

OBSERVATION INFLATION: DEINE HANDLUNGEN WERDEN ZU MEINEN.

**VERWECHSLUNG VON SELBST UND ANDEREM
BEI DER ERINNERUNG AN HANDLUNGEN.**

**Von der Humanwissenschaftlichen Fakultät der Universität zu Köln
angenommene Inauguraldissertation zur Erlangung des Doktorgrades
nach der Promotionsordnung vom 12. Juli 2007**

vorgelegt von Isabel Lindner aus Bochum

März 2009

Erstgutachter: Prof. Dr. Walter Hussy; Zweitgutachter: Prof. Dr. Gerald Echterhoff

INHALTSVERZEICHNIS

0	EINLEITUNG: KÖNNEN FREMDE HANDLUNGEN ZU EIGENEN WERDEN?	8
I	OBSERVATION INFLATION: PHÄNOMEN UND ERSTE EMPIRISCHE BEFUNDE AN JUNGEN ERWACHSENEN	11
I.1	Untersuchung falscher Erinnerungen – warum und wie?	11
	I.1.1 Was nutzt die Untersuchung falscher Erinnerungen?.....	11
	I.1.2 Paradigmen zur Untersuchung falscher Erinnerungen.....	12
	I.1.3 Fehlattribution von Quellenmerkmalen vs. Fehlattribution von Vertrautheit als Mechanismen falscher Erinnerungen....	14
I.2	Falsche Erinnerungen für Ereignisse und Handlungen: imagination inflation	16
	I.2.1 Imagination inflation für Kindheitsereignisse.....	16
	I.2.1.1 Pionierstudie.....	16
	I.2.1.2 Nachfolgende Untersuchungen zur Erhellung möglicher Mechanismen.....	17
	I.2.1.3 Résumé und Diskussion.....	20
	I.2.2 Imagination inflation für einfache Handlungen.....	21
	I.2.2.1 Pionierstudie.....	21
	I.2.2.2 Nachfolgende Untersuchungen zur Erhellung möglicher Mechanismen.....	23
	I.2.2.3 Résumé und Diskussion.....	26
	I.2.3 Exkurs: Korrekte Erinnerungen für Handlungen: enactment-Effekt.....	28
I.3	Vom Vorstellen zum Beobachten: Das Konzept der Simulation	31
	I.3.1 Vorstellen als Simulationsprozess.....	31
	I.3.2 Beobachten als Simulationsprozess.....	32
	I.3.2.1 Die Entdeckung von Spiegelneuronen beim Affen..	32
	I.3.2.2 Neuronale Korrelate des Beobachtens beim Menschen.....	33
	I.3.2.3 Mögliche Funktionen der Simulation.....	37
	I.3.2.4 Weitere empirische Befunde.....	38
	I.3.3 Hinterlässt Simulation Gedächtnisspuren?.....	39
I.4	Résumé und Ableitung der Fragestellung	41

I.5	Experiment I	49
I.5.1	Versuchsplanung.....	49
I.5.1.1	Operationalisierung und Variablenvalidität.....	49
I.5.1.2	Versuchsplananlage und Versuchsplan.....	56
I.5.1.3	Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang.....	57
I.5.1.4	Kontrolle von allgemeinen Störeffekten.....	58
I.5.1.5	Kontrolle von speziellen Stör- und von Material-effekten.....	59
I.5.1.6	Ethische Gesichtspunkte.....	62
I.5.2	Versuchsdurchführung.....	63
I.5.2.1	Stichprobe.....	63
I.5.2.2	Material.....	63
I.5.2.3	Hilfsmittel und Geräte.....	72
I.5.2.4	Ablauf.....	72
I.5.3	Ergebnisse.....	78
I.5.3.1	Deskriptive Statistiken.....	78
I.5.3.2	Manipulation-Checks.....	81
I.5.3.3	Hypothesenprüfung.....	85
I.5.3.4	Weitere Befunde.....	87
I.5.4	Diskussion.....	91
I.5.4.1	Diskussion unter methodischen Gesichtspunkten...	91
I.5.4.2	Diskussion unter theoretischen und praktischen Gesichtspunkten.....	95
I.5.4.3	Résumé und Ausblick.....	100
I.6	Experiment II	103
I.6.1	Voruntersuchung.....	103
I.6.1.1	Zielsetzung der Voruntersuchung.....	103
I.6.1.2	Veränderungen des Designs gegenüber Experiment I.....	104
I.6.1.3	Stichprobe.....	106
I.6.1.4	Ergebnisse.....	107
I.6.1.5	Résumé und Ausblick.....	110
I.6.2	Versuchsplanung.....	110
I.6.2.1	Operationalisierung und Variablenvalidität.....	110
I.6.2.2	Versuchsplananlage und Versuchsplan.....	114
I.6.2.3	Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang.....	116

