hierin jedoch eher motivationale oder fähigkeitsbezogene Personenmerkmale ab, wie etwaige Selbstpräsentationsmotive oder (verbale) Intelligenz. Zum einen könnte sich eine motivierte Bereitschaft zur Auseinandersetzung mit den Einstellungsobjekten in einem positiveren Selbstbericht und in umfangreicheren Geschichten äußern. Zum anderen ist anzunehmen, dass sich (verbal) intelligente Personen durch ein besseres Verständnis/ eine höhere Kontrollierbarkeit von Fragebogen-Items und eine vergleichsweise komplexere und vielseitigere Ausdrucksweise auszeichnen. Entsprechend zeichnet sich Post-Hoc-Analysen<sup>5</sup> zufolge insbesondere für die Motivation zu Vorurteilsfreiheit eine Bedeutung für den Zusammenhang zwischen Wortanzahl und expliziten Einstellungen. Die statistische Kontrolle von sozialer Erwünschtheit oder der verbalen Intelligenz haben einen vergleichbar unbedeutenden Einfluss in diesem Zusammenhang. Der ebenfalls stabile positive Zusammenhang zwischen dem Vorurteilsgehalt der Geschichten und den expliziten Einstellungen (Studie III) konnte dagegen als relativ unabhängig von sozialer Erwünschtheit bzw. der Motivation zur Vorurteilsfreiheit herausgestellt werden. Sollten weitere Untersuchungen eine stabile Replizierbarkeit dieser Beziehung

---

<sup>5</sup> Aufgrund der Verfügbarkeit entsprechender Kontrollvariablen wie Intelligenz, soziale Erwünschtheit und Motivation zur Vorurteilsfreiheit wurde an dieser Stelle lediglich eine Post-Hoc-Analyse für die Daten aus Studie III durchgeführt. Dabei muss jedoch berücksichtigt werden, dass hier keine signifikanten Zusammenhänge zwischen Geschichtenlänge und expliziten Einstellungen für das Einstellungsobjekt Homosexualität bestehen. Eine Verallgemeinerung der Befunde ist also nicht möglich.

auch unter Kontrolle weiterer möglicher Drittvariablen erbringen, kann eine zugrunde liegende Einstellungsdisposition als das gemeinsame Substrat des Zusammenhangs zwischen Geschichtenmessung und Selbstbericht angenommen werden.

Die Validität der offenen Assoziativmessung hinsichtlich *impliziter* Einstellungen ist weniger eindeutig zu bewerten. Zwischen den Geschichtenvariablen und den impliziten Einstellungsmaßen bestehen zunächst augenfällig weniger substanzielle Zusammenhänge, als für den Selbstbericht. Dies kann einerseits als Hinweis auf die relative Unabhängigkeit von Geschichtenmessung und impliziten Einstellungen betrachtet werden, andererseits sind jedoch eine Reihe weiterer möglicher Einflussfaktoren zu bedenken. In dieser Arbeit sind es, wie bereits erwähnt, insbesondere die Übergewichtsgeschichten, die in verhältnismäßig engem Zusammenhang zu expliziten Einstellungen zu Übergewicht stehen. Da jedoch die impliziten Messungen zu Übergewicht relativ problematisch hinsichtlich ihrer psychometrischen Eigenschaften (geringe Reliabilität) bzw. ihrer Interpretierbarkeit (situative Variabilität vs. problematische Einstellungsinhalt) sind, ist die Divergenz von Geschichtenmessung und impliziten Einstellungen zu Übergewicht nur begrenzt inhaltlich interpretierbar.

Für das Einstellungsobjekt Homosexualität spielt vor allem die Variable der Stereotyperkennung bzw. -thematisierung eine entscheidende Rolle im Zusammenhang zu impliziten Einstellungen. Die entsprechenden Korrelationen zu IAT und EAST weisen in die Richtung höherer impliziter Ablehnung von Homosexualität bei Personen, die den Stereotyp Homosexualität in den Geschichten nicht erkannt bzw. nicht deutlich thematisiert haben. Bei der betreffenden Geschichtenvariablen ist fraglich, ob es sich um ein Nichterkennen oder ein Nichtansprechen im Sinne von Vermeidung handelt, aufgrund der signifikant negativen Zusammenhänge zu beiden impliziten Maßen liegt jedoch eine mehr oder weniger bewusste Vermeidung der Homosexualitätsthematik näher.

An dieser Stelle ist als kritischer Aspekt die Abhängigkeit der offenen Assoziativmessung von der Darstellbarkeit des Einstellungsobjekts in einem schematisch-mehrdeutigem Bildmotiv anzuführen. Bei grafisch schwer fassbaren Einstellungsobjekten besteht die Gefahr einer Fehlinterpretation des Bildmotivs und somit einer inhaltlich wertlosen Geschichtenmessung. Eine missglückte Stereotypaktivierung könnte demnach prinzipiell Ursache für fehlende Zusammenhänge zwischen Geschichtenmessung und anderen Einstellungsmaßen sein. Insofern würde die relative Unmissverständlichkeit von Bildmotiven zum Einstellungsobjekt Übergewicht eine maßgebliche Voraussetzung für die enge Beziehung zwischen Übergewichtsgeschichten und expliziten Einstellungen zu Übergewicht sein. Dagegen könnte für die Homosexualitätsmotive eine Identifikationsproblematik hinsichtlich des thematisierten Einstellungsobjektes vorliegen. Da die Bildmotive neben der Aktivierung des Einstellungsobjektes einen ausreichenden Interpretationsspielraum bieten sollten, bleibt es für die Homosexualitätsmotive der individuellen Auslegung überlassen, zwei enge Freunde, Brüder oder gar Vater und Sohn zu sehen. Hier zeichnet sich ein regelrechtes Dilemma zielsicherer Stereotypaktivierung bei maximaler

Mehrdeutigkeit der Darstellung ab. Eine projektive Methode zur Erfassung bewusster und unbewusster, reflektierter und unreflektierter, überlegter und spontaner Kognitionsinhalte muss jedoch eine entsprechende Projektionsfläche individueller Interpretationen bieten. Die absolut unmissverständliche Darstellung von Homosexualität, z.B. in Form sich küssender Männer, müsste eine starke inhaltliche Einschränkung der Geschichten und ergo eine Varianzeinschränkung der Geschichtenvariablen zur Folge haben. Entsprechende Geschichten könnten inhaltlich wenig mehr widerspiegeln als eine Kusszene. Da in dieser Arbeit jedoch interindividuelle Differenzen im Fokus der Betrachtung liegen, sollte die offene Assoziativmessung sensitiv für interindividuelle Unterschiede in der inhaltlichen Auslegung der Bildmotive sein und entsprechende Variabilität aufweisen. Für das Einstellungsobjekt Homosexualität könnte die um einiges variabelere Interpretierbarkeit der Bildmotive entsprechend Ursache für die geringeren Korrespondenzen zwischen Homosexualitätsgeschichten und expliziten Einstellungen zu Homosexualität sein.

Hinsichtlich Reliabilität und Validität der offenen Assoziativmessung ist soweit zu resümieren, dass die Methodik des Thematischen Apperzeptionstest grundsätzlich erfolgreich auf die Messung von Einstellungen, Vorurteilen und vorurteilsrelevanten Motivationen adaptiert werden konnte. Zwischen dem Vorurteilsgehalt der Geschichten und den expliziten Vorurteilen bestehen substantielle positive Beziehungen, wobei hier bereits erklärbare Unterschiede zwischen beiden Einstellungsobjekten zu verzeichnen sind. Diese werden bei der Betrachtung der Zusammenhänge zwischen den Geschichtenvariablen und den impliziten Assoziationsmaßen deutlicher und können u.a. auf Unterschiede im Grad der bewussten Themenaktivierung und -verarbeitung durch die Bildmotive beider Einstellungsobjekte zurückgeführt werden. Insofern reflektiert die offene Assoziativmessung die Besonderheiten von Einstellungen zu Übergewicht bzw. Homosexualität. Während übergewichtige Personen problemlos aufgrund augenfälliger Merkmale als übergewichtig zu identifizieren sind, müssen sich homosexuelle Menschen regelrecht „outen“, um von der Gesellschaft „erkannt“ zu werden. Dies macht eine Vermeidung der Homosexualitätsthematik für Außenstehende prinzipiell leichter.

Abschließend ist zur Evaluation der offenen Assoziativmessung der Aspekt der Verfahrensökonomie zu berücksichtigen. Zunächst ist die einzurechnende Bearbeitungsdauer der Geschichtenmessung im Vergleich zu Fragebogenmessungen und impliziten Messverfahren deutlich höher, wenn keine Zeitvorgaben festgelegt werden. Entscheidender ist jedoch die qualitative Datenbasis der Assoziativmessung. Für die Untersuchung empirischer Fragestellungen, sind quantitative Inhaltskodierungen des Geschichtenmaterials vorzunehmen, was zur Erhöhung der Objektivität und entsprechend auch der Validität von mindestens zwei unabhängigen Ratern vorgenommen werden sollte. In dieser Hinsicht ist die Auswertung der offenen Assoziativmessung deutlich aufwendiger, als für Fragebogenskalen und computergestützte Experimente. Sollte sich der Zusammenhang zwischen Geschichtenlänge und Einstellungen als stabil replizierbar erweisen, wäre jedoch eine objektive und zudem problemlos zu

kodierende Geschichtenvariable zur Untersuchung interindividueller Einstellungsdifferenzen verfügbar.

Die Aspekte, deren zufolge die offene Assoziativmessung als vergleichsweise weniger ökonomisch zu bewerten ist, bedingen gleichzeitig das erhebliche Potential des Verfahrens. Aufgrund des offenen Antwortformats und der qualitativen Datenbasis ist die Geschichtenmessung hinsichtlich Informationsgehalt und Auswertungsmöglichkeiten zweifellos vielfältiger, als jede Fragebogenmessung oder experimentelle Reaktionszeit- und Fehlerdaten. Die Möglichkeiten des Verfahrens sind im Rahmen vorliegender Arbeit notwendigerweise nur begrenzt ausgeschöpft worden. Speziell in Hinblick auf die Messung vorurteilsbezogener *impliziter* Motive zusätzlich zu selbstberichteten *expliziten* Motiven (z.B. Banse & Gawronski, 2003; Plant & Devine, 1998) erscheint eine weiterführende Untersuchung der offenen Assoziativmessung viel versprechend.

#### 7.4 Die Rolle der ausgewählten Einstellungsobjekte

Für die Untersuchung der Konstruktvalidität von expliziten und impliziten Einstellungen wurden zwei soziale Minderheiten unserer Gesellschaft als Einstellungsobjekte ausgewählt, und zwar übergewichtige und homosexuelle Menschen. Diese beiden Personengruppen sind in vielen Kulturen nachweislich betroffen von Stereotypen, Vorurteilen und Diskriminierung seitens der sozialen Mehrheit (z.B. Crandall & Biernat, 1990; Crandall & Martinez, 1996; Crandall et al., 2001; Grover, Keel & Mitchell, 2003; Rothblum et al., 1990; Hayes, 1995; Hebl, et al., 2002; Herek, 1994; LaMar & Kite, 1998; Steffens & Wagner, 2004).

Ein diesbezüglich kritischer Aspekt aktueller Arbeit bezieht sich auf die Geschlechtsspezifität des Materials der Messinstrumente. Die Stimuli der impliziten Einstellungsmessungen stellen zum einen übergewichtige und normalgewichtige Frauen dar und zum anderen rein männliche und gemischtgeschlechtliche Paare. Während jedoch für das Einstellungsobjekt Homosexualität aus Repräsentativitätsgründen in allen Messungen männliche Homosexuelle fokussiert wurden, liegt eine, das Material betreffende Geschlechtsspezifität für das Einstellungsobjekt Übergewicht nur im Hinblick auf die impliziten Maße vor. Im Selbstbericht ist allgemein von „übergewichtigen Menschen“ bzw. „Dicken“ die Rede. Diese forcierte Selektivität des Stimulusmaterials diente der Vermeidung von ambivalenten Assoziationen zu männlichen und weiblichen Individuen, da sowohl übergewichtige *Frauen*, als auch homosexuelle *Männer* die jeweiligen Prototypen der ausgewählten sozialen Stereotype Homosexualität und Übergewicht darstellen (vgl. Abschnitt 3.2) und insbesondere die impliziten Maße auf Ambivalenzen im Stimulusmaterial sensibel reagieren sollten (z.B. de Liver, Wigboldus & van der Pligt, 2005). Dies ist umso bedeutsamer, derweil das Geschlecht an sich einen sozialen Stereotyp darstellt, zu dem divergente Einstellungen bzw. Assoziationen berichtet werden (z.B. Carpenter & Banaji, 2000; Eagly

& Mladinic, 1989; Sidanius, Pratto & Bobo, 1994; Skowronski & Lawrence, 2001). Eine Alternative zur weitestgehenden Vergleichbarkeit des verwendeten Materials hinsichtlich des Geschlechts wäre die ausdrückliche Fokussierung auf weibliche Übergewichtige im Selbstbericht, wobei dadurch Konfundierung des Stereotyps Übergewicht mit dem Stereotyp Frauen unauflösbar gewesen wäre. Fraglich ist, ob die Fragebogen-Items zu „übergewichtigen Menschen“ und das rein weibliche Stimulusmaterial für IAT- und EAST-Übergewicht derartige inhaltliche Abweichungen in den Einstellungen zur Folge haben können, die eine Erklärung für die mangelnden Korrespondenz zwischen expliziten und impliziten Einstellungen zu Übergewicht liefern. Dieser Frage kann über systematische empirische Untersuchungen nachgegangen werden.

Darüber hinaus ist im Rahmen vorliegender Arbeit nicht hinreichend zu klären, inwieweit Einstellungen gegenüber Übergewicht und gegenüber Homosexualität in ihrer *Konzeptualisierung* miteinander vergleichbar sind bzw. inwiefern fundamentale Unterschiede zwischen beiden Einstellungsrepräsentationen anzunehmen sind, weshalb eine gemeinsame Konstruktvalidierung nicht möglich wäre. Prinzipiell sollte eine Generalisierbarkeit der Konzeptionen über Einstellungsinhalte hinweg bestehen. Spezifischen Eigenschaften der Einstellungsobjekte selbst sind jedoch als Einflussfaktoren für die Beziehung zwischen expliziten und impliziten Einstellungen ausgewiesen worden (z.B. Gschwendner & Schmitt, 2005; Nosek, 2004), sie verdienen deshalb gesonderte Beachtung. Eine entscheidende Variable in vorliegender Arbeit ist entsprechend die *Selbstrelevanz* der Einstellungsobjekte für die Untersuchungsteilnehmer. Darüber hinaus sind an dieser Stelle Einstellungsunterschiede in Abhängigkeit von Personenmerkmalen, wie z.B. dem Geschlecht, zu berücksichtigen. So ist aufgrund der Körperschemamessung in Sinne von körperlichem Real- versus Idealbild die durchschnittliche Präferenz eines relativ schlanken Körpers für die Gesamtstichprobe in Studie III festzustellen, wobei Frauen eine höhere Abweichung zwischen Realität und Wunschvorstellung aufweisen als Männer. Diese allgemeine Tendenz verwundert ebenso wenig wie der spezifische Geschlechtsunterschied, da vor allem für Menschen aus westlichen Industriegesellschaften eine enorme Diskrepanz zwischen dem realen Gewichtsmittelwert der (erwachsenen) Bevölkerung und der medial verbreiteten Idealfigur besteht. Angesichts der konsensualen Präferenz eines extrem schlanken, ja häufig sogar abmagerten Körpers, sind Männer wie Frauen, jüngere wie ältere Menschen gleichsam mit einem zum Teil völlig übersteigerten Anspruch konfrontiert, möglichst dünn zu sein. Entsprechend berücksichtigen Ätiologiemodelle von Essstörungen wie *Anorexia Nervosa* oder *Bulimia Nervosa* den Einfluss des gesellschaftlichen Schlankheitsideals (vgl. Fichter & Quadflieg, 2000; Hänsel, 1995). Es liegt nahe, das Vorherrschen eines mehr oder weniger latenten „Dick-Seins-Gefühls“ bei vielen Menschen anzunehmen, auch wenn laut Body-Mass-Index keine Normabweichung in Richtung Übergewicht oder sogar bereits eine Normabweichung in Richtung Untergewicht vorliegt.

Eine solche irrealer Selbstidentifikation kann für Homosexualität dagegen nicht angenommen werden, was u.a. sicherlich darauf zurückzuführen ist, dass es sich bei der sexuellen Orientierung nicht um ein kontinuierlich verteiltes Merkmal wie das Körpergewicht bzw. der BMI handelt. Eine erwachsene Person kann im Normalfall vermutlich problemlos bestimmen, ob sie heterosexuell, homosexuell oder bisexuell orientiert ist, wobei die Kategorie „heterosexuell“ für die Mehrheit unserer Gesellschaft als Norm gilt, während homo- und bisexuell gleichermaßen deviant sind. Dagegen fällt es nicht schwer anzunehmen, dass viele Personen der Allgemeinbevölkerung der Meinung sind, hier und da nicht ganz dem figürlichen Wunschbild zu entsprechen. Darüber hinaus ist die Wahrscheinlichkeit eines sexuellen Orientierungswechsels relativ gering im Vergleich zu der Möglichkeit, Gewicht zuzunehmen und vielleicht selbst einmal als übergewichtig zu gelten, noch dazu, weil der BMI positiv mit dem Alter korreliert ist, was auch die Verschiebung der oberen und unteren Normschranken für den Normalgewichtsbereich um jeweils einen Punkt etwa alle 10 Jahre nach dem 25. Lebensjahr widerspiegelt. Empirische Hinweise auf die hier vermuteten unterschiedlichen Selbstidentifikationsphänomene der Merkmale Körpergewicht und sexuelle Orientierung liefert der Vergleich von Untersuchungen, denen zufolge homosexuelle Menschen einen Ingroup-Favorisierung aufweisen (z.B. Banse, Seise & Zerbes, 2001; Seise, Banse & Neyer, 2002), wogegen übergewichtige ebenso wie normalgewichtige Menschen negativere Einstellungen zu Übergewicht haben (z.B. Crandall, 1994).

Es kann also vermutet werden, dass Übergewicht für die Mehrzahl der Probanden von Studie III eine deutlich höhere Selbstrelevanz besitzt, als Homosexualität, noch dazu weil der Frauenanteil der Stichprobe relativ hoch ist, und insbesondere für die Frauen die Figur eine große Rolle spielt. Dies sollte berücksichtigt werden, da qualitative Unterschiede zwischen den Einstellungsobjekten Auswirkungen auf die Varianz der resultierenden Einstellungsmaße haben kann, wodurch die Vergleichbarkeit der beiden Einstellungsinhalte Übergewicht und Homosexualität gemindert werden kann. Um die Eigenschaften der Einstellungsobjekte näher bestimmen zu können, ist eine Untersuchung der kognitiven Repräsentationen der Einstellungsinhalte notwendig. Sollten jedoch die spezifischen Einstellungscharakteristika Hauptursache der divergenten Beziehungsmuster für Einstellungen zu Übergewicht im Vergleich zu Homosexualität in Studie III sein, wäre von einer allgemeingültigen, inhaltsunabhängigen Konzeptionalisierung von Einstellungen abzusehen. Nicht dokumentierten Zusammenhangsanalysen zufolge, resultieren aus einer statistischen Kontrolle des Geschlechts jedoch keine nennenswerten Veränderungen der Beziehungsmuster in Studie III.