	I.6.2.4	Kontrolle von allgemeinen Störeffekten.....	116
	I.6.2.5	Kontrolle von speziellen Stör- und von Material- effekten.....	116
	I.6.2.6	Ethische Gesichtspunkte.....	116
I.6.3		Versuchsdurchführung.....	117
	I.6.3.1	Stichprobe.....	117
	I.6.3.2	Material.....	117
	I.6.3.3	Hilfsmittel und Geräte.....	118
	I.6.3.4	Ablauf.....	118
I.6.4		Ergebnisse.....	123
	I.6.4.1	Deskriptive Statistiken.....	123
	I.6.4.2	Manipulation-Checks.....	126
	I.6.4.3	Hypothesenprüfung.....	126
	I.6.4.4	Weitere Befunde.....	127
I.6.5		Diskussion.....	132
	I.6.5.1	Diskussion unter methodischen Gesichtspunkten...	132
	I.6.5.2	Diskussion unter theoretischen und praktischen Gesichtspunkten.....	136
	I.6.5.3	Résumé und Ausblick.....	141
I.7		Experiment III	143
	I.7.1	Versuchsplanung.....	143
	I.7.1.1	Operationalisierung und Variablenvalidität.....	143
	I.7.1.2	Versuchsplananlage und Versuchsplan.....	149
	I.7.1.3	Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang.....	150
	I.7.1.4	Kontrolle von allgemeinen Störeffekten.....	150
	I.7.1.5	Kontrolle von speziellen Stör- und von Material- effekten.....	150
	I.7.1.6	Ethische Gesichtspunkte.....	151
I.7.2		Versuchsdurchführung.....	151
	I.7.2.1	Stichprobe.....	151
	I.7.2.2	Material.....	151
	I.7.2.3	Hilfsmittel und Geräte.....	152
	I.7.2.4	Ablauf.....	152
I.7.3		Ergebnisse.....	154
	I.7.3.1	Deskriptive Statistiken.....	154
	I.7.3.2	Manipulation-Checks.....	155

	I.7.3.3	Hypothesenprüfung.....	156
	I.7.3.4	Weitere Befunde.....	157
	I.7.4	Diskussion.....	162
	I.7.4.1	Diskussion unter methodischen Gesichtspunkten...	162
	I.7.4.2	Diskussion unter theoretischen und praktischen Gesichtspunkten.....	165
	I.7.4.3	Résumé und Ausblick.....	171
I.8		Zwischenrésumé und Ausblick zu den Experimenten I-III.....	173
II		ERWEITERUNG DER EMPIRISCHEN BASIS: OBSERVATION INFLATION BEI ÄLTEREN ERWACHSENEN.....	176
II.1		Veränderungen von Gedächtnisleistungen im Alter.....	176
	II.1.1	Übersicht über Veränderungen und deren vermutete Ursa- chen.....	176
	II.1.2	Veränderungen des Quellengedächtnisses und deren neu- ronale Korrelate.....	178
	II.1.3	Imagination inflation bei älteren TeilnehmerInnen.....	180
II.2		Befunde zu motorischer Simulation bei älteren Teilnehmer- Innen.....	185
II.3		Résumé und Ableitung der Fragestellung.....	186
II.4		Experiment IV.....	191
	II.4.1	Versuchsplanung.....	191
	II.4.1.1	Operationalisierung und Variablenvalidität.....	191
	II.4.1.2	Versuchsplananlage und Versuchsplan.....	192
	II.4.1.3	Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang.....	193
	II.4.1.4	Kontrolle von allgemeinen und speziellen Stör- sowie von Materialeffekten.....	194
	II.4.1.5	Ethische Gesichtspunkte.....	194
	II.4.2	Versuchsdurchführung.....	194
	II.4.2.1	Stichprobe.....	194
	II.4.2.2	Material, Hilfsmittel und Geräte.....	195
	II.4.2.3	Ablauf.....	198
	II.4.3	Ergebnisse.....	200
	II.4.3.1	Deskriptive Statistiken.....	200
	II.4.3.2	Manipulation-Checks.....	202
	II.4.3.3	Hypothesenprüfung.....	202

	II.4.3.4	Weitere Befunde.....	203
	II.4.4	Diskussion.....	210
	II.4.4.1	Diskussion unter methodischen Gesichtspunkten...	210
	II.4.4.2	Diskussion unter theoretischen und praktischen Gesichtspunkten.....	213
	II.4.4.3	Résumé und Ausblick.....	219
III		SYNOPSIS DER BEFUNDE UND AUSBLICK: (WIE) KÖNNEN FREMDE HANDLUNGEN ZU EIGENEN WERDEN?	221
IV		KURZZUSAMMENFASSUNG	226
V		LITERATURVERZEICHNIS	227
VI		ABBILDUNGS- UND TABELLENVERZEICHNIS	242
VII		ANHANG	245
	A	Materialien zum ersten Experiment	245
	A.1	Vertrag zwischen Untersuchungsleiterin und Untersu- chungsteilnehmerInnen.....	245
	A.2	Umgestellte Sätze.....	246
	A.3	Instruktionen.....	251
	B	Materialien zum zweiten Experiment	255
	B.1	Einverständniserklärung.....	255
	B.2	Aufklärungsformular.....	257
	B.3	Handlungsanweisungen.....	258
	B.4	Instruktionen (Voruntersuchung).....	263
	B.5	Instruktionen (Hauptuntersuchung).....	265
	C	Materialien zum dritten Experiment	269
	C.1	Handlungsanweisungen.....	269
	C.2	Instruktionen.....	271
	D	Materialien zum vierten Experiment	274
	D.1	Einverständniserklärung.....	274
	D.2	Aufklärungsformular.....	276
	D.3	Instruktionen.....	276

0 EINLEITUNG: KÖNNEN FREMDE HANDLUNGEN ZU EIGENEN WERDEN?



Copyright © 1965 bei Saul Steinberg

Abb. 1: Zeichnung von Saul Steinberg (abgeb. nach Aronson, 1994, S. 348).

Korda, Herausgeber der Autobiographie von Ex-Präsident Ronald Reagan, berichtete 1997:

We had to convince Reagan not to include the story about how he recorded the atrocities at the German death camps ... a story that he had told Yitzhak Shamir, bringing tears to Shamir's eyes, because as it happens, Reagan had spent the entire war in Hollywood ... He had seen some of the first footage taken by Army cinematographers of the ... camps and had somehow convinced himself that he'd been there (zit. n. Senkfor, Van Petten & Kutas, 2002, S. 402).

Offensichtlich unterlag der ehemalige Präsident einem Phänomen, das in der Gedächtnispsychologie treffenderweise als „falsche Erinnerung“ bezeichnet wird (z.B. Roediger, 1996). Genauer könnte man von einem Fehler der Quellenzuordnung sprechen: Ronald Reagan schrieb die lebhaften Eindrücke von den Lagern nicht der Quelle „Beobachtung anderer Personen“, sondern der Quelle „eigenes Erleben“ zu (z.B. Johnson, Hashtroudi & Lindsay, 1993).