## 7.5 Zusammenfassung und Ausblick

Mit vorliegender Arbeit wurden die Konzeptionalisierung und die Bedeutung expliziter und impliziter Einstellungen untersucht, woraus eine Reihe möglicher Anknüpfungspunkte für weiterführende Fragestellungen abgeleitet werden konnte. Es wird deutlich, dass die Einstellungsforschung mit der Unterscheidung zwischen bewusst verarbeiteten, kontrollierbaren und automatisierten, unkontrollierbaren Kognitionen einen komplexen Problemraum eröffnete, der nur durch gezielte Untersuchungen einzuengen ist. Der empirische Teil dieser Arbeit liefert Hinweise auf eine anzunehmende Divergenz expliziter und impliziter Einstellungen, indem gezeigt werden konnte, dass der Einfluss der Messmethode die inhaltliche Bedeutung überwiegt. Zudem wurden für explizite und implizite Einstellungen unterschiedliche Beziehungsmuster zu einstellungsrelevanten Konstrukten aufgedeckt. Der Rahmen für eine Generalisierung der Ergebnisse ist jedoch aufgrund verschiedener inhaltlicher und methodischer Aspekte begrenzt. Zum einen wurden in vorliegender Arbeit zwei spezifische Einstellungsobjekte fokussiert, die Relevanz für Vorurteile und soziale Diskriminierung haben. Notwendigerweise bleiben dadurch weitere Einstellungsinhalte unberücksichtigt; darüber hinaus müssen charakteristische inhaltliche Aspekte der ausgewählten Einstellungsobjekte Übergewicht und Homosexualität in Betracht gezogen werden (vgl. Abschnitt 7.4)

Die prozeduralen Aspekte in Studie III erschweren die Vergleichbarkeit der Messungen, zumindest für das Einstellungsobjekt Übergewicht, wobei es sehr fraglich erscheint, inwiefern die Erhebung expliziter und impliziter Maße für zwei oder gar mehr Einstellungsobjekte zu einem Messzeitpunkt realistisch ist bzw. andere unerwünschte Effekte wie Ermüdung oder Überforderung verursacht. Darüber hinaus muss bei der Verallgemeinerung der Befunde die Materialspezifik der impliziten Messungen kritisch berücksichtigt werden (vgl. Abschnitt 7.4).

Die psychometrischen Unzulänglichkeiten des EAST sind kein alleiniges Problem vorliegender Arbeit, wie die wenigen EAST-Publikationen in wissenschaftlichen Fachzeitschriften und die unveröffentlichten Ergebnissen der Teilnehmer am EAST-Workshop (08.-10. April, Oppurg) erkennen lassen. Den geringen Reliabilitäten des EAST zufolge liegen ebenso Einschränkung in der Generalisierbarkeit der Ergebnisse von Studie I und III vor (vgl. Abschnitt 7.2). Dies kann allerdings als eigenständiger Befund gewertet werden, dementsprechend es fraglich erscheint, inwieweit implizite (Einstellungs-) Messung hinreichend Gültigkeit besitzt.

## 8 Literaturverzeichnis

- Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*.
- Aidman, E. V. & Carroll, S. M. (2003). Implicit individual differences: relationships between implicit self-esteem, gender identity, and gender attitudes. *European Journal of Personality*, 17, 19-37.
- Allport, G. W. (1935). Attitudes. In C. Murchison, *A Handbook of Social Psychology* (798-844). Worcester, MA, US: Clark University Press.
- Altemeyer, B. (1981). *Right-Wing Authoritarianism*. Winnipeg: University of Manitoba Press.
- Altemeyer, B. (1988). *Enemies of freedom*. San Francisco: Jossey-Bass.
- Altemeyer, B. (1994). Reducing prejudice in right-wing authoritarians. In M. P. O.Zanna (Ed.), *The psychology of prejudice: The Ontario symposium, Vol. 7* (131-148).
- Altemeyer, B. (1996). *The authoritarian specter*. Cambridge: Harvard University Press.
- Altemeyer, B. (1998). The other "authoritarian personality". *Advances in Experimental Social Psychology*, 30, 47-92.
- Altemeyer and B. Hunsberger (1992). Authoritarianism, Religious Fundamentalism, Quest, and Prejudice, *International Journal for the Psychology of Religion* 2, 113-133.
- Amodio, D. M. H., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality & Social Psychology*, 84, 738-753.
- Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality & Social Psychology*, 84, 738-753.
- Anooshian, L. J. & Seibert, P. S. (1996). Conscious and unconscious retrieval in picture recognition: A framework for exploring gender differences. *Journal of Personality & Social Psychology*, 70(3), 637-645.
- Aronson, E., Wilson, T.D. & Akert, R.M. (2004). Sozialpsychologie und das Gesetz. In *Sozialpsychologie* (4. Aufl., S. 584-615). München: Pearson Studium.
- Asendorpf, J. (2004). *Psychologie der Persönlichkeit* (3. überarb. und aktual. Aufl.). Heidelberg: Springer-Verlag.
- 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 & Social Psychology*, 83, 380-393.
- Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality & Social Psychology*, 81, 789-799.
- Atkinson, J. W. & Feather, N. T. (1966): A theory of achievement motivation. New York: Wiley.
- Bagozzi, R.P. & Yi, Y. (1990). Assessing method variance in multitrait-multimethod matrices, the case of self-reported affect and perceptions at work. *Journal of Applied Psychology*, 75 (5), 547-560.
- Balk, D. E., Lampe, S., Sharpe, B., Schwinn, S., Holen, K., & Cook, L. D. R. I. (1998). TAT results in a longitudinal study of bereaved college students. *Death Studies*, 22, 3-21.
- Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion*, 15(4), 501-520.
- Banse, R. & Fischer, L. (2002). Implicit and explicit aggressiveness and the prediction of aggressive behavior. *Poster präsentiert auf der 11. Konferenz der European Association of Personality Psychology*. Jena.
- Banse, R. & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Psychometrische Eigenschaften und Validität. *Diagnostica*, 49, 4-13.

- Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift fuer Experimentelle Psychologie*, 48, 145-160.
- Bargh, J. A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of Personality & Social Psychology*, 43(3), 425-436.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh: *Unintended thought* (3-51). New York, NY, US: Guilford Press.
- Bargh, J. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer. & T. K. Srull: *Handbook of social cognition, Vol. 1: Basic processes, Vol. 2: Applications* (2nd ed. 1-40). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer: *Automaticity of everyday life: Advances in social cognition*, Vol. 10. Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Bargh, J. A., Chaiken, S., Govender, R. & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality & Social Psychology*, 62(6), 893-912.
- Baron, R. A. & Byrne, D. (2003). *Social Psychology* (10 ed.). Boston MA: Allyn & Bacon.
- Beck, A. T (1976). *Cognitive therapy and the emotional disorders*. Oxford, England: International Universities Press.
- Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88 (3), 588-606.
- Bessenoff, G. R. & Sherman, J. W. (2000). Automatic and controlled components of prejudice toward fat people: Evaluation versus stereotype activation. *Social Cognition*, 18, 329-353.
- Bierhoff, H.W. (1998). *Sozialpsychologie: Ein Lehrbuch* (2. Aufl.). Stuttgart: Kohlhammer.
- Blair, I. V. (2001). Implicit stereotypes and prejudice. In G.B.Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (359-374).
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality & Social Psychology*, 81, 828-841.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality & Social Psychology Review*, 6, 242-261.
- Block, J. & Block, J. (1951). An investigation of the relationship between intolerance of ambiguity and ethnocentrism. *Journal of Personality*, 19, 303-311.
- Boniecki, K. A., Lacks, J., & Zuwerink, J. R. (2002). The elusive relationship between measures of implicit and explicit prejudice. *Representative Research in Social Psychology*, 26, 1-14.
- Bortz, J. (2004). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.) Heidelberg: Springer.
- Boswell, D. A. (1980). Evaluation of transparencies for psychology instruction. *Teaching of Psychology*, 7(3), 171-173.
- 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 & Social Psychology*, 79, 631-643.
- Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology*, 4, 79-101.
- 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 & Social Psychology*, 81, 760-773.
- Brewer, M. B. (1988). A dual process model of impression formation. In, T. K. Srull & R.S. Wyer: *A dual process model of impression formation* (1-36). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Brockhaus Enzyklopädie (2005). Mannheim: Bibliographisches Institut & F. A. Brockhaus AG.
- Buchner, A. & Brandt, M. (2003). Further evidence for systematic reliability differences between explicit and implicit memory tests. *Quarterly Journal of Experimental Psychology: Section A*, 56, 193-209.

- Buchner, A. & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology*, 40(3), 227-259.
- Budner, S. (1962). Intoleranz of ambiguity as a personality variable. *Journal of Personality*, 30, 29-50.
- Busemeyer, J. R. & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 88, 237-244.
- Byrne, B.M. (2001). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Busemeyer, J.R. & Jones, L.E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 93, 549-562.
- Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality & Social Psychology*, 42, 116-131.
- Campbell, J.D. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Carpenter, S. & Banaji, M. R. (2000). Implicit gender attitudes : Group membership, cultural construal, and malleability. unpublished manuscript, Yale University, New Haven.
- Chambliss, H. O., Finley, C. E. & Blair, S. N. (2004). Attitudes toward Obese Individuals among Exercise Science Students. *Medicine & Science in Sports & Exercise*, 36 (3), 468-474
- Cloetta, B. (1983). Der Fragebogen zur Erfassung von Machiavellismus und Konservatismus. *Schweizerische Zeitschrift für Psychologie*, 42 (2/3), 127-159.
- Collins, L. M. (1991). The measurement of dynamic latent variables in longitudinal aging research: Quantifying adult development. *Experimental Aging Research*, 17(1), 13-20.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg K. & Groom, C. (2004, in press). Beyond automaticity and control: The Quad-Model of behavioral responses. Manuscript submitted for publication.
- Corbin, W. R., Bernat, J. A. & Calhoun, K. S. (2001). The role of alcohol expectancies and alcohol consumption among sexually victimized and nonvictimized college women. *Journal of Interpersonal Violence*, 16(4), 297-311.
- Cramer, P. (2001). Identification and Its Relation To Identity Development. *Journal of Personality*, 69, 667-688.
- Cramer, P. (2002). Defense Mechanisms, Behavior, and Affect in Young Adulthood. *Journal of Personality*, 70, 103-126.
- Crandall, C. S. (1994). Prejudice against fat people: Ideology and self-interest. *Journal of Personality & Social Psychology*, 66, 882-894.
- Crandall, C. S. & Biernat, M. (1990). The ideology of anti-fat attitudes. *Journal of Applied Social Psychology*, 20, 227-243.
- Crandall, C. S., D'Anello, S., Sakalli, N., Lazarus, E., Wiczorkowska, G., & Feather, N. T. (2001). An Attribution-Value model of prejudice: Anti-fat attitudes in six nations. *Personality & Social Psychology Bulletin*, 27, 30-37.
- Crandall, C.S., Eshleman, A. & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalisation. *Journal of Personality and Social Psychology*, 82 (3), 359-378.
- Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin*, 87, 546-563.
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit Attitude Measures: Consistency, Stability, and Convergent Validity. *Psychological Science*, 12, 163-170.

Czopp, A. M., Monteith, M. J., Zimmerman, R. S., & Lynam, D. R. (2004). Implicit Attitudes as Potential Protection From Risky Sex: Predicting Condom Use With the IAT. *Basic & Applied Social Psychology*, *26*, 227-236.

Dalbert, C. (1999). The world is more just for me than generally: About the personal belief in a just world scale's validity. *Social Justice Research*, *12*(2), 79-98.

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 & Social Psychology*, *81*, 800-814.

Deci, E. L. & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*, 109-134.

Deci, E. L. & Ryan, R. M. (2000). The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*, 227-268.

DeCharms, R., Morrison, H. W., Reitman, W. R. & McClelland, D.C. (1955). Behavioral correlates of directly and indirectly measured achievement motivation. In D. C. McClelland (Ed.), *Studies in motivation* (414-423). New York: Appleton-Century-Crofts.

De Houwer, J. (in press). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage Publishers.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology*, *37*, 443-451.

De Houwer, J. (2002). The implicit association test as a tool for studying dysfunctional associations in psychopathology. *Behavior Therapy and Experimental Psychiatry*, *53*, 115-133.

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

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

De Houwer, J. (in press). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage Publishers.

De Houwer, J., Crombez, G., Koster, E. H. W., & De Beul, N. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *Journal of Behavior Therapy & Experimental Psychiatry*, *35*, 275-286.

De Houwer, J. & De Bruycker, E. (2005). Do EAST scores provide a measure of interindividual differences in attitudes? *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.

De Houwer, J. & Eelen, P. (1998). An Affective Variant of the Simon Paradigm. *Cognition & Emotion*, *12*, 45-62.

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition & Emotion*, *16*, 643-666.

de Jong, W. (1980). The stigma of obesity: The consequences of naive assumptions concerning the causes of physical deviance. *Journal of Health & Social Behavior*, *21*, 75-87.

de Jong, P. J. (2002). Implicit self-esteem and social anxiety: differential self-favouring effects in high and low anxious individuals. *Behaviour Research & Therapy*, *40*(5), 501-508.

de Jong, P. J. & Huijding, J. (2005). Pictorial vs. verbal stimuli in the EAST: Sensitivity to generally affective and phobia-relevant stimuli in high and low spider fearful individual. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.

de Jong, P. J., Pasma, W. & Kindt, M. (2001). A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research & Therapy*, *39*(1), 101-113.

Degner, J. (in prep.). Prejudice towards obese people. Dissertation an der Universität des Saarlandes in Saarbrücken.

- Degner, J. & Wentura, D. (2005). Assessing intergroup attitudes with the EAST. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- de Liver, Y., Wigboldus, D. & van der Pligt, J. (2005). Ambivalent attitudes and the EAST: the influence of evaluative context. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Denz, H. (1988). Strukturen des Konservatismus. *Angewandte Sozialforschung*, 15, 75-82.
- De Raedt, R., Schacht, R. & De Houwer, J. (2004). Depressed patients have positive implicit self-esteem, submitted for publications.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality & Social Psychology*, 56, 5-18.
- Devine, P.G. (2001). Implicit prejudice and stereotyping : How automatic are they? Introduction to the special section. *Journal of Personality and Social Psychology*, 81 (5), 757-759.
- Devine, P. G. & Baker, S. M. (1991). Measurement of racial stereotype subtyping. *Personality & Social Psychology Bulletin*, 17, 44-50.
- Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality & Social Psychology*, 60, 817-830.
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality & Social Psychology*, 82, 835-848.
- Dollinger, S. J. (2003). Need for uniqueness, need for cognition and creativity. *Journal of Creative Behavior*, 37, 99-116.
- Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. *Journal of Social Issues*, 57(4), 829-849.
- Dovidio, J. F. & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In J.M.Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (204-237).
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 33 (5), 510-540.
- Downing, L.L. & Monaco, N.R. (1986). In-Group/Out-Group Bias as a Function of Differential Contact and Authoritarian Personality. *The Journal of Social Psychology*, 126 (4), 445-452.
- Duckitt, J. H. (1992). *The social psychology of prejudice*. Westport, CT, US: Praeger Publishers/Greenwood Publishing Group, Inc.
- Duckitt, J. (2001). A dual process cognitive-motivational theory of ideology and prejudice. In M.P. Zanna (Hrsg.), *Advances in experimental social psychology*, Vol 33, (41-113). San Diego, CA: Academic Press.
- Duckitt, J., Wagner, C. & du Plessis, I. (2002). The psychological bases of ideology and prejudice: Testing a dual process model. *Journal of Personality & Social Psychology*, 83(1), 75-93.
- Duden - *Das Fremdwörterbuch* (2005). Bibliographisches Institut & F.A. Mannheim: Brockhaus AG.
- Dunton, B. C. & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality & Social Psychology Bulletin*, 23, 316-326.
- Duriez, B. & Van Hiel, A. (2002). The march of modern fascism. A comparison of social dominance orientation and authoritarianism. *Personality & Individual Differences*, 32, 1199-1213.
- Durrheim, K. (1998). The relationship between tolerance of ambiguity and attitudinal conservatism: A multidimensional analysis. *European Journal of Social Psychology*, 28(5), 731-753.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL, US: Harcourt Brace Jovanovich College Publishers.
- Eagly, A. H. & Mladinic, A. (1989). Gender stereotypes and attitudes towards women and men. *Personality and Social Psychological Bulletin*, 15, 543-558.

- Eckhardt, W. (1991). Authoritarianism. *Political Psychology*, 12, 97-124.
- Edwards, A.L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, 65 (2), 241-261.
- Eid, M., Lischetzke, T., Nussbeck, F. & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods*, 8(1), 38-60.
- Egloff, B. & Schmukle (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality & Social Psychology*, 83, 1441-1455.
- Egloff, B. & Schmukle, S. C. (2004). Gender differences in implicit and explicit anxiety measures. *Personality & Individual Differences*, 36, 1807-1815.
- Egloff, B., Schwerdtfeger, A. & Schmukle, S. C. (2005). Temporal Stability of the Implicit Association Test--Anxiety. *Journal of Personality Assessment*, 84(1), 82-88.
- Egloff, B., Wilhelm, F. H., & Neubauer, D. (2002). Implicit anxiety measure predicts cardiovascular reactivity to an evaluated speaking task. *Emotion*, 2, 3-11.
- Ekehammar, B., Akrami, N., & Araya, T. (2003). Gender differences in implicit prejudice. *Personality & Individual Differences*, 34, 1509-1523.
- Ellwart, T., Rinck, M. & Becker, E. (2005). Comparing indirect measures of associations: Application of STIAT and EAST to fear of spiders. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Eysenck, H.J. (1954). *The Psychology of Politics*. London: Routledge & Kegan Paul.
- Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In: C. Hendrick & M. S. Clark: *Research methods in personality and social psychology* ( 74-97). Thousand Oaks, CA, US: Sage Publications, Inc.
- 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 & Social Psychology*, 69, 1013-1027.
- Fazio, R. H. & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology*, 54, 297-327.
- Fazio, R. H., Roskos-Ewoldsen, D. R., & Powell, M. C. (1994). *Attitudes, perception, and attention*. San Diego: Academic Press, Inc.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality & Social Psychology*, 50, 229-238.
- Fazio, R. H. & Towles-Schwen, T. (1990). The MODE model of attitude-behavior processes. In S. Chaiken, & Y. Trope: *Dual-process theories in social psychology* (97-116). New York, NY, US: Guilford Press.
- Feather, N.T. (1996). Reactions to Penalties for an Offense in Relation to Authoritarianism, Values: Perceived Responsibility, perceived Seriousness, and Deservingness. *Journal of Personality and Social Psychology*, 71 (3), 571-587.
- Fenigstein, A., Scheier, M. F. & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting & Clinical Psychology*, 43(4), 522-527.
- Fichter, M. M. & Quadflieg, N. (2000). Comparing self- and expert rating: A self-report screening version (SIAB-S) of the Structured Interview for Anorexic and Bulimic Syndromes for DSM-IV and ICD-10 (SIAB-EX). *European Archives of Psychiatry & Clinical Neuroscience*, 250(4), 175-185.
- Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *Journal of Personality*, 18, 108-143.
- Frindte, W., Funke, F. & Jacob, S. (1997). Autoritarismus, Wertorientierungen und jugendliche Identifikation. *Gruppendynamik*, 28, 273-289.