Die geschilderte Gedächtnistäuschung erscheint aus zwei Gründen besonders frappierend: Erstens besteht für Ereignisse oder Handlungen ein Erinnerungsvorteil gegenüber anderen Inhalten, z.B. verbal enkodiertem Material (*enactment-Effekt*, z.B. Cohen, 1981; Engelkamp & Krumnacker, 1980). Zweitens bereitet es Menschen normalerweise keinerlei Probleme, Selbst

und Andere(n) als Akteure sowohl aktuellen als auch vergangenen Geschehens zu unterscheiden (z.B. Jeannerod & Pacherie, 2004; Decety, & Chaminade, 2003).

Falsche Erinnerungen stellen aber nicht nur erstaunliche, sondern auch wissenschaftlich interessante Phänomene dar, denn sie können Hinweise auf die Funktionsweise des Gedächtnisses liefern (z.B. Roediger & McDermott, 2000). Darüber hinaus ist die Untersuchung von Gedächtnistäuschungen auch von praktischer Relevanz. Falsche Erinnerungen daran, ein Ereignis erlebt oder eine Handlung ausgeführt zu haben, können gravierende Konsequenzen für das alltägliche Leben nach sich ziehen, aber auch ernste Folgen außerhalb des Alltages entfalten – z.B. wenn sie im Kontext von Zeugenaussagen auftreten (z.B. Buchner & Brandt, 2002).

Innerhalb von kognitionspsychologischen Studien wurden derartige Gedächtnistäuschungen durch das *Vorstellen* von Ereignissen bzw. Handlungen hervorgerufen; ein Phänomen, das als *imagination inflation* bekannt geworden ist (Garry, Manning, Loftus & Sherman, 1996; Goff & Roediger, 1998). Der imagination-inflation-Effekt wurde sowohl für komplexere Ereignisse als auch einfachere Handlungen untersucht. Letzterer ist bisher vornehmlich auf die perzeptuelle Ähnlichkeit der beiden Quellen „Ausführung“ und „Vorstellung“ zurückgeführt worden (z.B. Thomas, Bulevich & Loftus, 2003). Denn sowohl während der Ausführung als auch während der Vorstellung einfacher Handlungen werden sensorische (z.B. taktile oder visuelle) Eindrücke enkodiert, deren Ursprung (wahrgenommen oder vorgestellt) später nicht immer eindeutig zuzuordnen ist. Neurowissenschaftliche Befunde wurden zur Erklärung des Phänomens nicht herangezogen. Dabei weisen gerade diese auf eine weitere Ähnlichkeit von Ausführung und Vorstellung hin, die a posteriori zu Schwierigkeiten bei der Diskrimination führen könnte: die motorischen Repräsentationen beider Prozesse. In zahlreichen Studien wurde nachgewiesen, dass das Vorstellen einer Handlung nicht nur als eine Form der perzeptuellen, sondern auch der motorischen Simulation dieser Handlung begriffen werden kann, so dass sich die Aktivität motorischer Hirnareale während der tatsächlichen und der bloß vorgestellten bzw. simulierten Handlungsausführung ähnelt (siehe insb. die Metaanalyse von Grèzes & Decety, 2001).

Neurowissenschaftliche Studien zeigen zudem, dass Menschen, die die Handlungen anderer *beobachten*, diese automatisch simulieren, d.h. eine korrespondierende motorische Repräsentation entwickeln (z.B. Rizzolatti & Craighero, 2004). Demnach weist auch die Beobachtung ähnliche Überlappungen mit der Ausführung auf wie die Vorstellung (z.B. Grèzes & Decety, 2001). Dass darüber hinaus motorische Hirnareale beim Abruf reaktiviert werden und daher als Basis von Erinnerungsurteilen dienen könnten, wurde für alle drei Prozesse – Ausführung, Vorstellung und Beobachtung – nachgewiesen (z.B. Senkfor, Van Petten & Kutas, 2002; Nils-

son, Nyberg, Klingberg, Åberg, Persson & Roland, 2000). Die drei Verarbeitungsformen ähneln sich aber nicht nur auf dieser Ebene. Beim Beobachten einer anderen Person werden – ähnlich wie beim Ausführen und Vorstellen – sensorische Details, perzeptuelle Eindrücke, enkodiert. So können bspw. Geräusche der eigenen, der fremden sowie der vorgestellten Handlungsausführung entstammen.

Da also die drei Prozesse des Ausführens, Vorstellens und Beobachtens einer Handlung bei allen Unterschieden (z.B. hinsichtlich des aktiven bzw. passiven Status' der Person oder hinsichtlich des Spektrums und der Qualität enkodierter sensorischer Informationen) starke neurokognitive Überlappungen aufweisen, geht die vorliegende Arbeit der Frage nach, ob nicht – ebenso wie das Vorstellen – auch das Beobachten von Handlungen falsche Erinnerungen an die Ausführung dieser Handlungen hervorrufen könnte, ob also analog zum imagination-inflation- ein *observation-inflation-Effekt* existiert; so wie in dem eingangs zitierten Beispiel. Anders formuliert: Können fremde Handlungen zu eigenen werden? Und wie geschieht dies gegebenenfalls?

Der erste Teil der Arbeit widmet sich der Ableitung dieser Fragestellung aus dem theoretisch-empirischen Hintergrund, der Entwicklung einer angemessenen Untersuchungsmethodik sowie der Darstellung von Daten, die zur Beantwortung der Fragen an jungen Erwachsenen erhoben wurden. Im Rahmen von drei Studien werden Phänomen und Mechanismus des observation-inflation-Effekts eingegrenzt und theoretische sowie praktische Konsequenzen diskutiert. Im zweiten Teil werden diese Befunde durch Untersuchungen an älteren Erwachsenen ergänzt. Hierdurch sollen weitere Hinweise auf den zugrundeliegenden Mechanismus generiert sowie Möglichkeiten, die Kenntnisse für den Alltag nutzbar zu machen, abgeleitet werden. Im letzten Teil schließlich werden die gewonnenen Daten zusammengeführt und kritisch hinsichtlich ihrer Bedeutung für Theorie und Praxis beleuchtet.

I OBSERVATION INFLATION: PHÄNOMEN UND ERSTE EMPIRISCHE BEFUNDE AN JUNGEN ERWACHSENEN

I.1 Untersuchung falscher Erinnerungen – warum und wie?

I.1.1 Was nutzt die Untersuchung falscher Erinnerungen?

Von einer Gedächtnistäuschung bzw. einer falschen Erinnerung spricht man, wenn ein erinnertes Ereignis nicht mit dem Originalereignis übereinstimmt (z.B. Buchner & Brandt, 2002). Von solchen Illusionen geht seit jeher eine Faszination aus – zeigen sie doch, dass unser subjektives Erleben nicht (immer) mit objektiv Vorhandenem übereinstimmt: „Illusions mock our belief that what we perceive, remember, and know is in perfect accord with the state of the external world“ (Roediger, 1996, S. 76). Während man sich in der Gedächtnispsychologie zunächst auf die Untersuchung korrekter Erinnerungsleistungen konzentrierte, begannen WahrnehmungspsychologInnen schon sehr früh (ab 1850), sich intensiv mit optischen Täuschungen zu beschäftigen. Dabei hat das Studium derartiger Illusionen einen großen Beitrag zum Verständnis der Funktionsweise unserer Wahrnehmung geleistet. So zeigt bspw. die Ponzotäuschung, wie die durch den Kontext induzierte Tiefe bei der Größenschätzung berücksichtigt wird und illustriert daher das Phänomen der Größenkonstanz auf sehr eindrückliche Weise (z.B. Müsseler, 2002; Wohlschläger & Prinz, 2006).