- Frindte, W., Funke, F. & Jacob, S. (1999). Neu-alte Mythen über Juden: Ein Forschungsbericht. In R. Dollase, T. Kliche & H. Moser (Hrsg.), *Politische Psychologie der Fremdenfeindlichkeit. Opfer-Täter-Mittäter* (119-130). Weinheim: Juventa.
- Frings, C. & Wentura, D. (2003). Who is watching Big Brother? TV consumption predicted by masked affective priming. *European Journal of Social Psychology*, 33, 779-791.
- Funke, F. (1996). *Autoritarismus und Ambiguitätsintoleranz: Die Bedeutung der politischen Orientierung für die Kovariation beider Konstrukte*. Unveröffentlichte Diplomarbeit. Friedrich-Schiller-Universität Jena.
- Gaertner, S. L. & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly*, 46, 23-30.
- Garrow, J. (1974). *Energy balance and obesity in man*. New York: American Elsevier.
- Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology*, 49, 171-180.
- Gawronski, B. & Conrey, F. R. (2004). Der Implizite Assoziationstest als Maß automatisch aktivierter Assoziationen: Reichweite und Grenzen. *Psychologische Rundschau*, 55(3), 118-126.
- Gemar, M. C., Segal, Z. V. & Sagrati, S. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology*, 110(2), 282-289.
- Grayson, D. & Marsh, H.W. (1994). Identification with deficient rank loading matrices in confirmatory factor analysis: Multitrait-Multimethod-Models. *Psychometrika*, 59 (1), 121-134.
- Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Greenwald, A. G., Banaji, M. R., & Rudman, L. A. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109, 3-25.
- Greenwald, A. G. & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality & 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 & Social Psychology*, 74, 1464-1480.
- Grover, V. P., Keel, P. K., & Mitchell, J. P. (2003). Gender Differences in Implicit Weight Identity. *International Journal of Eating Disorders*, 34, 125-135.
- Gschwendner-Lukas T., Hofmann W. & Schmitt M. (2004). Motivation zur Vorurteilkontrolle und soziale Normen: Eine Validierungsstudie. *Poster, präsentiert auf dem 44. Kongress der Deutschen Gesellschaft für Psychologie*, Göttingen, September, 2004.
- Haddock, G, Zanna, M.P. & Esses, V.M. (1993). Assessing the structure of prejudicial attitudes. The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology*, 65 (6), 1105-1118.
- Hänsel D. (1995): Eß-Störungen. In: V. Faust: *Psychiatrie - Ein Lehrbuch für Klinik, Praxis und Beratung* (655-664). Stuttgart: Gustav-Fischer-Verlag
- Hänze, M. (2002). Beduerfnis nach Struktur und Furcht vor Festlegung: Psychometrische Analysen einer deutschsprachigen Skala zur Erfassung der Konstrukte "Personal Need for Structure", "Personal Fear of Invalidity" und "Need for Closure". *Zeitschrift fuer Differentielle und Diagnostische Psychologie*, 23, 327-338.
- Harris, M. B., Walters, L. C., & Waschull, S. (1991). Gender and ethnic differences in obesity-related behaviors and attitudes in a college sample. *Journal of Applied Social Psychology*, 21, 1545-1566.
- Harris, M. B. & Foltz, S. (1999). Attitudes toward weight and eating in young women tennis players, their parents, and their coaches. *Eating Disorders: The Journal of Treatment & Prevention*, 7, 191-205.
- Hassan, M.K. & Khalique, A. (1987). A study of prejudice in Hindu & Muslim college students. *Psychologia: Annual International Journal of Personality in the Orient*, 30, 80-84.
- Hayes, B. C. (1995). Religious identification and moral attitudes: The British case. *British Journal of Sociology*, 46, S. 457-474.



- Heaven, P. C. L. & Quintin, D. St. (2003). Personality factors predict racial prejudice. *Personality & Individual Differences*, 34, 625-634.
- Heaven, P. C. L. & Connors, J. R. (2001). TI- A note on the value correlates of social dominance orientation and right-wing authoritarianism. *Personality & Individual Differences*, 31, 925-930.
- Hebl, M. R., Foster, J. B., Mannix, L. & Dovidio, J. (2002). Formal and interpersonal discrimination: A field study of bias toward homosexual applicants. *Personality & Social Psychology Bulletin*, 28(6), 815-825.
- Hebl, M. R. & Mannix, L. M. (2003). The weight of obesity in evaluating others: A mere proximity effect. *Personality & Social Psychology Bulletin*, 29, 28-38.
- Herek, G. M (1994). Assessing heterosexuals' attitudes toward lesbians and gay men: A review of empirical research with the ATLG scale. In: Greene, B. & Herek, G. M, *Lesbian and gay psychology: Theory, research, and clinical applications* (206-228). Thousand Oaks, CA, US: Sage Publications, Inc.
- Herek, G. M & Capitano, J. P. (1996). "Some of my best friends": Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians. *Personality & Social Psychology Bulletin*, 22(4), 412-424.
- Herek, G. M. (2000). The Psychology of Sexual Prejudice. *Current Directions in Psychological Science*, 9.
- Hess, E. H. (1965). Attitude and pupil size. *Scientific American*, 212(4), 46-54.
- Hibbard, S., Farmer, L., & Wells, C. (1994). Validation of Cramer's Defense Mechanism Manual for the TAT. *Journal of Personality Assessment*, 63, 197.
- Hibbard, S., Mitchell, D., & Porcelli, J. (2001). Internal Consistency of the Object Relations and Social Cognition Scales for the Thematic Apperception Test. *Journal of Personality Assessment*, 77, 408-419.
- Hoffmann, J., Kiesel, A. & Sebald, A. (2003). Task switches under Go/NoGo conditions and the decomposition of switch costs. *European Journal of Cognitive Psychology*, 15 (1), 101-128.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. (in press). A Meta-Analysis on the Correlation between the Implicit Association Test and Explicit Self-Report Measures. *Personality and Social Psychology Bulletin*.
- Hofmann, W., Gschwendner, T. & Schmitt, M. (2005). On implicit-explicit consistency: The role of individual differences in awareness and adjustment. *European Journal of Personality*, 19, 25-49.
- Holtgraves, T. (2004). Social Desirability and Self-Reports: Testing Models of Socially Desirable Responding. *Personality and Social Psychology Bulletin*, 30 (2), 161-172.
- Hoyle, R.H. (1995). *Structural equation modelling: Concepts, issues and applications*. Thousand Oaks: Sage Publications.
- Hugenberg, K. & Bodenhausen, G. V. (2004). Ambiguity in Social Categorization: The role of prejudice and facial affect in race categorization. *Psychological Science*, 15, 342-345.
- Huijding, J. (2005). The malleability of automatic spider-fear related affective associations over the course of an exposure treatment. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Huijding, J., De Jong, P. J. & ter Hart, L. (2003). Specific Predictive Power of Implicit Associations for Automatic Fear Behavior, submitted for publication.
- Hummert, M. L., Garstka, T. A., & O'Brien, L. T. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology & Aging*, 17, 482-495.
- Jacoby, L. L., Lindsay, D. S. & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist*, 47, 802-809.
- James W. P. T, Ferro-Luzzi A. & Waterlow J. C. (1988): Definition of chronic energy deficiency in adults. *European Journal of Clinical Nutrition*, 42, 969-981.
- Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and Explicit Measures of Sexual Orientation Attitudes: Ingroup Preferences and Related Behaviors and Beliefs Among Gay and Straight Men. *Personality & Social Psychology Bulletin*, 30, 629-642.

- Judd, C. M., Drake, R. A. & Downing, J. W. (1991). Some dynamic properties of attitude structures: Context-induced response facilitation and polarization. *Journal of Personality & Social Psychology*, 60(2), 193-202.
- Karpinski, A. & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality & Social Psychology*, 81, 774-788.
- Katz, I. & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality & Social Psychology*, 55, 893-905.
- Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality & Social Psychology Bulletin*, 24, 407-416.
- Kawakami, K. & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality & Social Psychology Bulletin*, 27, 212-225.
- Kawakami, K., Dovidio, J. F., & Dijksterhuis, A. (2003). Effect of social category priming on personal attitudes. *Psychological Science*, 14, 315.
- Kawakami, K., Spears, R., & Dovidio, J. F. (2002). Disinhibition of stereotyping: Context, prejudice, and target characteristics. *European Journal of Social Psychology*, 32, 517-530.
- Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activations. *Personality & Social Psychology Bulletin*, 28, 3-15.
- Kämpfe, N. (2002). *Persönlichkeit, soziale Einstellungen und Fremdenfeindlichkeit: Eine empirische Untersuchung mit Selbst- und Bekannteneinschätzung*.
- Kenny, D.A. & Kashy, D.A. (1992). Analysis of the Multitrait-Multimethod Matrix by confirmatory factor analysis. *Psychological Bulletin*, 112 (1), 165-172.
- Kirscht, J.P. & Dillehay, R.C. (1967). *Dimensions of authoritarianism*. Lexington, Ky: University of Kentucky Press.
- Kischkel, K.-H. (1983). Zum Effekt einiger Ausbildungsfaktoren auf die Ambiguitätstoleranz von Lehrern und die Beziehung zwischen Ambiguitätstoleranz und Konservatismus. *Zeitschrift für Empirische Pädagogik*, 7, 23-41.
- Kite, M. E. & Whitley, B. E. Do heterosexual women and men differ in their attitudes toward homosexuality? A conceptual and methodological analysis. In G. M. Herek: *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals* (39-61). Thousand Oaks, CA, US: Sage Publications, Inc.
- Klein, A. (2000). *Moderatormodelle: Verfahren zur Analyse von Moderator-effekten in Strukturgleichungsmodellen*. Hamburg: Kovac.
- Klein, O., Snyder, M. & Livingston, R. W. (2004). Prejudice on the stage: Self-monitoring and the public expression of group attitudes. *British Journal of Social Psychology*, 43(2), 299-314.
- Kline, R. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Koestner, R., Weinberger, J., & McClelland, D. C. (1991). Task-intrinsic and social-extrinsic sources of arousal for motives assessed in fantasy and self-report. *Journal of Personality*, 59, 57-82.
- Kraiger, K., Hakel, & Cornelius. Exploring Fantasies of TAT Reliability. *Journal of Personality Assessment* 48[4], 365. 1984. Lawrence Erlbaum Associates.
- Kruglanski, A. W., Atash, M. N., DeGrada, E., Manetti, L., Pierro, A., & Webster, D. M. (1997). Psychological theory testing versus psychometric nay-saying: Comment on Neuberg et al.'s (1997) critique of the Need for Closure Scale. *Journal of Personality & Social Psychology*, 73, 1005-1016.
- Kruglanski, A. W. & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review*, 103, 263-283.
- Kruglanski, A. W. & Webster, D. M. (2000). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review*, 103, 263-283.

- Kühnen, U., Schiessl, M., Bauer, N., Paulig, N., Poehlmann, C., & Schmidhals, K. (2001). How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. *Zeitschrift fuer Experimentelle Psychologie*, *48*, 135-144.
- Kwon, P., Campbell, & Williams (2001). Sociotropy and Autonomy: Preliminary Evidence for Construct Validity Using TAT Narratives. *Journal of Personality Assessment*, *77*, 128.
- LaMar, L. & Kite, M. (1998). Sex differences in attitudes toward gay men and lesbians: A multidimensional perspective. *Journal of Sex Research*, *35*, 189-196.
- Lance, C.E., Noble, C.L. & Scullen, S.E. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. *Psychological Methods*, *7* (2), 228-244.
- Laythe, B., Finkel, D., Bringle, R., & Kirkpatrick, L. A. (2002). Religious fundamentalism as a predictor of prejudice: A two-component model. *Journal for the Scientific Study of Religion*, *41*, 623-635.
- Leak, G. K. & Randall, B. A. (1995). Clarification of the link between right-wing authoritarianism and religiousness: The role of religious maturity. *Journal for the Scientific Study of Religion*, *34*(2), 245-252.
- Lehrl, S., Merz, J., Burkard, G. & Fischer, B. (1991). Mehrfachwahl-Wortschatz-Intelligenztest (MWT-A). Erlangen: perimed Fachbuch-Verlagsgesellschaft mbH.
- Leigh, J., Westen, D., & Barends, A. (1992). The assessment of complexity of representations of people using TAT and interview data. *Journal of Personality*, *60*, 809-837.
- Leone, C., Wallace, H. M., & Modglin, K. (1999). The need for closure and the need for structure: Interrelationships, correlates, and outcomes. *Journal of Psychology*, *133*, 553-562.
- Lepore, L. & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality & Social Psychology*, *72*, 275-287.
- Levin, S. & Sidanius, J. (1999). Social dominance and social identity in the United States and Israel: In-group favoritism or outgroup derogation? *Political Psychology*, *20*(1), 99-126.
- Lippa, R. & Arad, S. (1999). Gender, Personality, and Prejudice: The Display of Authoritarianism and Social Dominance in Interviews with College Men and Women. *Journal of Research in Personality*, *33*, 463-493.
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality & Social Psychology*, *81*, 842-855.
- Mackenzie, M. (1984). *Fear of fat*. New York: Columbia University Press.
- Madigan, S. (1983). Picture memory, In J.C. Yuille (Ed.), *Imagery, memory and cognition: Essays in honor of Allan Paivio*. Hillsdale, NJ: Erlbaum.
- Maison, D., Greenwald, A. G., & Bruin, R. (2001). The Implicit Association Test as a measure of implicit consumer attitudes. *Polish Psychological Bulletin*, *32*, 61-69.
- Marlowe, D. & Crowne, D.P (1960). *The approval motive-studies in evaluative dependence*. New York: Wiley.
- Marsh, H.W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, *13*, 335-361.
- Marsh, H.W. & Craven, R.G. (1991). Self-other agreement on multiple dimensions of preadolescent self-concept inferences by teachers, mothers, and fathers. *Journal of Educational Psychology*, *83* (3), 393-404.
- May, K. & Hittner, J. B. (2003). On the relation between power and reliability of difference scores. *Perceptual & Motor Skills*, *97*(3), 905-908.
- McClelland, D. C. (1985). How motives, skills, and values determine what people do. *American Psychologist*, *40*, 812-825.
- McClelland, D. C. (1995). Achievement motivation in relation to achievement-related recall, performance, and urine flow, a marker associated with release of vasopressin. *Motivation & Emotion*, *19*, 59-76.
- McClelland, D. C. & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied Psychology*, *67*(6), 737-743.

- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690-702.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J.F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (91-125).
- McConnell, A. R. & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435-442.
- McFarland, S.G. & Adelson, S. (1996). *An omnibus study of personality, values, and prejudice. Vortrag präsentiert auf dem Kongress der International Society for Political Psychology*, Vancouver, British Columbia, July 1996.
- McFarland, S. G. & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, *20*(6), 483-510.
- McGivern, R. F., Mutter, K. L., Anderson, J., Wideman, G., Bodnar, M., & Huston, P. J. (1998). Gender differences in incidental learning and visual recognition memory: Support for a sex difference in unconscious environmental awareness. *Personality & Individual Differences*, *25*, 223-232.
- McGregor, I., Zanna, M.P., Holmes, J.G. & Spencer, S.J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology*, *80* (3), 472-488.
- Meier, B. & Perrig, W. J. (2000). Low Reliability of Perceptual Priming: Consequences for the Interpretation of Functional Dissociations Between Explicit and Implicit Memory. *Quarterly Journal of Experimental Psychology: Section A*, *53* (1), 211-233.
- Meyers Enzyklopädisches Lexikon in 25 Bänden (1971-1981, 9. Aufl.). Mannheim: Bibliographisches Institut.
- Meyer, D. E., Glass, J. M. & Mueller, S. T. (2001). Executive-process interactive control: A unified computational theory for answering 20 questions (and more) about cognitive ageing. *European Journal of Cognitive Psychology*, *13*(1-2), 123-164.
- Mierke, J., Musch, J. & Klauer, K. C. (2003). The positivity proportion effect: A list context effect in masked affective priming. *Memory & Cognition*, *31*(6), 953-967.
- Mierke, J. & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift fuer Experimentelle Psychologie*, *48* (2), 107-122.
- Mierke, J. & Klauer, K. C. (2003). Method-Specific Variance in the Implicit Association Test. *Journal of Personality & Social Psychology*, *85* (6), 1180-1192.
- Millman, J. (1980). Implications of minimum competence testing programs for achievement and achievement testing. *New Directions for Testing & Measurement*, *5*, 31-38.
- Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*, 455-469.
- Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (2000). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. In C. Stangor (Ed.), *Stereotypes and prejudice: Essential readings* (305-322).
- Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition*, *19*, 395-417.
- Morrison, T. G., McLeod, L. D., Morrison, M. A. (1997). Gender stereotyping, homonegativity, and misconceptions about sexually coercive behavior among adolescents: Erratum. *Youth & Society*, *29*(1), 134.
- Morrison, T. G., Parriag, A. V. & Morrison, M. A. (1999). The psychometric properties of the homonegativity scale. *Journal of Homosexuality*, *37*(4), 111-126.
- Muller, J. (1997). *Conservatism. An anthropology of social and political thought*. Princeton: University Press.
- Murray, H. A. (1943). *Thematic apperception test*. Cambridge, MA, US: Harvard University Press.

- Musch, J. & Klauer, K. C. (2003). *The psychology of evaluation: Affective processes in cognition and emotion*. Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Musch, J., Brockhaus, R., & Broeder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit. *Diagnostica*, 48, 121-129.
- Neubauer, A. C. (1997). The mental speed approach to the assessment of intelligence In: J. Kingma and W. Tomic, (Eds), *Advances in cognition and educational practice: Reflections on the concept of intelligence* (149–174). Greenwich, Connecticut: JAI Press.
- Neuberg, S. L., Judice, T. N., & West, S. G. (1997). What the Need for Closure Scale measures and what it does not: Toward differentiating among related epistemic motives. *Journal of Personality & Social Psychology*, 72, 1396-1412.
- Neuberg, S. L., West, S. G., Thompson, M. M., & Judice, T. N. (1997). On dimensionality, discriminant validity, and the role of psychometric analyses in personality theory and measurement: Reply to Kruglanski et al.'s (1997) defense of the Need for Closure Scale. *Journal of Personality & Social Psychology*, 73, 1017-1029.
- Neuberg, S. N. & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality & Social Psychology*, 65, 113-131.
- Neumann, R., Hülsenbeck, K. & Seibt, B. (2004). Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of Experimental Social Psychology*, 40(4), 543-550.
- Neumann, R. & Seibt, B. (2001). The structure of prejudice: Associative strength as a determinant of stereotype endorsement. *European Journal of Social Psychology*, 31, 609-620.
- Nosek, B. A. (2003). Moderators of the relationship between implicit and explicit attitudes. *Dissertation Abstracts International: Section B: The Sciences & Engineering*, 63.
- Nosek, B. A. & Banaji, M. R. (2002). (At least) two factors moderate the relationship between implicit and explicit attitudes. In R.K.Ohme & M. Jarymowicz (Eds.), *Natura Automatyzmow* (49-56). Warszawa: WIP PAN & SWPS.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math not = me. *Journal of Personality & Social Psychology*, 83, 44-59.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics*, 6, 101-115.
- Nosek, B. A. & Banaji, M. R. (2001). The GO/NO-GO Association Task. *Social Cognition*, 19, 625-668.
- Nosek, B. A., Greenwald, A. G. & Banaji, M. R. (2005). Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity. *Personality & Social Psychology Bulletin*, 31(2), 166-180.
- Nosek, B. A., & Hansen, J. J. (2004). *The associations in our heads belong to us: Measuring the multifaceted attitude construct in implicit social cognition*. Unpublished manuscript. University of Virginia.
- Oesterreich, D. (1993). *Autoritäre Persönlichkeit und Gesellschaftsordnung. Der Stellenwert psychischer Faktoren für politische Einstellungen – eine empirische Untersuchung von Jugendlichen in Ost und West*. Weinheim: Juventa.
- Olson, M. A. & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science*, 14, 636-639.
- Olson, M. A. & Fazio, R. H. (2004). Reducing the Influence of Extrapersonal Associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality & Social Psychology*, 86, 653-667.
- Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitudes and racism: Effects of word familiarity and frequency on the implicit association test. *Social Cognition*, 19, 97.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255-287.
- Paulhus, D. L. & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality & Social Psychology*, 60, 307-317.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York, NY, US: Cambridge University Press.