Die Erkenntnis, dass gerade das Studium falscher bzw. fehlgeleiteter Kognitionen geeignet ist, die Funktionsweise eben dieser zu erhellen (z.B. Helmholtz, 1881; Gregory & Gombrich, 1973), setzte sich jedoch in den späten 1960er/70er Jahren auch innerhalb der Gedächtnispsychologie durch; Ausgangspunkt waren falsche Erinnerungen und Suggestibilität bei Kindern. Heutzutage stellt die Erforschung von Erinnerungstäuschungen ein zentrales Teilgebiet der Gedächtnispsychologie dar und hat uns auf faszinierende Weise gezeigt, dass es sich beim Erinnern um einen (re)konstruktiven Prozess handelt (z.B. Roediger, 1996; Roediger & McDermott, 2000). Dabei greift dieser nicht mehr wegzudenkende Forschungsstrang insbesondere auf drei Paradigmen zurück, die im Folgenden näher erläutert werden sollen.

I.1.2 Paradigmen zur Untersuchung falscher Erinnerungen

Als einer der ersten beschrieb Bartlett 1932 in seinem Buch *Remembering* Erinnern als einen (re)konstruktiven Prozess, der schemageleitet ist. Er hatte StudentInnen die Indianersage „War of the Ghosts“ vorgelesen und sie anschließend zu mehreren Zeitpunkten gebeten, diese wiederzugeben. Es zeigte sich, dass die Studierenden Details der Geschichte veränderten bzw. hinzufügten, um diese kohärenter zu gestalten – kohärenter mit Schemata, die ihnen vertraut waren. So wurde bspw. die Robbenjagd durch das Angeln von Fischen ersetzt (z.B. Roediger, 1996; Buchner & Brandt, 2002). In diesem Kontext ist das erste der drei grundlegenden Paradigmen anzusiedeln: das *Deese-Roediger-McDermott Paradigma* (DRM-Paradigma). Dieses stellt folgendes Phänomen in den Mittelpunkt: Wenn Personen eine Liste stark verwandter Informationen, z.B. semantisch assoziierter Wörter wie Faden, Ohr, spitz, Garn usw., dargeboten bekommen, werden sie bei einem anschließenden Gedächtnistest dazu tendieren, nicht dargebotene, aber ebenso verwandte Informationen zu erinnern, z.B. zu behaupten, das Wort Nadel wäre auf der Liste gewesen. Deese wies diesen Effekt erstmals 1959 anhand der Methode des freien Abrufs nach. Die Prozedur wurde erst viele Jahre später von Roediger und McDermott (1995) wieder aufgegriffen und weiterentwickelt und ist seither unter der genannten Bezeichnung bekannt. In einem Großteil der nach 1995 durchgeführten Studien wurden die Fehlerinnerungen anhand von Rekognitionsurteilen nachgewiesen – d.h. unter Vorlage von Items, die daraufhin beurteilt werden sollten, ob sie in der ursprünglichen Liste waren oder nicht. Auch Studien, die falsche Erinnerungen aufgrund von logischen (Bransford & Franks, 1971) oder pragmatischen Implikationen (Brewer, 1977) nachgewiesen haben, können diesem Forschungsstrang zugerechnet werden. Unzählige Studien haben sich seither mit der Frage beschäftigt, welcher Mechanismus dieser Form von falschen Erinnerungen zugrunde liegt, und ob bzw. wie diese Fehlerinnerungen minimiert werden könnten (z.B. McDermott, 1996; Neuschatz, Benoit & Payne, 2003; Watson, McDermott & Balota, 2004). Auch wurde das Paradigma bei der Untersuchung verschiedener PatientInnengruppen eingesetzt, z.B. um zur Erhellung der spezifischen Beeinträchtigungen amnestischer Menschen beizutragen (z.B. Schacter, Verfaellie & Pradere, 1996; Budson, Michalska, Sullivan, Rentz, Daffner & Schacter, 2003; Budson, Sullivan, Mayer, Daffner, Black & Schacter, 2002). Da der vorwiegend im Bereich des semantischen Gedächtnisses angesiedelte Forschungsstrang für die vorliegende Arbeit jedoch nicht weiter relevant ist, soll an dieser Stelle nicht näher auf ihn eingegangen werden.