- Pelham, B. W., Koole, S. L., Hardin, C. D., Hetts, J. J., Seah, E. & DeHart, T. (2005). Gender moderates the relation between implicit and explicit self-esteem. *Journal of Experimental Social Psychology*, 41(1), 84-89.
- Pentony, J.F., Petersen, K.S.E., Philips, O., Leong, C., Harper, P., Bakowski, A., Steward, S. & Gonzales, R. (2000). A comparison of authoritarianism in the United States, England, and Hungary with selected non-random samples. *European Psychologist*, 5 (4), 259-268.
- Perugini, M., O'Gorman, R. & Connor, M. (2005). The EAST as an inter-individual difference measure: Promising results from three studies. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Petty, R. E. & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Wm. C. Brown Company Publishers.
- Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology*, 48, 609-647.
- Phelps, E. A.; O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C. & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12 (5), 729-738.
- Plant, E. A. & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality & Social Psychology*, 75, 811-832.
- Pratto, F. (1994). Consciousness and automatic evaluation. , In P. M. Niedenthal & S. Kitayama: *Heart's eye: Emotional influences in perception and attention* (115-143). San Diego, CA, US: Academic Press
- Pratto, F. & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality & Social Psychology*, 61(3), 380-391.
- Pratto, F. & Shih, M. (2000). Social Dominance Orientation and Group Context in Implicit Group Prejudice. *Psychological Science*, 11, 515-518.
- Pratto, F., Sidanius, J., Stallworth, L.M. & Malle, B.F. (1994). Social dominance orientation, A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67 (4), 741-763.
- Prohansky, H. M. (1966). The development of intergroup attitudes. In L. W. Hoffman & M. L. Hoffman (Eds.), *Review of child development research*, Vol. 2, (311-371). New York: Russell Sage Foundation.
- Ray, J.J. & Lovejoy, F.H. (1986). The generality of racial prejudices. *Journal of Social Psychology*, 126, 563-564.
- Reis, J. (1997). *Ambiguitätstoleranz: Beiträge zur Entwicklung eines Persönlichkeitskonstruktes*. Heidelberg: A-sanger.
- Riemann, R., Grubich, C., Hempel, S., & Mergl, S. (1993). Personality and attitudes towards current political topics. *Personality & Individual Differences*, 15, 313-321.
- Rippl, S. & Seipel, C. (1998). Autoritarismus und Fremdenfeindlichkeit bei ost- und westdeutschen Jugendlichen. *Zeitschrift für Politische Psychologie*, 6 (3), 273-288.
- Roefs, A. & Jansen, A. (2002). Implicit and explicit attitudes toward high-fat foods in obesity. *Journal of Abnormal Psychology*, 111, 517-521.
- Rogosa, D., Brandt, D. & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, 92(3), 726-748.
- Ronan, G. F. (1990). Manual for the Personal Problem-Solving-System for Scoring Thematic Apperception Test responses Personal-Problem-Solving System for Scoring TAT Responses. Unpublished manuscript.
- Ronan, G. F., Colavito, V. A., & Hammontree, S. R. (1993). Personal Problem-Solving System for Scoring TAT Responses: Preliminary Validity and Reliability Data. *Journal of Personality Assessment*, 61, 28-40.
- Ronan, G. F., Date, A. L., & Weisbrod, M. (1995). Personal problem-solving scoring of the TAT: Sensitivity to training. *Journal of Personality Assessment*, 64, 119-131.

- Ronan, G. F., Senn, J., & Date, A. L. (1996). Personal Problem-Solving Scoring of TAT Responses: Known-Groups Validation. *Journal of Personality Assessment*, 67, 641-653.
- Rosenberg, S. D., Blatt, S. J., & Oxman, T. E. (1994). Assessment of Object Relatedness Through a Lexical Content Analysis of the TAT. *Journal of Personality Assessment*, 63, 345-362.
- Rosenberg, M. J., Hovland, C. I. & McGuire, W. J. (1960). *Attitude organization and change: An analysis of consistency among attitude components (Yales studies in attitude and communication. Vol. III.)*. Oxford, England: Yale University Press.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Göttingen: H. Huber.
- Rothblum, E. D., Brand, P. A., Miller, C. T. & Oetjen, H. A. (1990). The relationship between obesity, employment discrimination, and employment-related victimization. *Journal of Vocational Behavior*, 37(3), 251-266.
- Rothermund, K. & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift fuer Experimentelle Psychologie*, 48, 94-196.
- Rothermund, K. & Wentura, D. (2004). Underlying Processes in the Implicit Association Test: Dissociating Salience From Associations. *Journal of Experimental Psychology: General*, 133, 139-165.
- Rowatt, W. C. & Franklin, L. M. (2004). Christian orthodoxy, religious fundamentalism, and right-wing authoritarianism as predictors of implicit racial prejudice. *International Journal for the Psychology of Religion*, 14, 125-138
- Rowatt, W. C., Franklin, L. M., & Cotton, M. (2005). Patterns and Personality Correlates of Implicit and Explicit Attitudes Toward Christians and Muslims. *Journal for the Scientific Study of Religion*, 44, 29-43.
- Ruch, W. & Hehl, F.-J. (1983). Intolerance of Ambiguity as a factor in the appreciation of humour. *Personality & Individual Differences*, 4(5), 443-449.
- Rudman, L. A., Greenwald, A. G., & Mellott, D. S. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition*, 17, 437-465.
- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality & Social Psychology*, 81, 856-868.
- Rudman, L. A., Feinberg, J. M., & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition*, 20, 294-320.
- Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, 13, 79-82.
- Sanbonmatsu, D. M. & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality & Social Psychology*, 59(4), 614-622.
- Sarma, C. (1973). Prejudice and personality. In T.E. Stranumgam (Hrsg.). *Research in personality and social problems* (96-114). Madras: University of Madras Press.
- Saucier, G. (2000). Isms and the structure of social attitudes. *Journal of Personality and Social Psychology*, 78(2), 366-385.
- Scheffer, D., Kuhl, J. & Eichstaedt, J. (2003). Der Operante Motiv-Test (OMT): Inhaltsklassen, Auswertung, psychometrische Kennwerte und Validierung. In F. Rheinberg & J. Stiensmeier-Pelster (Hrsg.): *Diagnostik von Motivation und Selbstkonzept* (S. 151-168). Göttingen; Hogrefe.
- Schenk, J. (1980). Clericalism and conservatism. *Psychologische Beiträge*, 22(1), 145-162.
- Schmalt, H.-D. & Sokolowski, K. (2000). Zum gegenwärtigen Stand der Motivdiagnostik. *Diagnostica*, 46, 115-123.
- Schmidt, T.E. (1995). *Straight & narrow? Compassion & clarity in the homosexual debate*. Aus Sonderdruck Verlag VBG 1997 (Hrsg: H. Gasser & W. Gasser).
- Schmitz, F., Teige, S., Voss A. & Klauer, C. (2005). Cognitive capacity, task shifts, and measurement issues in the EAST procedure. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop*, Oppurg, 08.-10. April, 2005.
- Schmukle, S. C. & Egloff, B. (2004). Does the Implicit Association Test for Assessing Anxiety Measure Trait and State Variance? *European Journal of Personality*, 18(6), 483-494.

- Schmukle, S. C. & Egloff, B. (2005). A Latent State-Trait Analysis of Implicit and Explicit Personality Measures. *European Journal of Psychological Assessment, 21*(2), 100-107.
- Schuette, R. A. & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality & Social Psychology Bulletin, 21*(7), 704-710.
- Schultheiss, O. C. & Brunstein, J. C. (2001). Assessment of Implicit Motives With a Research Version of the TAT: Picture Profiles, Gender Differences, and Relations to Other Personality Measures. *Journal of Personality Assessment, 77*, 71-86.
- Seise, J., Banse, R., & Neyer, F. (2002). Individuelle Unterschiede in impliziten und expliziten Einstellungen zur Homosexualität: Eine empirische Studie. *Zeitschrift fuer Sexualforschung, 15*, 21-42.
- Sharkey, K. J. & Ritzler (1985). Comparing Diagnostic Validity of the TAT and a New Picture Projective Test. *Journal of Personality Assessment, 49*, 406-412.
- Sherman, S. J., Chassin, L., Presson, C., Rose, J., & Koch, K. (1999). The relationship between the implicit and explicit attitudes of smokers and nonsmokers toward cigarette smoking. *Vortrag präsentiert bei der Midwest Psychological Association, Chicago, April, 1999*.
- Sherrod, D. & Nardi, P. M. (1998). Homophobia in the courtroom: An assessment of biases against gay men and lesbians in a multiethnic sample of potential jurors. In G. M. Herek: *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals* (24-38). Thousand Oaks, CA, US: Sage Publications, Inc.
- Sidanius, J. (1978). Intolerance of ambiguity and socio-political ideology: A multidimensional analysis. *European Journal of Social Psychology, 8*(2), 215-235.
- Sidanius, J. & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*.
- Sidanius, J., Pratto, F., & Bobo, L. (1994). Social dominance orientation and the political psychology of gender: A case of invariance? *Journal of Personality & Social Psychology, 67*, 998-1011.
- Sinha, R.R. & Hassan, M.K. (1975). Some personality correlates of social prejudice. *Journal of Social and Economic Studies, 3*, 225-231.
- Skowronski, J. J. & Lawrence, M. A. (2001). A Comparative Study of the Implicit and Explicit Gender Attitudes of Children and College Students. *Psychology of Women Quarterly, 25*, 155-165.
- Six, W. (1996). Generalisierte Einstellungen. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie. Differentielle und Persönlichkeitsforschung. Bd. 3: Temperaments- und Persönlichkeitsunterschiede* (1-50). Göttingen: Hogrefe.
- Six, B., Wolfradt, U. & Zick, A. (2001). Autoritarismus und Soziale Dominanzorientierung als generalisierte Einstellungen. *Zeitschrift für Politische Psychologie, 9* (2-3), 23-40.
- Snyder, M. (1992). Motivational foundations of behavioral confirmation. In M. P. Zanna: *Advances in experimental social psychology*, Vol. 25, (67-114). San Diego, CA, US: Academic Press.
- Son Hing, L. S., Li, W. & Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology, 38*(1), 71-78.
- Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*, 140-154.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 3*, 271-295.
- Stahl, C. & Degner, J. (2005). Automatic and controlled processes in the EAST. *Vortrag präsentiert auf dem DFG-geförderten EAST-Workshop, Oppurg, 08.-10. April, 2005*.
- Stangor, C. (2000). *Stereotypes and prejudice: Essential readings*. New York, NY, US: Psychology Press.
- Steffens, M. C. (2004). Is the Implicit Association Test Immune to Faking? *Experimental Psychology, 51*(3), 165-179.
- Steffens, M. C. & Buchner, A. (2003). Implicit Association Test: Separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology, 50*, 33-48.



- Steffens, M. C., Lichau, J. & Still, Y. (2004). Individuum oder Gruppe, Exemplar oder Kategorie?: Ein Zweifaktorenmodell zur Erklärung der Reaktionszeitunterschiede im Implicit Association Test(IAT). *Zeitschrift für Psychologie*, 212(2), 57-65.
- Steffens, M. C. & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift fuer Experimentelle Psychologie*, 48, 123-134.
- Steffens, M. C. & Wagner, C. (2004). Attitudes Toward Lesbians, Gay Men, Bisexual Women, and Bisexual Men in Germany. *Journal of Sex Research*, 41(2), 137-149.
- Stoeber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität. *Diagnostica*, 45, 173-177.
- Strack, F. & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality & Social Psychology Review*, 8, 220-247.
- Stunkard, A. J., Sorensen, T. I., Hanis, C., & Teasdale, T. W. (1986). "An adoption study of human obesity": Reply. *New England Journal of Medicine*, 315, 130.
- Suslow, T., Ohrmann, P. & Arolt, V. (2001). Ausgestaltung so genannter affektiver Voraktivierungseffekte in der evaluativen Entscheidungsaufgabe: Hinweise auf automatische Vigilanz für negative Informationen. *Zeitschrift für Psychologie*, 209(2), 137-152.
- Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behaviour consistency for stigmatised behaviour. *Cognition & Emotion*, 15, 207-230.
- Tabachnik, B.G. & Fidell, L.S. (2001). *Using multivariate statistics* (4. Ausg.). Boston: Allyn & Bacon.
- Teachman, B. A., Gapinski, K. D., Brownell, K. D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology*, 22, 68-78.
- Teachman, B. A. & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology*, 112, 100-109.
- Tedeschi, J. T., Schlenker, B. R., & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, 26, 685-695.
- Teige, S., Schnabel, K. Banse, R. & Asendorpf, J. (2004). Assessment of Multiple Implicit Self-Concept Dimensions Using the Extrinsic Affective Simon Task (EAST). *European Journal of Personality*, 18(6), 495-520.
- Thrash, T. M. & Elliot, A. J. (2002). Implicit and self-attributed achievement motives: Concordance and predictive validity. *Journal of Personality*, 70, 729-755.
- Tiggemann, M. & Rothblum, E. D. (1988). Gender differences in social consequences of perceived overweight in the United States and Australia. *Sex Roles*, 18(1-2), 75-86.
- Tiggemann, M. & Rothblum, E. D. (1988). Gender differences in social consequences of perceived overweight in the United States and Australia. *Sex Roles*, 18(1-2), 75-86.
- Trautner, H. M. (1992). The development of sex-typing in children: A longitudinal analysis. *German Journal of Psychology*, 16(3), 183-199.
- van Hiel, A. & Mervielde, I. (2002). Explaining conservative beliefs and political preferences: A comparison of social dominance orientation and authoritarianism. *Journal of Applied Social Psychology*, 32 (5), 965- 976.
- van Ijzendoorn, M.H. (1989). Moral Judgement, Authoritarianism, and Ethnocentrism. *The Journal of Social Psychology*, 129 (1), 37-45.
- Voss, A., Rothermund, K., & Wentura, D. (2003). Estimating the valence of single stimuli: A new variant of the affective Simon task. *Experimental Psychology*, 50, 86-96.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Oxford, England: Rand McNally.
- Webster, D. M. & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality & Social Psychology*, 67, 1049-1062.

- Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality & Social Psychology*, 39, 186-200.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- Widaman, K.F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied psychological measurement*, 9, (1), 1-26.
- Wiers, R. W., Van Woerden, N., Smulders, F. T. Y & de Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, 111(4), 648-658. Wikipedia, freie Enzyklopädie. <http://de.wikipedia.org/wiki>.
- Wigboldus, D., van Knippenberg, A. & Holland, R. (2001). The difference between relative and absolute implicit associations. *Vortrag präsentiert auf dem dritten Treffen des European Social Cognition Network*, Houfalize, Belgium.
- Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, 120, 3-24.
- Williams, R. H. & Zimmerman, D. W. (1996). Are simple gain scores obsolete: Commentary on the commentaries of Collins and Humphreys. *Applied Psychological Measurement*, 20(3), 295-297.
- Wilson, G.D. (1970). Is there a general factor in social attitudes? Evidence from a factor analysis of the Conservatism scale. *British Journal of Social and Clinical Psychology*, 9, 101-107.
- Wilson, G.D. (1973). *The Psychology of Conservatism*. London: Academic Press.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality & Social Psychology*, 72, 262-274.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. *Journal of Experimental Social Psychology*, 37, 244-252.
- Wojke, B. & Polo, M. (2001). Motive-Related Memories: Content, Structure, and Affect. *Journal of Personality*, 69, 391-415.
- Wooley, S. C., Wooley, O. W. & Dyrenforth, S. R. (1979). Theoretical, practical, and social issues in behavioral treatments of obesity. *Journal of Applied Behavior Analysis*, 12(1), 3-25.
- Zick, A. (1997). *Vorurteile und Rassismus: eine Sozialpsychologische Analyse*. Berlin: Waxmann.
- Zimbardo, P. G. (1995). *Psychologie* (6. Aufl.). Berlin: Springer Verlag.
- Zimbardo, P. G. & Perrig, R. J. (2004). *Psychologie: Eine Einführung*. Pearson Studium.
- Zimmerman, D. W. (1996). An efficient alternative to the Wilcoxon signed-ranks test for paired nonnormal data. *Journal of General Psychology*, 123(1), 29-40.
- Zacker, J. (1973). Authoritarian avoidance of ambiguity. *Psychological Reports*, 33(3), 901-902.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151-175.

## – Anhang –

<b>A.4.2</b>	Stimulusmaterial für IAT und EAST in Studie I	2
<b>A.6.2</b>	Stimulusmaterial für IAT und EAST in Studie III	4
	„Heimset“:           Einstellungsprädiktoren	7
	Offene Assoziativmessung	8
	„Laborset“	9
	Kodierschema für offene Assoziativmessung in Studie III	10
<b>A.6.4</b>	Systematischer Modellvergleich zur MTMM-Analyse	
	- Ausgangsdaten	11
	- Modifikation	14
<b>A.6.5</b>	Basismodelle der Vorhersage von Einstellungen	15

**A.4.2: Zielkonzeptstimuli für IAT- & EAST-Homosexualität  
in Studie I und III (schwarzweiß bzw. blau/grün schattiert)**



**A.4.2: Evaluatives Stimulusmaterial für IAT & EAST in Studie I**

**A.6.2: Zielkonzeptstimuli für IAT- & EAST- Übergewicht  
in Studie III (schwarzweiß bzw. blau/grün schattiert)**





**A.6.2: Evaluatives Stimulusmaterial für IAT & EAST in Studie III  
„unangenehm“:**



**A.6.2: Evaluatives Stimulusmaterial für IAT & EAST in Studie III  
„angenehm“:**





## **A.6.2: „Heimset“: Einstellungsprädiktoren**

Liebe Teilnehmer,

Vor Ihnen liegt nun ein Fragebogen zu verschiedenen Themen des alltäglichen Lebens. Wir sind an Ihrer persönlichen Meinung dazu interessiert.

Es gibt keine richtigen und keine falschen Antworten! Auch wenn Sie bei einer Aussage unsicher sind: Kreuzen Sie einfach spontan und offen an, was Ihrem persönlichen Empfinden am ehesten entspricht.

Für die Bewertung der Aussagen stehen Ihnen immer mehrere Antwortmöglichkeiten von völliger Ablehnung bis hin zu vollkommener Zustimmung zur Verfügung.

Die Antwortkästchen sind für eine bessere Übersichtlichkeit unterschiedlich schattiert:

Helleres Grau signalisiert *Ablehnung*,  
Dunkleres Grau signalisiert *Befürwortung* der jeweiligen Aussage.

Wir bitten Sie, keine Antwort auszulassen.

Wie weisen Sie nochmals darauf hin, dass die Behandlung aller Untersuchungselemente selbstverständlich absolut anonym erfolgt.

Beginnen Sie hier:

Bitte kreuzen Sie an, inwieweit Sie folgenden Aussagen zustimmend bzw. ablehnend gegenüber stehen		lehne voll- kommen ab	weder... noch...	stimme voll- kommen zu
1.	Ich probiere gerne Dinge aus, auch wenn nicht immer etwas dabei herauskommt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	Ich beschäftige mich nur mit Aufgaben, die lösbar sind	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	Ich mag es, wenn unverhofft Überraschungen auftreten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.	Ich lasse die Dinge gerne auf mich zukommen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.	Ich habe es gerne, wenn die Arbeit gleichmäßig verläuft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.	Ich warte geradezu darauf, dass etwas aufregendes passiert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.	Wenn um mich herum alles drunter und drüber geht, fühle ich mich richtig wohl	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8.	Ich weiß gerne, was auf mich zukommt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

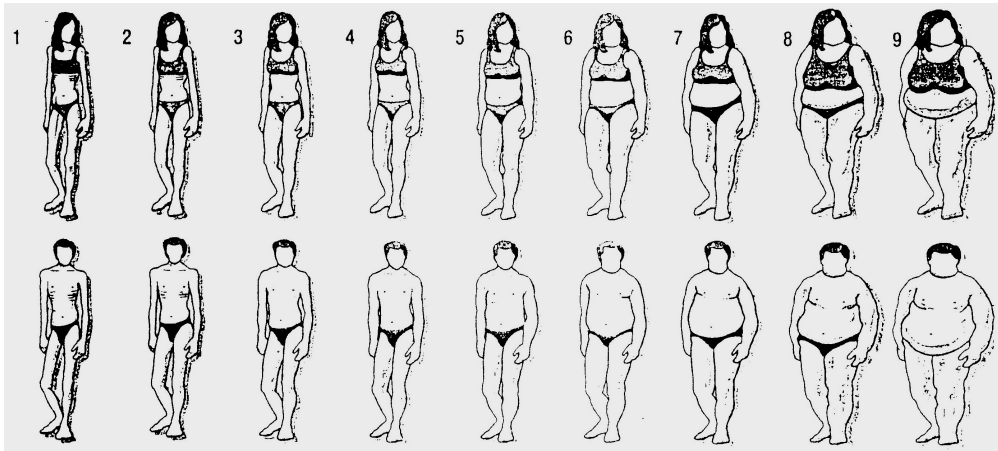
9.	Man sollte seine eigenen Moralvorstellungen über ‚Gut und Böse‘ entwickeln und weniger der Bibel oder anderen alten, traditionellen Glaubenssätzen Beachtung schenken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	Was wir in unserem Land anstelle von mehr ‚Bürgerrechten‘ wirklich brauchen, ist eine anständige Portion Recht und Ordnung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.	Die Zeiten, in denen sich Frauen ihren Männern unterzuordnen hatten, sollten der Vergangenheit angehören. Der ‚Platz einer Frau‘ in der Gesellschaft sollte sein, wo immer sie möchte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.	Die Abkehr von der Tradition wird sich eines Tages als fataler Fehler herausstellen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.	Es gibt kein Verbrechen, das die Todesstrafe rechtfertigen würde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.	Gehorsam und Achtung vor der Autorität sind die wichtigsten Tugenden, die Kinder lernen sollten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.	Gleichgeschlechtliche Lebensgemeinschaften sollten der Ehe gleichgestellt werden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.	Was unser Land wirklich braucht, ist ein starker, entschlossener Kanzler, der das Übel zerschlagen und uns wieder auf unseren rechten Weg bringen wird	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.	Es ist gut, dass die jungen Leute heutzutage größere Freiheiten haben, „ihr eigenes Ding zu machen“ und gegen Dinge zu protestieren, die sie nicht mögen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18.	Tugendhaftigkeit und Gesetzestreue bringen uns auf lange Sicht weiter, als das ständige Infragestellen der Grundfesten unserer Gesellschaft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19.	Es ist wichtig, die Rechte von Radikalen und Abweichlern in jeder Hinsicht zu wahren	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20.	Die wahren Schlüssel zum „guten Leben“ sind Gehorsam, Disziplin und Tugend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21.	Wir sollten danach streben, das Arbeitseinkommen so gleich wie möglich zu gestalten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22.	Gruppengleichheit sollte unser Ideal sein	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23.	Es ist in Ordnung, wenn einige Gruppen größere Chancen im Leben haben als andere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24.	Um im Leben vorwärts zu kommen, ist es manchmal notwendig, andere für sich einzuspannen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25.	Wir sollten unser möglichstes tun, um gleiche Bedingungen für unterschiedliche Gruppen zu schaffen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26.	Es ist im Grunde richtig, dass einige Gruppen besser gestellt sind als andere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27.	Unterlegene Gruppen sollten unter sich bleiben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28.	Wir hätten weniger Probleme, wenn wir alle Menschen gleicher behandeln würden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29.	Es wäre gut, wenn alle Gruppen gleichwertig wären	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30.	Um seine Ziele zu erreichen, ist es manchmal erforderlich, auf andere Menschen Druck auszuüben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31.	Alle Gruppen sollten im Leben die gleichen Chancen bekommen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32.	Wenn einige Gruppen unter sich bleiben würden, hätten wir weniger Probleme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33.	Soziale Gleichheit sollte zunehmen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34.	Ab und zu müssen andere Gruppen in ihre Schranken verwiesen werden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35.	Einige Menschen sind anderen gegenüber einfach unterlegen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36.	Keine einzelne Gruppe sollte in der Gesellschaft dominieren	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bitte geben Sie auch noch zu folgenden Aussagen den Grad Ihrer Zustimmung bzw. Ablehnung an		lehne voll- kommen ab	weder... noch...	stimme voll- kommen zu
37.	Ich würde komplexe Probleme den einfach zu lösenden vorziehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38.	Ich trage gern die Verantwortung für Situationen, die sehr viel Nachdenken erfordern	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39.	Denken entspricht nicht dem, was ich unter Spaß verstehe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
40.	Ich würde lieber etwas tun, das wenig Denken erfordert, als etwas, das mit Sicherheit meine Denkfähigkeit herausfordert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
41.	Ich versuche Situationen vorauszuahnen und zu vermeiden, in denen die Wahrscheinlichkeit groß ist, dass ich intensiv überlegen muss	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
42.	Es befriedigt mich, lang und angestrengt über etwas nachzudenken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
43.	Ich überlege nur so intensiv wie unbedingt nötig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
44.	Ich bevorzuge kleine, alltägliche Projekte gegenüber längerfristigen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
45.	Ich mag Tätigkeiten, die nur noch wenig Nachdenken erfordern, wenn man sie einmal gelernt hat	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
46.	Die Vorstellung, mich auf mein Denkvermögen zu verlassen, um es zu etwas zu bringen, reizt mich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47.	Ich mag Tätigkeiten, die das Suchen neuer Problemlösungen beinhalten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
48.	Mein Leben sollte bevorzugt voller Rätsel sein, die ich lösen muss	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
49.	Die Vorstellung, abstrakt zu denken, spricht mich an	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
50.	Ich ziehe Tätigkeiten vor, die Intelligenz erfordern, schwierig und bedeutend sind, als solche, die zwar irgendwie wichtig sind, aber nicht viel Überlegung verlangen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
51.	Wenn ich eine Aufgabe erledigt habe, die viel geistige Anstrengung erfordert hat, fühle ich mich eher erleichtert als befriedigt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
52.	Es reicht mir, dass etwas funktioniert, ich muss nicht wissen wie oder warum es funktioniert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
53.	Ich finde mich häufig wieder, wie ich über etwas nachdenke, auch wenn es mich nicht persönlich betrifft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
54.	Der erste Eindruck, den ich von anderen Menschen gewinne, bewahrheitet sich meistens	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
55.	Manchmal lüge ich, wenn ich muss	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
56.	Ich bin nicht immer mir selber gegenüber ganz ehrlich gewesen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
57.	Es ist schon einmal vorgekommen, dass ich jemanden ausgenutzt habe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
58.	Ich weiß immer, warum ich etwas mag	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
59.	Ich fluche niemals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
60.	Es fällt mir schwer, einen beunruhigenden Gedanken beiseite zu drängen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
61.	Manchmal zahle ich es lieber anderen heim, als dass ich vergebe und vergesse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
62.	Manchmal verpasse ich etwas, weil ich mich einfach nicht schnell genug entscheiden kann	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
63.	Ich habe schon einmal zuviel Wechselgeld herausbekommen, ohne es der Verkäuferin zu sagen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
64.	Ich bin ein vollkommen rational denkender Mensch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
65.	Ich würde grundsätzlich alles angeben, was ich zu verzollen hätte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
66.	Ich kann Kritik selten vertragen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
67.	Manchmal fahre ich schneller, als erlaubt ist	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
68.	Ich bin mir meiner Urteile sehr sicher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
69.	Ich habe Dinge getan, von denen ich anderen nichts erzähle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

70.	An meinen Fähigkeiten als Liebhaber/in habe ich schon gelegentlich gezweifelt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
71.	Ich nehme niemals Dinge an mich, die mir nicht gehören	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
72.	Ich weiß nicht immer die Gründe für mein Handeln	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
73.	Jeder Mensch braucht etwas, woran er voll und ganz glauben kann	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
74.	Unter Freiheit muss bei den heutigen Bedingungen zuerst einmal Freiheit von gesellschaftlichen und staatlichem Zwang verstanden werden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
75.	Kriege wird es auch weiterhin geben, die Menschen sind nun einmal so	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
76.	Es gehört zur unveränderlichen Natur des Menschen, nach persönlichem Besitz zu streben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
77.	Die Gründe für zwischenmenschliche Konflikte und Auseinandersetzungen muss man in erster Linie in gesellschaftlichen Verhältnissen und nicht in der individuellen Moral der Menschen suchen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
78.	Der Mensch wird immer das Bedürfnis nach einer übergreifenden Ordnung haben, die seinem Denken und Handeln einen festen Halt gibt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
79.	Unsere Gesellschaft verhindert immer noch die Befriedigung wichtiger menschlicher Bedürfnisse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
80.	Es gibt für den Menschen wirklich wichtigere Aufgaben, als alles mit dem Verstand durchdringen zu wollen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
81.	Der Mensch sollte nur diejenigen staatlichen Beschränkungen seiner Freiheit befolgen, deren Berechtigung er auch selber als vernünftig ansieht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
82.	Nicht die Vernunft, sondern das Gewissen ist die verlässliche Instanz zur Beurteilung von Gut und Böse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
83.	Bei Menschen, die nicht ständig gegen das Böse in sich ankämpfen, gewinnt es rasch die Oberhand	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
84.	Ständige intensive Kritik – und nicht Vertrauen – gewährleistet das Funktionieren einer demokratischen Gesellschaft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
85.	Auch wenn man die Menschen vollständig über ein Problem informieren würde, wären die meisten von ihnen nicht in der Lage, eine vernünftige Entscheidung zu treffen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
86.	Kriege sind nur eine Folge von gesellschaftlichen und politischen Gegebenheiten und haben mit der „Natur des Menschen“ nichts zu tun	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
87.	Dass der Mensch ein Bedürfnis nach Unterordnung unter eine wahre Autorität habe, behaupten allein diejenigen, die ihn in Unmündigkeit halten wollen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
88.	Ein noch so geschulter und kritischer Verstand kann letzten Endes doch keine echte innere Befriedigung geben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Im Folgenden möchten wir Sie um eine Selbsteinschätzung Ihrer Person bitten:**

Schauen Sie sich bitte die folgenden schematischen Abbildungen an:



89. In welchem dieser Bilder finden Sie sich wieder? (Geben Sie bitte die Nummer an)

90. Welche Figur entspricht am ehesten Ihrem Idealbild? (Geben Sie bitte die Nummer an)

Zum Abschluss bitten wir Sie um ein paar Angaben zu Ihrer Person:

91. Alter:

92. Geschlecht:  Weiblich  Männlich

Kreuzen Sie bitte an, wie stark Sie sich mit Ihrem Geschlecht identifizieren, also wie sehr Sie sich als Mann bzw. Frau fühlen:

93.

absolut als Mann weder, noch absolut als Frau

94. a) Wenn Sie eine **Frau** sind, kreuzen Sie nun hier bitte an, welcher partnerschaftlichen Orientierung Sie sich am nächsten fühlen:

lesbisch bisexuell heterosexuell

94. b) Wenn Sie ein **Mann** sind, kreuzen Sie nun hier bitte an, welcher partnerschaftlichen Orientierung Sie sich am nächsten fühlen:

schwul bisexuell heterosexuell

95. Höchster Bildungsabschluss:

<input type="checkbox"/> Hauptschulabschluss (8. Klasse)	<input type="checkbox"/> Berufsausbildung
<input type="checkbox"/> Realschulabschluss (10. Klasse)	<input type="checkbox"/> Hochschulabschluss (z.B. Magister, Diplom)
<input type="checkbox"/> Fachhochschulreife	<input type="checkbox"/> Höherer akademischer Grad
<input type="checkbox"/> Allgemeine Hochschulreife (Abitur)	<input type="checkbox"/> Ohne Abschluss <input type="checkbox"/> keine Angabe

96. Wo sind Sie aufgewachsen?

<input type="checkbox"/> Baden-Württemberg	<input type="checkbox"/> Niedersachsen
<input type="checkbox"/> Bayern	<input type="checkbox"/> Nordrhein-Westfalen
<input type="checkbox"/> Berlin	<input type="checkbox"/> Rheinland-Pfalz
<input type="checkbox"/> Brandenburg	<input type="checkbox"/> Saarland
<input type="checkbox"/> Bremen	<input type="checkbox"/> Sachsen
<input type="checkbox"/> Hamburg	<input type="checkbox"/> Sachsen-Anhalt
<input type="checkbox"/> Hessen	<input type="checkbox"/> Schleswig-Holstein
<input type="checkbox"/> Mecklenburg-Vorpommern	<input type="checkbox"/> Thüringen

nicht in Deutschland  keine Angabe

97. Was studieren Sie?

98. In welchem Semester studieren Sie?

Bringen Sie den ausgefüllten Fragebogen gemeinsam mit den anderen Unterlagen, die Sie mit nach Hause bekommen haben, zum letzten Teil der Untersuchung unbedingt wieder mit.

Herzlichen Dank für Ihre Teilnahme!

### **A.6.2: „Heimset“: Offene Assoziativmessung**



Liebe Teilnehmer,

im folgenden bitten wir Sie, Ihrer Fantasie und Kreativität freien Lauf zu lassen.

Vor Ihnen liegt ein Heft mit vier thematischen Skizzen bzw. Zeichnungen. Die Bilder könnten zur Illustration eines Kurzgeschichtenbandes dienen. Ihre Aufgabe ist es, die Zeichnungen nacheinander zu betrachten und selbst eine kleine Geschichte zu der jeweiligen Illustration zu verfassen. Es geht uns dabei *nicht* um künstlerische Aspekte der Bilder, sondern um ihren Inhalt, ihre Aussagekraft, also darum, was Sie persönlich mit dem Bild in Verbindung bringen. Versuchen Sie, so gut es geht, eine reine Beschreibung der Abbildungen zu vermeiden und sich stattdessen zu einer kleinen Erzählung durch das Bild anregen zu lassen

Als kleine Hilfestellung für Ihre eigenen Geschichten sollen Ihnen folgende Fragen dienen:

Was stellt die Illustration dar?

Was denken und fühlen die dargestellten Personen?

Wie kam es zu der abgebildeten Szene?

Wie könnte sich die Situation weiter entwickeln?

Nehmen Sie sich für jedes Bild und die dazu zu erzählende Geschichte etwa fünf Minuten Zeit und schreiben Sie Ihre Geschichte bitte in das dafür vorgesehene Feld neben der Abbildung. Blättern Sie anschließend um und fahren Sie mit dem nächsten Bild in gleicher Weise fort.

Sollte der bereitgestellte Platz einmal nicht ausreichen, verwenden Sie bitte die Rückseite des entsprechenden Blattes.

Bitte versuchen Sie, deutlich zu schreiben.

Ihre Geschichten ebenso wie Ihre übrigen Angaben fließen absolut anonymisiert in eine gruppenstatische Auswertung (also keine Einzelfallbetrachtung) ein.

Ihrer Fantasie sind keine Grenzen gesetzt.

Bringen Sie dieses Heft gemeinsam mit den anderen Unterlagen, die Sie mit nach Hause bekommen haben, zum letzten Teil der Untersuchung unbedingt wieder mit.











**A.6.2: „Laborset“: Explizite Einstellungs- & Motivationskriterien**



seit 1558

Vom Versuchsleiter auszufüllen: Vpnr:.....

Form B

Institut für Psychologie  
Humboldtstr. 11  
07743 Jena

Friedrich-Schiller-Universität Jena

**Ansprechpartner:**

Dipl. Psych. Nicole Kämpfe und Dr. phil. Kristin Mitte

Tel.: (03641) 9-45163 oder 9-45164

E-Mail: [Nicole.Kaempfe@uni-jena.de](mailto:Nicole.Kaempfe@uni-jena.de) oder [mail@kristin-mitte.de](mailto:mail@kristin-mitte.de)

Liebe Teilnehmer,

im Anschluss an die Experimente erhalten Sie nun noch einen Fragebogen mit Aussagen, denen man so oder so ähnlich im Alltag begegnet und denen man zustimmend oder ablehnend gegenüberstehen kann.

Bei der Beurteilung solcher Aussagen können eigene und allgemeine Meinung voneinander abweichen. Es kommt häufig vor, dass man sich in seinem Standpunkt zu manchen Themen mehr oder weniger stark von der Mehrzahl der Menschen unterscheidet. Dies wollen wir mit vorliegendem Fragebogen berücksichtigen und möchten Sie daher bitten, jede Aussage in zweifacher Hinsicht zu beurteilen: Zum Einen sind wir daran interessiert, wie Sie **persönlich** darüber denken. Zum Anderen möchten wir Sie bitte, die **allgemeine gesellschaftliche Meinung** zu den präsentierten Aussagen einzuschätzen. Für jede Beurteilungsperspektive – also Ihre eigene Sicht und die der Allgemeinheit – steht Ihnen eine Antwortspalte zur Verfügung.

Um die Aussagen zu bewerten, können Sie immer zwischen fünf Antwortmöglichkeiten von völliger Ablehnung bis hin zu vollkommener Zustimmung wählen. Die Antwortkästchen sind für eine bessere Übersichtlichkeit unterschiedlich schattiert:

Helleres Grau signalisiert *Ablehnung*, Dunkleres Grau signalisiert *Befürwortung* der jeweiligen Aussage.

Auch wenn Sie bei einer Aussage unsicher sein sollten: Kreuzen Sie einfach spontan und offen an, was einerseits Ihrem **persönlichen** Empfinden und andererseits der **Mehrheitsmeinung** am ehesten entspricht.

Wir bitten Sie, keine Antwort auszulassen.

Bei der Behandlung des Fragebogens folgen wir selbstverständlich den Maßgaben des Datenschutzes; mit all Ihren Daten wird völlig anonymisiert verfahren.

Bitte beurteilen Sie folgende Aussagen zum einen nach Ihrer persönlichen Meinung und zum anderen gemäß der Meinung innerhalb unserer Gesellschaft	ICH					UNSERE GESELLSCHAFT				
	lehne ich vollkommen ab		weder... noch...		stimme ich vollkommen zu	lehnt man vollkommen ab		weder... noch...		stimmt man vollkommen zu
1. Übergewichtige Menschen sind genauso attraktiv wie Normalgewichtige	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Selbst wenn ich versuche, es zu unterdrücken: Wenn ich eine wirklich dicke Person sehe, finde ich sie irgendwie abstoßend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Im Durchschnitt sind dicke Menschen weniger fleißig als andere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Ich habe schon oft Sympathie für dicke Menschen empfunden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Ich fände es entsetzlich, würde ich selbst übergewichtig werden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Dicke Menschen sind in ihrer Arbeit meist passiv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Übergewicht wird in der Regel durch einen Mangel an Willenskraft verursacht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Ich kann gut nachvollziehen, dass die meisten Normalgewichtigen keine Liebesbeziehung zu einer übergewichtigen Person wollen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Niemand muss dick sein; wer dick ist, ist selbst dafür verantwortlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Ich halte übergewichtige Menschen im beruflichen Kontext für nicht besonders durchsetzungsfähig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Ich fühle mich unbehaglich in der Nähe von Dicken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Dicke sollten stärker an den Kosten beteiligt werden, die sie dem Gesundheitswesen verursachen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Übergewichtige Menschen verfügen bei ihrer Arbeit erfahrungsgemäß nur über sehr wenig Ausdauer und Durchhaltevermögen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Normalgewichtige Mitarbeiter sind meist ehrgeiziger als ihre übergewichtigen Kollegen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Ich könnte mir vorstellen, ein Liebesverhältnis mit einer deutlich übergewichtigen Person zu haben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Es ist abstoßend, wenn eine wirklich dicke Person am Strand in Badebekleidung herumläuft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Die meisten Dicken sind selbst Schuld an ihrem Übergewicht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Anstatt sich über mangelnde Eleganz von Übergroßen-Kleidung zu beklagen, sollten Dicke ihre Energie lieber nutzen, um abzunehmen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Übergewichtige erhalten im Gesundheitswesen mehr Aufmerksamkeit, als sie bekommen sollten (z.B. durch Diätprogramme, Ernährungsberatung usw.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Beurteilen Sie bitte auch die folgenden Aussagen in zweifacher Hinsicht

ICH

UNSERE GESELLSCHAFT

	ICH					UNSERE GESELLSCHAFT				
	lehne ich vollkommen ab		weder... noch...		stimme ich vollkommen zu	lehnt man vollkommen ab		weder... noch...		stimmt man vollkommen zu
20. Homosexuelle Männer führen ein zu ausschweifendes und hemmungsloses Leben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. Schwule Männer sind für leitende Funktionen weniger geeignet als heterosexuelle Menschen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. Schwule Paare sollten das Recht erhalten zu heiraten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. Schwule Männer sollten nicht mit Kindern oder Jugendlichen arbeiten dürfen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Homosexualität ist wie Heterosexualität eine gleichwertige Form der menschlichen Sexualität	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Ich könnte mir vorstellen, eine homosexuelle Partnerschaft einzugehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Ich würde einen schwulen Mann in ein öffentliches Amt wählen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. Wenn ein Mann homosexuelle Gefühle hat, sollte er alles tun, um diese zu überwinden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Homosexualität ist nur eine andere Form des Lebensstils, die nicht verurteilt werden sollte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. Ich finde es gut, wenn Menschen ihre Homosexualität ausleben können	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. Schwulen Paaren sollte es genauso wie heterosexuellen Paaren erlaubt werden, Kinder zu adoptieren	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Schwule sind einfach nicht fähig, eine normale Beziehung zu führen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. Wenn mir jemand sagt, er sei schwul, kann ich gut damit umgehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. Ich möchte nicht, dass Schwule mit ihrer Sexualität in die Öffentlichkeit gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34. Wenn mir mein Kind gesteht, dass es homosexuell ist, würde ich es auf diesem Weg unterstützen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35. Ich könnte mir ein sexuelles Verhältnis zu einer gleichgeschlechtlichen Person vorstellen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36. Homosexualität ist krankhaft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37. Menschen mit homosexueller Orientierung sind in meinen Augen eine diskriminierte Minderheit in unserer Gesellschaft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

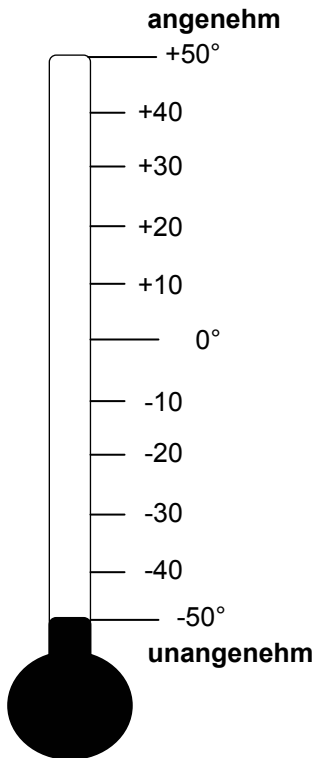
Nachfolgend werden Ihnen kurz neun **Situationen** präsentiert. Versuchen Sie sich, so gut es geht, in die jeweilige Lage hineinzusetzen und geben Sie an, wie angenehm bzw. unangenehm die Situation

a) für Sie wäre und b) für den Durchschnittsbürger unserer Gesellschaft wäre.

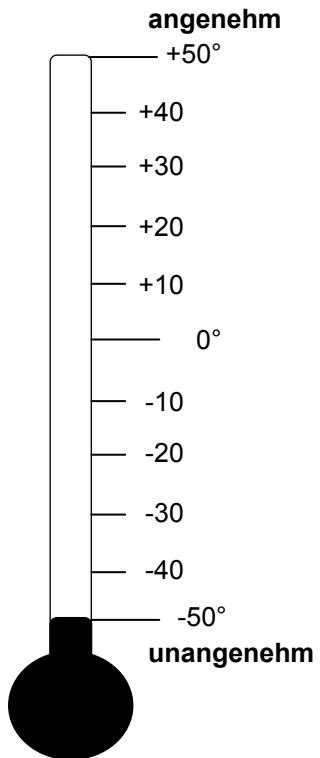
Bitte beurteilen Sie folgende Situationen nach Ihrem persönlichen Empfinden und entsprechend dem Empfinden innerhalb unserer Gesellschaft	ICH					ANDERE				
	wäre mir sehr angenehm		weder... noch...		wäre mir sehr unangenehm	wäre anderen sehr angenehm		weder... noch...		wäre anderen sehr unangenehm
38. Im Bus setzt sich eine übergewichtige Frau neben mich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39. Eine übergewichtige Person lädt mich zu einem Rendezvous ein	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
40. Eine dicke Bekannte möchte mit mir gemeinsam einen Einkaufsbummel machen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
41. Die Kollegin, mit der ich ein Büro teile, ist stark übergewichtig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
42. Eine übergewichtige Person betritt die Sauna, in der ich mich entspanne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
43. An meinem Sportkurs nehmen mehr als die Hälfte übergewichtige Leute teil	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
44. Vor mir auf einer schmalen Treppe läuft ein dicker Mann	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
45. Auf einer Party bittet mich eine übergewichtige Person, mit ihr zu tanzen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
46. In der Warteschlange im Supermarkt steht eine dicke Person direkt vor mir	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47. Im Restaurant setzt sich eine übergewichtige Familie an den Nachbartisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
48. Ich bin in einer Gruppe von schwulen Männern	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
49. Ein schwuler Mann/eine lesbische Frau verliebt sich in mich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
50. Ein Freund will mit mir in eine schwules Café gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
51. Ich werde mit einem offensichtlich schwulem Mann in der Öffentlichkeit gesehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
52. Ich erfahre, dass ein Freund von mir schwul ist	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
53. Ein schwules Paar küsst sich in meiner Nähe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
54. Mein Sohn erzählt mir, dass er schwul ist	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
55. Ich sehe zwei Männer in der Öffentlichkeit Hand in Hand gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
56. Ich erfahre, dass der Lehrer meines Kindes homosexuell ist	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

57.) Bitte geben Sie auf den folgenden Thermometerskalen an, bis zu welchem Grad Ihnen die genannten Personengruppen vergleichsweise angenehm bzw. unangenehm sind. Dabei stehen +50° für maximal angenehm und -50° für maximal unangenehm. Füllen Sie das Thermometer so aus, dass es Ihr persönliches Empfinden über-, unter- und normalgewichtigen sowie homo- und heterosexuellen Menschen gegenüber widerspiegelt.

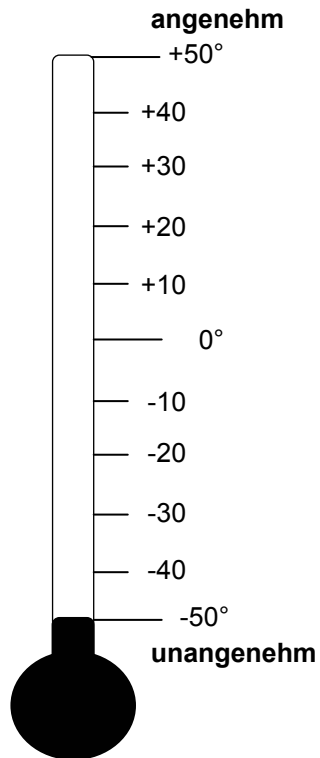
a) übergewichtige Menschen sind mir...



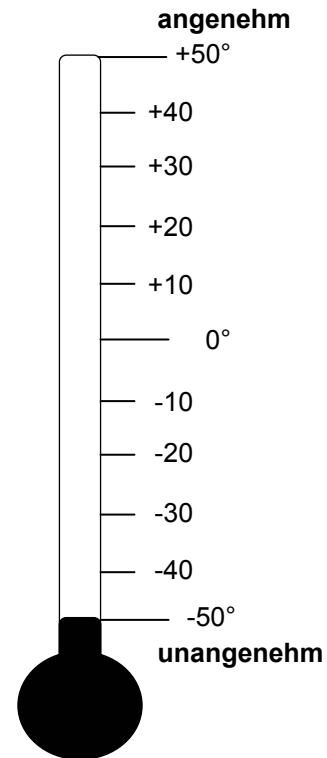
b) normalgewichtige Menschen sind mir...



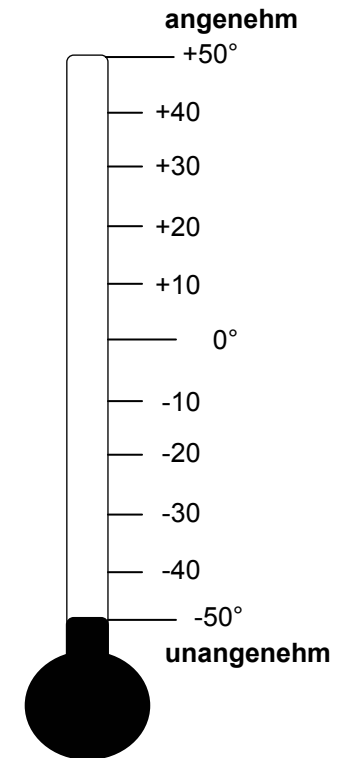
c) untergewichtige Menschen sind mir...



d) homosexuelle Menschen sind mir...



e) heterosexuelle Menschen sind mir...





Bitte beurteilen Sie folgende Aussagen zum einen nach Ihrer persönlichen Meinung und zum anderen gemäß der Meinung in unserer Gesellschaft	ICH					UNSERE GESELLSCHAFT				
	lehne ich vollkommen ab		weder... noch...		stimme ich vollkommen zu	lehnt man vollkommen ab		weder... noch...		stimmt man vollkommen zu
58. Es ist einfach, den Ärger sozialer Minderheiten zu verstehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
59. Während der letzten Jahre haben manche Minderheiten wirtschaftlich mehr erhalten, als ihnen zusteht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
60. In ihren Bestrebungen für Gleichberechtigung sind die meisten sozialen Minderheiten noch nicht fordernd genug	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
61. Einige Minderheiten haben zu viel gesellschaftlichen Einfluss	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
62. Während der letzten Jahre haben Regierung und Medien mehr Respekt gegenüber manchen Minderheitsgruppen gezeigt, als sie verdienen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
63. Mitglieder sozialer Randgruppen sollten nicht sich dort hinein drängen, wo sie nicht erwünscht sind	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
64. Diskriminierung von Minderheiten ist ein Problem in Deutschland	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
65. Vorurteile gegenüber Minderheiten beruhen meist auf Tatsachen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
66. Die meisten Leute versuchen, vorurteilsfrei gegenüber Minderheiten zu erscheinen, weil es den heutigen Standards entspricht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
67. Weil es für mich persönlich wichtig ist, versuche ich, mich ohne Vorurteile gegenüber anderen zu verhalten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
68. Ich versuche negative Gedanken gegenüber Minderheiten zu verbergen, um ablehnende Reaktionen anderer Leute zu vermeiden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
69. Ich glaube fest daran, dass man keine Vorurteile gegenüber anderen Gruppen haben sollte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
70. Wenn ich mich Minderheiten gegenüber vorurteilsvoll verhalten würde, wäre ich besorgt, den Ärger anderer auf mich zu ziehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
71. Ich bin davon überzeugt, dass es in Ordnung ist, gewisse Vorurteile gegenüber Minderheiten zu haben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
72. Ich vermeide Vorurteile, weil das heutzutage so erwartet wird	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
73. Ich persönlich finde es falsch, Vorurteile zuzulassen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Fahren Sie bitte fort, folgende Aussagen nach Ihrer persönlichen Meinung und gemäß der Meinung in unserer Gesellschaft zu beurteilen**

		ICH					UNSERE GESELLSCHAFT					
		lehne ich vollkommen ab		weder... noch...		stimme ich vollkommen zu		lehnt man vollkommen ab		weder... noch...		stimmt man vollkommen zu
74.	Ich versuche, so wenig Vorurteile wie möglich zu zeigen, damit andere mich nicht ablehnen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
75.	Ohne Vorurteile zu sein, ist wichtig für meine Meinung über mich selbst	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
76.	Ich vermeide Vorurteile aufgrund des Drucks, den andere deswegen ausüben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
77.	Ich vermeide Vorurteile, weil ich es so von mir erwarte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
78.	Man sollte sich nie durch Vorurteile leiten lassen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
79.	Ich achte darauf, dass mein Verhalten nicht durch Vorurteile beeinflusst wird	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
80.	Es ist mir egal, wenn jemand glaubt, dass ich Vorurteile gegenüber Minderheiten habe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
81.	Wenn man über Minderheiten spricht, sollte man abwertende Bezeichnungen vermeiden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
82.	Es lohnt sich nicht, sich ständig darüber Sorgen zu machen, ob man sich gerade irgendwem gegenüber vorurteilsvoll verhält	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
83.	Jeder Mensch hat Vorurteile. Es kommt drauf an, sich nicht davon leiten zu lassen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
84.	Man sollte sich seine eigenen Vorurteile bewusst machen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
85.	Man sollte sich besonders fair verhalten, wenn man mit jemandem zu tun hat, der vermutlich häufig unter Vorurteilen zu leiden hat	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
86.	Man sollte in Gesellschaft nichts Negatives über Minderheiten sagen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
87.	Ich ärgere mich über mich selbst, wenn ich etwas denke oder fühle, was für vorurteilsvoll gehalten werden könnte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
88.	Man sollte nicht über Witze lachen, die auf Kosten von Minderheiten gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
89.	Es wäre mir unangenehm, wenn jemand glauben würde, dass ich Vorurteile gegenüber Minderheiten hätte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
90.	Es macht mich wütend, wenn jemand Vorurteile über Minderheiten äußert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
91.	Ich finde es wichtiger zu sagen, was man denkt, als sich ständig darüber Sorgen zu machen, ob man jemandem damit zu nahe tritt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
92.	Es ist mir sehr unangenehm, jemanden zu verletzen, daher versuche ich immer, Rücksicht auf die Gefühle anderer zu nehmen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Fahren Sie bitte mit den letzten Aussagen fort, sie gemäß Ihrer persönlichen Meinung und der Meinung in unserer Gesellschaft zu beurteilen**

		ICH					UNSERE GESELLSCHAFT						
		lehne ich vollkommen ab		weder... noch...		stimme ich vollkommen zu			lehnt man vollkommen ab		weder... noch...		stimmt man vollkommen zu
93.	Wenn ich Gedanken oder Gefühle habe, die andere diskriminieren, behalte ich sie für mich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
94.	Bei einem Gespräch mit einer homosexuellen Person, versuche ich nicht über seine/ ihre Homosexualität nachzudenken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
95.	Bei der Beurteilung von fremden Personen ist es mir wichtig, mich nicht von äußeren Merkmalen leiten zu lassen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
96.	Mir persönlich ist es wichtig, Menschen mit einer körperlichen oder geistigen Behinderung möglichst offen zu begegnen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
97.	Wenn ich einem Freund eine mir bekannte Person beschreibe, versuche ich, ihre Nationalität nicht zu erwähnen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
98.	Ich möchte nicht ausländerfeindlich oder sexistisch erscheinen, auch mir selbst gegenüber	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
99.	Wenn ich einer Person anderer Nationalität begegne, versuche ich nicht, über ihre andere Herkunft nachzudenken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
100.	Ich möchte keine abwertenden Begriffe über fremde Nationalitäten verwenden, selbst, wenn mir solche Worte durch den Kopf gehen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Nun möchten wir Sie bitten, zu verschiedenen Fragen eine Schätzungen abzugeben.**

**Diese Fragen wurden bei einer repräsentativen Meinungsumfrage in der deutschen Bevölkerung gestellt.  
Die Angaben der Befragten sehen Sie neben jeder Frage.**

**Lesen Sie sich bitte jede Frage durch und geben Sie eine eigene Schätzung an,  
so als ob Sie die präsentierten Ergebnisse nicht kennen würden.**

**Tragen Sie ihre Antwort in die letzte, grau schattierte Spalte ein.**

	Umfrageergebnis	Ihre Schätzung
101. Wie viel Prozent der Frauen geben an, regelmäßig Eifersucht ihrem Partner gegenüber zu verspüren?	62	%
102. Wie viele Minuten am Tag lachen dicke Menschen durchschnittlich?	43	min
103. Wie viel Prozent der Homosexuellen gehen einem künstlerischen Beruf nach?	25	%
104. Wie hoch ist der Anteil an Krankschreibungen, die auf Sportunfälle zurückzuführen sind?	13	%
105. Um wie viele Jahre verringert sich die durchschnittliche Lebenserwartung aufgrund von starkem Übergewicht und dessen Folgeerscheinungen?	11	Jahre
106. Wie hoch ist der Anteil an Frauen in handwerklichen Berufen?	16	%
107. Wie hoch ist der Anteil an Partnerschaften unter Homosexuellen, die weniger als 3 Monate dauern?	60	%
108. Wie viel Prozent der Übergewichtigen geben an, ihr Leben voll und ganz zu genießen?	75	%
109. Wie viel Prozent der Frauen fühlen sich von schwulen Männern besser verstanden als von einer Freundin?	33	%
110. Wie viel Prozent der Ehen in Deutschland werden vor dem 15. Ehejahr geschieden?	41	%
111. Wie viel Prozent der Homosexuellen geben an, in ihrer Kindheit schwere Konflikte der Eltern erlebt zu haben?	54	%
112. Wie viel Prozent der Deutschen befinden sich derzeit in psychotherapeutischer Betreuung?	10	%
113. Wie hoch ist der Anteil unter den übergewichtiger Menschen, die kein Abitur haben?	82	%
114. Wie viel Prozent der deutschen Studenten geben an, einer ehrenamtlichen Tätigkeit nachzugehen?	35	%
115. Wie hoch ist der Anteil Homosexueller, der angibt, sich in seinem Freundes- und Bekanntenkreis rundum wohl zu fühlen?	48	%
116. Wie viel Prozent der Männer würden lieber eine dicke Partnerin haben?	29	%
117. Wie häufig denkt ein Mensch am Tag durchschnittlich ans Essen?	18	Mal
118. Wie hoch ist der Anteil Schwuler, der lieber heterosexuell orientiert sein würde?	31	%
119. Wie viel Prozent der Männer geben an, mit ihrem Körper zufrieden zu sein?	23	%
120. Wie hoch ist die prozentuale Belastung unseres Gesundheitssystems infolge von Übergewicht?	37	%

**Bitte geben Sie nun noch eine Schätzung über den prozentualen Anteil folgender Personen in unserer Gesellschaft ab:**

121. untergewichtige Männer: .....%	übergewichtige Männer: .....%	untergewichtige Frauen: .....%	übergewichtige Frauen: .....%
homosexuelle Männer: .....%	bisexuelle Männer: .....%	homosexuelle Frauen: .....%	bisexuelle Frauen: .....%
hoch ängstliche Männer: .....%	niedrig ängstliche Männer: .....%	hoch ängstliche Frauen: .....%	niedrig ängstliche Frauen: .....%

**Zum Abschluss möchten wir noch ein paar Angaben zu Ihrer Person erbitten. Lassen Sie bitte keine Frage aus!  
Auch die folgenden Angaben gehen ausschließlich in eine *gesamtstatistische* Auswertung ein.**

122. **Geschlecht:**  Weiblich  Männlich

123. **Alter:**

124. **Höchster Bildungsabschluss:**

- |                                                          |                                                             |                                                                     |
|----------------------------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------------|
| <input type="checkbox"/> Ohne Abschluss                  | <input type="checkbox"/> Fachhochschulreife                 | <input type="checkbox"/> Hochschulabschluss (z.B. Magister, Diplom) |
| <input type="checkbox"/> Hauptschulabschluss (8. Klasse) | <input type="checkbox"/> Allgemeine Hochschulreife (Abitur) | <input type="checkbox"/> Höherer akademischer Grad (z.B. Promotion) |
| <input type="checkbox"/> Realschulabschluss (10. Klasse) | <input type="checkbox"/> Berufsausbildung                   |                                                                     |

**Wo sind Sie aufgewachsen?**

125.  alte Bundesländer  neue Bundesländer  nicht in Deutschland

126. **Falls zutreffend,  
was studieren Sie?**

127. **In welchem Semester sind Sie?**

128. **Sind Sie Rechts- oder Linkshänder?**  rechtshändig  linkshändig  beidhändig

129. **Haben Sie eine Rot-Grün-Sehschwäche?**

ja  nein  nein, aber eine andere Sehschwäche, nämlich:

130. **Haben Sie bereits an ähnlichen Laborexperimenten wie den vorhergehenden teilgenommen?**  ja  nein

131. **Wie bezeichnen Sie Ihre sexuelle Orientierung?**  Heterosexuell  Homosexuell  Bisexuell

132. **Wie viele Ihrer Freunde und Bekannten sind homo- oder bisexuell?**

<input type="checkbox"/> keine	<input type="checkbox"/> etwa 3 bis 5	<input type="checkbox"/> etwa 8 bis 10
<input type="checkbox"/> etwa 1 bis 2	<input type="checkbox"/> etwa 6 bis 8	<input type="checkbox"/> mehr als 11

133. **Ihre Körpergröße in cm:**

**Ihr Gewicht in kg:**

134. **Wie viele Ihrer Freunde und Bekannten sind übergewichtig?**

<input type="checkbox"/> keine	<input type="checkbox"/> etwa 3 bis 5	<input type="checkbox"/> etwa 8 bis 10
<input type="checkbox"/> etwa 1 bis 2	<input type="checkbox"/> etwa 6 bis 8	<input type="checkbox"/> mehr als 11

**Vielen Dank für Ihre Teilnahme!**

## A.6.2: Kodierschema der offenen Assoziativmessung in Studie III

Kodiervariablen	Variablenname	Skala & Wertelabel
<b>Wortanzahl</b>	„Anzahl“	kontinuierlich
<b>Wertigkeit der Geschichte</b> - Wortwahl - Sarkasmus/ Zynismus/ Spott - positiv/ negativ besetzte Themen	„Valenz“	1 = negativ 2 = leicht negativ 3 = neutral 4 = leicht positiv 5 = positiv
<b>Individueller Vorurteilsgehalt</b> persönliche Vorurteile des Autors	„VOR <sup>x</sup> “	1 = keine eigenen Vorurteile 2 = wenig ~ 3 = mittel ausgeprägte ~ 4 = deutlich erkennbare ~ 5 = starke ~
<b>Antizipierter Vorurteilsgehalt</b> Thematisierte gesellschaftliche Vorurteile	„VOR <sup>d</sup> “	1 = keine antizipierten Vorurteile 2 = wenig ~ 3 = mittel ausgeprägte ~ 4 = deutlich erkennbare ~ 5 = starke ~
<b>Stereotyp erkannt/thematisiert</b>	„Stereo“	1 = Stereotyp nicht erkannt/ thematisiert 2 = nicht eindeutig bzw. verdeckt ~ 3 = Stereotyp eindeutig ~
<b>Kreativität der Geschichte</b>	„Kreativ“	1 = sehr un kreativ 2 = un kreativ 3 = mittelmäßig kreativ 4 = kreativ 5 = sehr kreativ

Tabelle A.6.2, Kodierschema für die Geschichten der offenen Assoziativmessung

## Anhang A.6.4: Systematischer Modellvergleich MTMM-Analyse nach (Widaman, 1985; siehe Abschnitte 6.4.4 und 6.4.6)

### Grundlagen

Die Systematik hierarchisch genesteter Strukturgleichungsmodelle von Widaman (1985) ermöglicht durch sukzessive Vergleiche ineinander genesteter Modelle, die Validierung von Trait- und Methodenfaktoren. Dabei wird der Nutzen einer Einbuße von Freiheitsgraden aufgrund der Modellierung mehrerer Trait- bzw. Methodenfaktoren getestet. Ist, ausgehend von Modellen einfacher Struktur (General faktormodell), der Verlust an Freiheitsgraden durch die Einführung von Modellkomponenten im Sinne der Konstruktvalidität gerechtfertigt, so können die modellierten Faktoren als konvergent valide bezeichnet werden. Der entsprechende Modellvergleich erfolgt über die Differenz der  $\chi^2$ -Werte zweier ineinander genesteter Modelle. Eine nicht signifikante  $\chi^2$ -Differenz führt zur Beibehaltung des sparsameren Modells. Die Restringiertheit der Modelle nimmt von A nach C bzw. von 1 nach 3 ab und ihre Komplexität steigt. Modelle, die sich links oder oberhalb von anderen Modellen befinden sind also in die rechts bzw. unter ihnen stehenden Modelle genestet.

		Methodenstruktur							
		A		B		B'		C	
Traitstruktur	1	Nullmodell		1 MF		$m$ MF (orth.)		$m$ MF (obl.)	
	2	1 TF		1 TF + 1 MF		1 TF + $m$ MF (orth.)		1 TF + $m$ MF (obl.)	
	2'	$t$ TF (orth.)		$t$ TF (orth.) + 1 MF		$t$ TF (orth.) + $m$ MF (orth.)		$t$ TF (orth.) + $m$ MF (obl.)	
	3	$t$ TF (obl.)		$t$ TF (obl.) + 1 MF		$t$ TF (obl.) + $m$ MF (orth.)		$t$ TF (obl.) + $m$ MF (obl.)	

**Abbildung A.6.4** MTMM Analyse mit Strukturgleichungsmodellen nach Widaman (1985, S.6), TF: Traitfaktor, MF: Methodenfaktor, orth.: orthogonal (unkorreliert), obl.: oblique (korreliert); bei vorliegender Datenstruktur:  $t=2$ ;  $m=4$

Bedeutsam in einer latenten MTMM-Analyse ist neben dem globalen Vergleich verschiedener Modellierungen die Betrachtung der Parameter des endgültigen Modells. Notwendig für *konvergente Validität* der interessierenden Konstrukte, hier also Einstellungen zu Übergewicht und Homosexualität, sind signifikante Ladungen der latenten Trait-Faktoren auf die jeweiligen Indikatoren, die in ihrem Betrag die Ladungen der Methodenfaktoren übersteigen sollten. Die Produkte dieser Ladungen entsprechen den Schätzungen der konvergenten Validitätskoeffizienten (z.B. Pedhazur & Schmelkin, 1991). Als *diskriminant valide* werden die Konstrukte bezeichnet, wenn die Korrelation ihrer latenten Trait-Faktoren verschieden von 1 ist, bzw. keine Konvergenz in einem gemeinsamen Faktor vorliegt (Schmitt & Stults, 1986; Widaman, 1985). Cote, Buckley & Best (2001) raten zusätzlich zur Betrachtung des Anteils erklärter Varianz an der Gesamtvarianz durch die Trait- und Methodenfaktoren. Diese sollte größer sein, als die nicht erklärte Varianz, die sich in den Fehlertermen niederschlägt.

### *Modellvergleich für vorliegende Fragestellung*

Im Folgenden werden zunächst das Nullmodell (1A) und das Generalfaktormodell (2A) als Ausgangsbasis berechnet und miteinander verglichen (Modellvergleich 1). Die konvergente Validität der Einstellungstrait wird über die Vergleiche des Generalfaktormodells mit den Modellen orthogonaler (2'A) und korrelierter Traits (3A) geprüft (Modellvergleiche 2a und 2b). Die Korreliertheit der Traitfaktoren prüft Modellvergleich 2c; von diesem sollte jedoch keine endgültige Aussage abgeleitet werden. Die konvergente Validität der Methodenfaktoren wird analog zu 2a und 2b über die Vergleiche des Generalfaktormodells mit den Modellen orthogonaler (1B') und korrelierter Methodenfaktoren (1C) getestet (Modellvergleiche 3a und 3b). Die Korreliertheit der Methodenfaktoren wird in Modellvergleich 3c getestet, wobei das Fehlen der Traitfaktoren auch hier keine endgültige Aussage über die Beziehungen zwischen den Methodenfaktoren zulässig macht. Eine enge Beziehung zwischen den erfassten Einstellungstrait würde sich an dieser Stelle in hohen Korrelationen der Methodenfaktoren niederschlagen. Die Bedeutung von Trait- *und* Methodenfaktoren untersucht Modellvergleich 4a mit der Gegenüberstellung von Modell 2'A und 2'B' (orthogonal) sowie 4b mit dem Vergleich von Modell 3A und 1C (oblique). Die Korreliertheit von Trait- und Methodenfaktoren prüft Modellvergleiche 5.

#### **Exkurs A.6.4**

##### **– Die Problematik des CFA-Ansatzes zur latenten MTMM Analyse –**

Die für den systematischen Vergleich nach Widaman (1985) betrachtete MTMM-Modellgruppe ist mit Correlated Trait-Correlated Method (CT-CM) Modelle benannt. Sie zeichnet sich durch die Modellierung  $t$  erhobener Traitfaktoren und  $m$  eingesetzter Methodenfaktoren aus, wobei keine Korrelationen zwischen Trait- und Methodenfaktoren vorgesehen sind. Vor der schrittweisen Analyse der Modelle muss auf die Problematik der CT-CM Modelle hingewiesen werden.

Die CT-CM Modelle sind mit  $m+t$  Informationen unteridentifiziert (Kenny & Kashy, 1992), was bei der Modellierung des kompletten CT-CM-Modells (Modell 3C) zu akuten Identifikations- und Schätzbarkeitsproblemen führt. Die Modelle sind entweder grundsätzlich nicht identifiziert oder sie weisen unplausible Parameterschätzungen wie negative Varianzen oder standardisierte Pfadkoeffizienten größer als 1 – so genannte „Heywood Cases“ – auf (Kenny & Kashy, 1992; Marsh, 1989; Grayson & Marsh, 1994). Die Wahrscheinlichkeit für das Auftreten solcher Heywood cases ist in kleinen Stichproben ( $N < 400$ ) erhöht (Marsh & Bailey, 1991; Lance, Noble & Scullen, 2002).

Ein weiteres unerwünschtes Ergebnis der CT-CM Modellierung ist die mögliche Überschätzung der Methodenvarianz und somit eine Unterschätzung der Reliabilität der Indikatoren (Eid, Lischetzke, Nussbeck & Trierweiler, 2003; Marsh, 1989; vgl. auch Eid, 2000; Kenny & Marsh, 1992; Kline 1998; Lievens & Conway, 2001; Maruyama, 1998; Wothke, 1984).



### Ergebnisse des systematischen Modellvergleichs

Für die aktuelle Datenmodellierung sind starke Restriktionen der Modelle 2A bis 1C vonnöten, um dem Identifikationsproblem (siehe Exkurs 6-1) entgegen zu wirken und eine Vergleichbarkeit der Modelle zu gewährleisten, während diese Restriktionen in den Modellen 2'B' bis 3C, in denen Trait- und Methodenfaktoren modelliert werden, zu Schätzbarkeitsproblemen führen. Für das am wenigsten restringierte Modell 3C, das korrelierten Methoden- und Traitfaktoren nahe legt und gemäß dem systematischen Modellvergleich die Daten am besten repräsentiert (Tab. A.6.4-1), werden schließlich „Out-of-Range“-Parameterschätzungen erhalten (Abb. 6-33 in Abschnitt 6.4.4), die keine gültige Interpretation des Modells erlauben. Der Vergleich der CT-CM Modelle nach Widaman bietet demzufolge nach im vorliegenden Fall keine hinreichenden Informationen zu Konstruktvalidierung der Einstellungsmaße.

Modellstatistik		$\chi^2$	df	$\chi^2/df$	<i>p</i>	RMSEA	NFI
1A	(Nullmodell)	562.826	120	4.690	.000	.150	.000
2A	(Generalfaktor)	403.666	119	3.392	.000	.120	.283
2'A	(TF orthogonal)	314.200	118	2.663	.000	.100	.442
3A	(TF oblique)	286.814	117	2.451	.000	.094	.490
1B'	(MF orthogonal)	342.493	116	2.953	.000	.109	.391
1C	(MF oblique)	324.417	110	2.949	.000	.109	.424
2'B'	(TF & MF orth.)	136.509	92	1.484	.002	.054	.757
3B'	(TF obl. & MF orth.)	136.130	91	1.496	.002	.055	.758
2'C	(TF orth. & MF obl.)	114.452	86	1.331	.022	.045	.797
3C	(TF & MF obl.)	110.378	85	1.299	.034	.043	.804

Modellvergleich		$\Delta\chi^2/ \Delta df$	
1	Basisanalyse konvergenter Validität	2A gegen 1A	159.16/ 1 ***
		2a) 2A gegen 2'A	89.47/ 1 ***
2	Bedeutung der Traitfaktoren	2b) 2A gegen 3A	116.85/ 2 ***
		2c) 2'A gegen 3A	27.39/ 1 ***
		3a) 2A gegen 1B'	61.17/ 3 ***
3	Bedeutung der Methodenfaktoren	3b) 2A gegen 1C	79.25/ 8 ***
		3c) 1B' gegen 1C	27.39/ 1 ***
		4a) 2'A gegen 2'B'	177.69/ 26 ***
4	Bedeutung von Trait- und Methodenfaktoren	4b) 3A gegen 3C	176.44/ 32 ***
		5a) 2'B' gegen 3B'	.04/ 1 n.s.
5	Korreliertheit von Trait- und Methodenfaktoren	5b) 2'B' gegen 2'C	22.06/ 6 **
		5c) 2'C gegen 3C	4.07/ 1 *

**Tabelle A.6.4-1**, Systematischer Modellvergleich nach Widaman (1985), \*\*\*:  $p < .001$ ; \*\*:  $p < .01$ , \*:  $p < .05$ , n.s.: nicht signifikant

### Modifizierter Modellvergleich nach CFA-Ansatz

Mit den Modifikationen auf Indikatorebene (vgl. Abschnitt 6.4.6) wird nachfolgend erneut ein Modellvergleich in Anlehnung an die Systematik von Widaman (1985) vorgenommen. Das Vorgehen entspricht dem oben dargestellten für die vollständigen Daten, der *Modellvergleich* wird an dieser Stelle jedoch nicht dokumentiert. Die Gütekriterien der geprüften Modelle berichtet Tabelle A.6.4-2.

Modellstatistik	$\chi^2$	df	$\chi^2/df$	<i>p</i>	RMSEA	NFI	R <sup>2</sup>
1A (Nullmodell)	606.34	66	9.187	.000	.223	.000	
2A (Generalfaktor)	359.24	54	6.653	.000	.185	.408	8.22%
2'A (TF orthogonal)	239.41	54	4.433	.000	.144	.605	14.36%
3A (TF oblique)	222.05	53	4.190	.000	.139	.634	13.26%
1B' (MF orthogonal)	248.40	54	4.600	.000	.148	.590	13.62%
1C (MF oblique)	220.08	51	4.315	.000	.142	.637	17.44%
2'B' (TF & MF orth.)	67.80	42	1.614	.007	.061	.888	16.35%
3B' (TF obl. & MF orth.)	70.05	41	1.708	.003	.066	.884	18.50%
2'C (TF orth. & MF obl.)	54.87	39	1.407	.047	.050	.909	18.09%
3C (TF & MF obl.)	51.31	38	1.35	.073	.046	.915	19.90%

**Tabelle A.6.4-2**, Systematischer Modellvergleich für die Einstellungsmaße direkter Messung, offener Assoziativmessung und indirekter IAT-Messung

er alleinigen Berücksichtigung der  $\chi^2$ -Werte wäre dem Modellvergleich zufolge ein Modell korrelierender Methoden- und orthogonaler Traitfaktoren (Modell 2'C) anzunehmen, da die entsprechende Veränderung von Modell 2'C auf 3C nicht signifikant ist ( $\Delta\chi^2=3.56$ ,  $\Delta df=1$ ,  $p>.05$ ). Jedoch weisen allein die Einstellungstrait eine signifikante Beziehung zueinander auf, während in keinem der kompletten Modelle Korrelationen der Methodenfaktoren statistisch bedeutsam sind. Zudem weist Modell 3C keine signifikante  $\chi^2$ -Statistik auf ( $\chi^2=51.31$ ,  $df=38$ ,  $p<.05$ ) und erzielt mit etwa 20% die höchste Varianzaufklärung durch die latenten Trait- und Methodenfaktoren. Es tritt kein Heywood Case auf, jedoch sind die Varianzen des Einstellungsfaktors „Homo“ ( $s^2=.027$ ,  $p>.110$ ) sowie die Methodenfaktoren „direkt“ ( $s^2=.000$ ,  $p=1.00$ ) und „offen“ ( $s^2=.162$ ,  $p>.120$ ) statistisch nicht bedeutsam. Dies sind Hinweise auf eine vorliegende Überfaktorisation. Der Anteil erklärter Varianz durch Trait- und Methodenfaktoren beträgt 51.55% und übersteigt somit den Anteil nicht erklärter Varianz.<sup>1</sup>

Auch in diesem modifizierten Modell liegt keine befriedigende Konvergenz der verschiedenen Einstellungsmaße in jeweils einem Einstellungsfaktor vor: Alle Pfadladungen der latenten Traits zu den IAT-Indikatoren sind nicht signifikant, wogegen die methodenspezifische Varianz („iat“) substantiell ist ( $s^2=.011$ ,  $p<.001$ ). Die Unbedeutendheit der Methodenfaktoren „direkt“ und „offen“ zeigt offensichtlich das Vorliegen traitspezifischer Methodeneffekte an.

<sup>1</sup> Im Vergleich dazu wird mit den Fehlermaßen für den IAT nur unter strengeren Restriktionen (Orthogonalität der Trait- und Methodenfaktoren) eine Modellkonvergenz erreicht. Der Modellfit ist entsprechend schlechter:  $\chi^2=62.34$ ,  $df=42$ ,  $p=.02$ ,  $\chi^2/df=1.48$ , RMSEA=.054, NFI=.874. Der Anteil erklärter Varianz beträgt 46.51%.

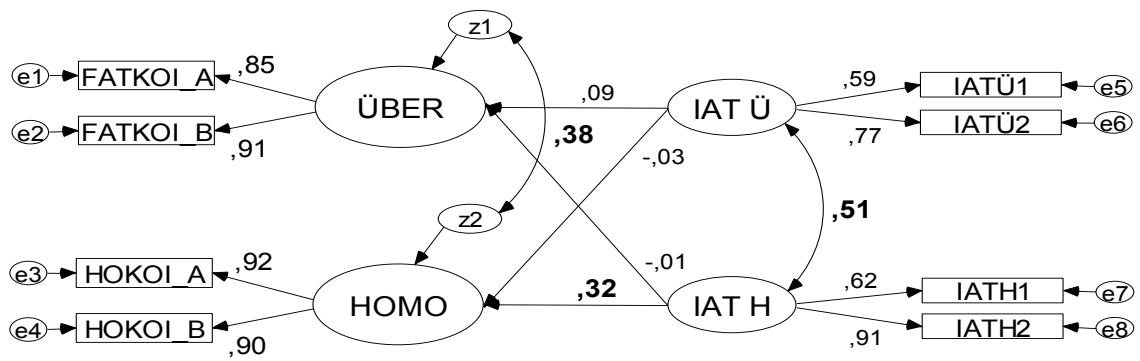
## Anhang A.6.5: Basismodelle der Vorhersage von Einstellungen (siehe Abschnitt 6.5.1)

An dieser Stelle wird die Vorhersage expliziter Einstellungen durch IAT bzw. EAST (Prädiktion 1a, 1b), die Vorhersage expliziter Einstellungen durch den Vorurteilsgehalt der Geschichten (Prädiktion 2) sowie die Vorhersage der Geschichtenmessung durch IAT bzw. EAST (Prädiktion 3a, 3b) unter statistischer Kontrolle des Messfehlers geprüft. Von zentralem Interesse sind die jeweiligen Anteile erklärter Varianz. Es wird von der relativen „Zwischenstellung“ der offenen Assoziativmessung ausgegangen, die sowohl explizite, als auch implizite Einstellungsanteile erfassen sollte. Prädiktion 1a stellt eine Replikation der Fragestellung von Gawronski (2002) mit zwei anderen Einstellungsobjekten dar. Prinzipiell werden in den Basismodellen Vorhersagepfade zwischen den Einstellungsobjekten zugelassen und auf ihre statistische Bedeutsamkeit geprüft. Diese bieten die Möglichkeit, diskriminante Validität bspw. der IAT-Indikatoren für Einstellungen zu Übergewicht bzw. Homosexualität in der Vorhersage expliziter Einstellungen zu Übergewicht bzw. Homosexualität zu untersuchen. Diskriminante Validität kann bei signifikanten Vorhersagepfaden innerhalb eines Einstellungsobjekts und nicht signifikanten Kreuzpfaden angenommen werden.

Der EAST wird trotz seiner unbefriedigenden psychometrischen Eigenschaften, die bereits die MTMM-Analyse deutlich heraushob, als Prädiktor mit berücksichtigt. Dies ist der Vollständigkeit der Datendokumentation und der Zugänglichkeit weiterer Informationen zur Evaluation des relativ unerforschten, impliziten Messverfahrens geschuldet. Aufgrund der vorangegangenen Analysen muss jedoch von einer geringen Bedeutung des EAST zur Vorhersage der relativ direkteren Einstellungsmessungen ausgegangen werden.

### 6.1.1.1 Prädiktion 1a: Vorhersage der direkten durch die indirekte Messung via IAT

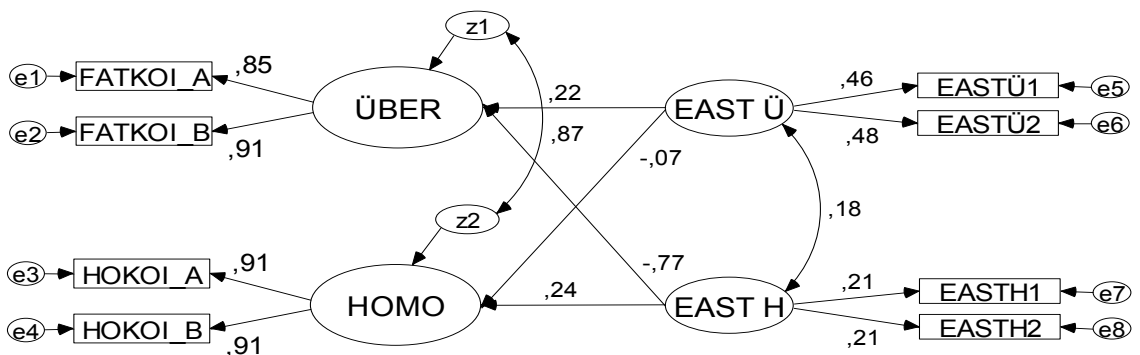
Das Modell zur Vorhersage expliziter durch implizite Einstellungen gemäß IAT gelangt zu einem akzeptablen Fit. Wie der grafischen Modelllösung zu entnehmen ist, besteht lediglich zwischen impliziten und expliziten Einstellungen zu Homosexualität eine signifikante Pfadladung. Wie auch in Gawronski's (2000) Modell bestehen keine substanziellen Kreuzladungen zwischen den Einstellungsobjekten (im Modell nicht eingezeichnet). Die Korrelationen zwischen beiden IAT Faktoren ist substanziell ( $r=.51$ ,  $p<.001$ ), ebenso wie die Korrelation der nicht vorhergesagten Anteile der direkten Einstellungsmessung ( $r=.37$ ,  $p<.001$ ). Beide Zusammenhänge repräsentieren methodenspezifische Varianz. Die Varianzaufklärung an der Fragebogenmessung ist mit 0.8% für Einstellungen zu Übergewicht (*Über*) nahezu Null und mit 9.7% für Einstellungen zu Homosexualität (*Homo*) sehr gering.



**Abbildung A.6.5-1**, Prädiktion 1a: Vorhersage direkter durch indirekte Einstellungsmessung (IAT)  
 Modell-Fit:  $\chi^2=32.30$ ,  $df=18$ ,  $p=.02$ ,  $\chi^2/df=1.80$ ,  $RMSEA=.069$ ,  $NFI=.938$

### 6.1.1.2 Prädiktion 1b: Vorhersage der direkten durch die indirekte Messung via EAST

Das Modell zur Vorhersage expliziter durch implizite Einstellungen gemäß EAST weist zwar ebenfalls einen guten Fit auf, jedoch ist kein Parameter des Strukturgleichungsmodells statistisch bedeutsam. Wiederum werden sehr geringe Pfadladungen der EAST-Faktoren auf ihre Indikatoren beobachtet. Die Varianzaufklärung von ÜBER ist mit 58.5% relativ hoch, bedingt durch den varianzstarken, jedoch nicht signifikanten negativen Vorhersagepfad ausgehend von EAST H. Für HOMO werden dagegen lediglich 5.7% der Varianz aufgeklärt. Die hohe Korrelation der Vorhersageresiduen ist nicht signifikant ( $r=.87$ ,  $p=.138$ ).



**Abbildung A.6.5-2**, Prädiktion 1b: Vorhersage direkter durch indirekte Einstellungsmessung (EAST)  
 Modell-Fit:  $\chi^2=29.96$ ,  $df=18$ ,  $p=.04$ ,  $\chi^2/df=1.67$ ,  $RMSEA=.063$ ,  $NFI=.925$

### 6.1.1.3 Prädiktion 2: Vorhersage der direkten Messung durch die Geschichtenmessung

Ein in vielerlei Hinsicht unterschiedliches Bild ergibt sich bei der Vorhersage expliziter Einstellungen durch die Einstellungsmessung gemäß den kodierten Geschichten. Das Modell gelangt nach allen Kriterien zu einem sehr guten Fit. Die Pfadgewichte der Vorhersage sind signifikant, wobei mit 35.2% wiederum eine höhere Varianzaufklärung für das Einstellungsobjekt Übergewicht erzielt wird. An der Varianz von HOMO werden dagegen 15.6% durch die Geschichtenmessung erklärt. Die Kreuzpfade der Vorhersage zwischen beide Einstellungsobjekte sind nicht signifikant.

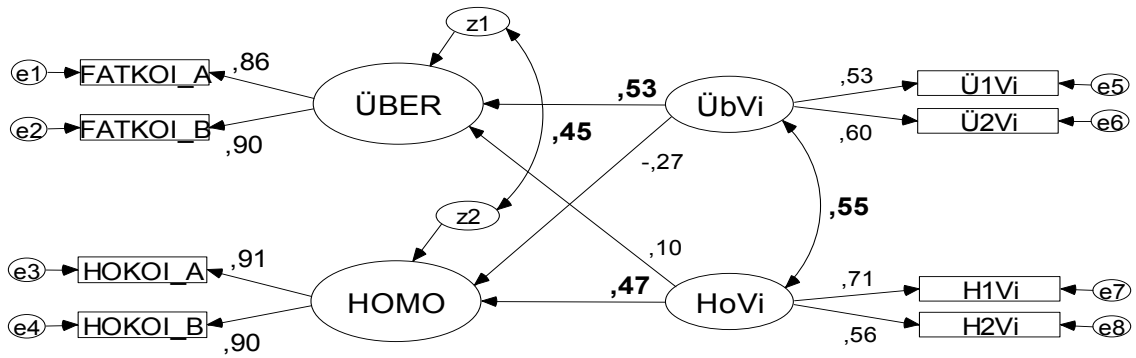


Abbildung A.6.5-3, Prädiktion 2: Vorhersage direkter Einstellungsmessung durch Geschichtenmessung  
Modell-Fit:  $\chi^2=11.74$ ,  $df=18$ ,  $p=.86$ ,  $\chi^2/df=.65$ ,  $RMSEA=.000$ ,  $NFI=.975$

#### 6.1.1.4 Prädiktion 3a: Vorhersage der Geschichtenmessung durch die indirekte Messung vis IAT

Hier stellt sich ein erwartungskonträrer signifikant negativer Zusammenhang zwischen der impliziten und der Geschichtenmessung von Einstellungen zu Homosexualität heraus, während der positive Zusammenhang für Einstellungen zu Übergewicht nicht statistisch bedeutsam ist und eine relativ hohe positive, jedoch ebenfalls nicht signifikante Ladung zwischen IATÜ und der Geschichtenmessung zu Homosexualität besteht. Die Varianzaufklärung der Geschichtenmessungen durch die IAT-Maße ist mit 6.9% für ÜbVi und 7.3% für HoVi relativ gering.

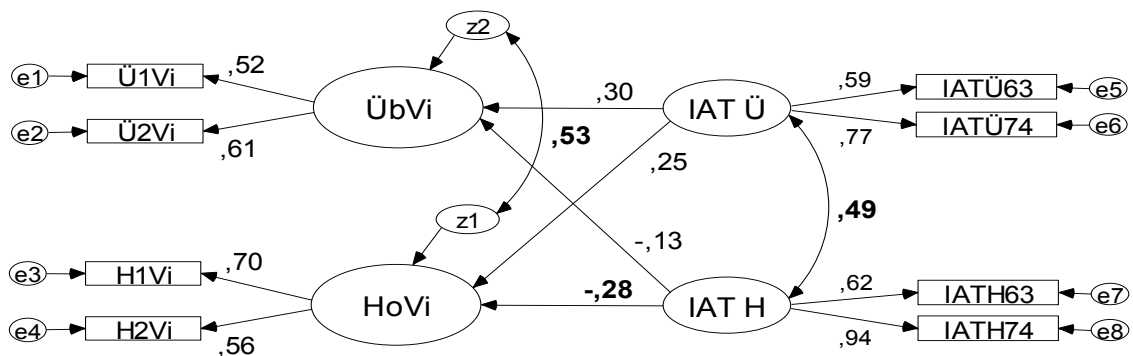


Abbildung A.6.5-4, Prädiktion 3a: Vorhersage Geschichtenmessung durch indirekte Einstellungsmessung (IAT); Modell-Fit:  $\chi^2=23.48$ ,  $df=18$ ,  $p=.17$ ,  $\chi^2/df=1.30$ ,  $RMSEA=.043$ ,  $NFI=.891$

#### 6.1.1.5 Prädiktion 3a: Vorhersage der Geschichtenmessung durch die indirekte Messung via EAST

Der EAST klärt 72.1% der Varianz der Homosexualitätsgeschichten auf, jedoch weist das Pfadgewicht von EAST H zu HoVi ein negatives Vorzeichen auf und ist zudem nicht signifikant. Von ÜbVi werden lediglich 5.2% Varianz aufgeklärt, vergleichbar zur Vorhersage der Geschichtenmessung durch die IAT-Maße in Prädiktion 3a. Wiederum erreicht in der Strukturgleichungsmodellierung mit den EAST-Daten kein Parameter der Vorhersage statistische Bedeutsamkeit. Die Korrelationen zwischen den EAST-Faktoren bzw. den Vorhersage-Residuen sind nicht signifikant.

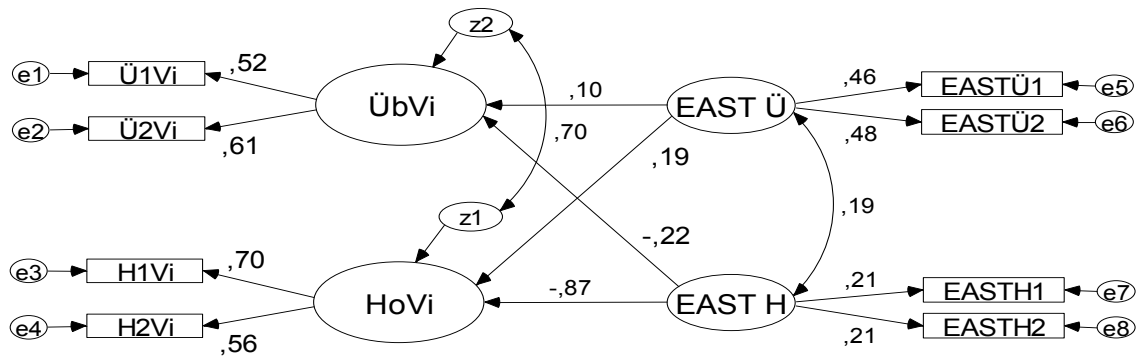


Abbildung A.6.5-5, Prädiktion 3b: Vorhersage Geschichtenmessung durch indirekte Einstellungsmessung (EAST) Modell-Fit:  $\chi^2=23.84$ ,  $df=18$ ,  $p=.16$ ,  $\chi^2/df=1.32$ , RMSEA=.044, NFI=.754

## **Ehrenwörtliche Erklärung**

Mir ist die geltende Promotionsordnung bekannt. Ich habe die Dissertation selbst angefertigt, keine Hilfe seitens eines Promotionsberaters in Anspruch genommen und alle von mir benutzten Hilfsmittel und Quellen in meiner Arbeit angegeben. Ich habe alle Personen, die mich unterstützt haben, in meiner Arbeit aufgeführt. Ich versichere, dass keine dritten Personen mittelbar oder unmittelbar geldwerte Leistungen für Arbeiten, die mit dem Inhalt der vorgelegten Dissertation in Zusammenhang stehen, von mir erhalten haben. Die Dissertation wurde für keine staatliche oder andere wissenschaftliche Prüfung eingereicht und ist ebenso wenig bei einer anderen Hochschule bzw. Fakultät als Dissertation eingereicht worden, auch nicht anteilig.

Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe

Nicole Kämpfe,

Jena den 15.04.2005

## – CURRICULUM VITAE –

Name	<i>Kämpfe</i>
Vorname	<i>Nicole</i>
Geburtsdatum	<i>16. 04. 1979</i>
Geburtsort	<i>Mühlhausen/ Thür.</i>
Staatsangehörigkeit	<i>deutsch</i>
Familienstand	<i>ledig</i>
Anschrift	<i>Schluppenstraße 3 07749 Jena</i>
Telefon dienstlich	<i>03641. 945073</i>
privat	<i>03641.511757</i>
E-mail	<i>Nicole.Kaempfe@uni-jena.de</i>
Vater	<i>Detlef Kämpfe geb.: 03.11.1950 Diplomingenieur für Maschinenbau derzeit: im Außendienst des Vertriebs von Aluminiumprofilen</i>
Mutter	<i>Dorothea Kämpfe (geborene Ernst) geb.: 01.09.1953 Facharbeiterin für Textiltechnik derzeit: Regionalgebietsleiterin im Einzelhandel</i>
Geschwister	<i>keine</i>
Schulausbildung	1985 – 1992 <i>Polytechnische Oberschule in Mühlhausen</i>
	1992 – 1997 <i>Naturwissenschaftliche Gymnasium Albert Schweitzer in Mühlhausen</i>
Schulabschluss	1997 <i>Abitur</i>
Note	<i>1.2</i>



Studium	1997	<i>Beginn des Diplom-Studienganges Psychologie an der Friedrich-Schiller-Universität in Jena (FSU)</i>
	1999	<i>Vordiplom</i>
Schwerpunktfächer im Hauptstudium		<i>Pädagogische Psychologie Arbeits-, Betriebs- und Organisationspsychologie</i>
Wahlpflichtfächer		<i>Interkulturelle Wirtschaftskommunikation</i>
Zusatzausbildungen		<i>Kommunikationspsychologie (FSU) Europakommunikation am Jean-Monet-Lehrstuhl (FSU)</i>
Praktika	08.02. – 08.04.1999	<i>Kinder- und Jugendpsychiatrie in Jena (hauptsächlich im Bereich Diagnostik)</i>
	01.11. – 01.07.2001	<i>Staatliches Schulamt Jena (Personalentwicklung und –weiterbildung über das Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien, ThILLM)</i>
Erwerbliche Nebentätigkeit während des Studiums	1998 – 2002	<i>Beschäftigung als wissenschaftliche Hilfskraft am Lehrstuhl für Psychologische Methodenlehre und Evaluationsforschung (Mitwirken im Projekt Evaluation der Lehre an der FSU, Erstellung und Evaluation von Fragebögen, Erstellung vorlesungsbegleitender Skripte, Verwalten von Literaturdatenbanken)</i>
	2000 – 2001	<i>Tutorin im Grund- und Hauptstudium für die Vorlesungen in Methodenlehre und Evaluationsforschung</i>
Diplomarbeit		<i>„Persönlichkeit, soziale Einstellungen und Fremdenfeindlichkeit: Eine empirische Untersuchung mit Selbst- und Bekannteneinschätzung“</i>
Studienabschluss	31.09.2002	<i>Diplom-Psychologin</i>
Note		<i>1 „Sehr gut“</i>
Berufliche Tätigkeit	seit 15.10. 2002	<i>wissenschaftliche Angestellte nach BAT-O IIa in einem DFG-Projekt: „Quellen interindividueller Differenzen in sozialen Einstellungen, sozialer Identität und Diskriminierung“ innerhalb der DFG-Forschergruppe „Discrimination and Tolerance in Intergroup Relations“</i>
Derzeitige Tätigkeit	seit 01.10. 2004	<i>wissenschaftliche Angestellte nach BAT-O IIa am Lehrstuhl Differentielle Psychologie, Persönlichkeitspsychologie und psychologische Diagnostik, Institut für Psychologie</i>