In den 1970er Jahren zeigten Loftus und Palmer (1974) TeilnehmerInnen eine Videodokumentation eines Autounfalls und ließen sie anschließend die Geschwindigkeit der Autos einschätzen. Die Raffinesse war, dass sie bei der Frage nach der Geschwindigkeit das Verb variierten: Die Frage lautete z.B. „About how fast were the cars going when they *smashed* each other?“

oder „About how fast were the cars going when they *contacted* each other?“ (daneben gab es drei weitere Formulierungen). In Abhängigkeit vom verwendeten Verb unterschied sich die durchschnittliche Geschwindigkeitseinschätzung deutlich: Bei Benutzung von *smash* lag sie bei 40.8 km/h, während sie bei Benutzung von *contacted* mit 31.8 km/h beziffert wurde. Neben dem Einfluss von Suggestivfragen wurde später auch die Bedeutung von falschen, nachträglich gegebenen Informationen auf die Erinnerungsleistung intensiv untersucht. Das Vorgehen ist unter dem Begriff *Falschinformationsparadigma* bekannt geworden. Klassischerweise enkodieren TeilnehmerInnen eine Originalinformation (verfolgen z.B. per Video einen Autounfall) und werden anschließend nachträglichen Informationen über dieses Ereignis ausgesetzt, die teilweise falsch sind (lesen z.B. eine Beschreibung des Vorfalls, die u.a. falsche Informationen, wie ein nicht vorhandenes Vorfahrtsschild, enthält; z.B. Loftus, Miller & Burns, 1978). Danach werden sie gebeten, zu entscheiden, ob präsentierte Items aus dem Originalereignis stammen oder nicht. Werden die falschen Informationen dem Originalereignis zugeordnet, spricht man vom sog. *Falschinformationseffekt* (*misinformation effect*; Loftus, 1975). Neben Rekognitions können auch Quellengedächtnisurteile erhoben werden, die nicht nur die Zuordnung zu einer Quelle (dem Originalereignis), sondern zu allen möglichen Quellen und deren Kombinationen erfordern (Originalereignis, nachträgliche Information, Originalereignis und nachträgliche Information, weder Originalereignis noch nachträgliche Information; z.B. Lindsay & Johnson, 1989). Diese Forschungsrichtung ist im Überschneidungsbereich von episodischem und semantischem Gedächtnis anzusiedeln und hat – ebenso wie das DRM-Paradigma – eine enorme Fülle an Untersuchungen nach sich gezogen. Das starke Interesse rührt dabei sicherlich auch daher, dass die beschriebenen Erinnerungsfehler von großer praktischer Relevanz sein können, wobei hier insbesondere an rechtspsychologische Belange zu denken ist. Auch auf diesen Forschungsstrang soll an dieser Stelle jedoch nicht näher eingegangen werden (eine Übersicht geben z.B. Ayers & Reder, 1998).

Neben diesen beiden existiert das für die vorliegende Arbeit zentrale *imagination-inflation-Paradigma* zur Untersuchung von falschen Erinnerungen an Erlebnisse bzw. Handlungen, die durch bloßes Vorstellen hervorgerufen werden (Garry, Manning, Loftus & Sherman, 1996; Goff & Roediger, 1998). Aufgrund seiner besonderen Relevanz für die vorliegende Arbeit ist dem Paradigma ein eigener Abschnitt gewidmet. Bevor jedoch näher hierauf eingegangen werden soll, sollen zwei Mechanismen aufgezeigt werden, die zur Erklärung verschiedener Formen von Erinnerungsfehlern (ebenso wie zur Erklärung korrekter Erinnerungsleistungen) herangezogen werden.

I.1.3 Fehlattribution von Quellenmerkmalen vs. Fehlattribution von Vertrautheit als Mechanismen falscher Erinnerungen

Die geschilderten Gedächtnistäuschungen werfen die Frage auf, wie eigentlich unser Gedächtnissystem seine Funktionalität bewahrt, d.h. eben doch (auch) realitätsadäquate Erinnerungsleistungen hervorbringt. Das *source memory framework* (Johnson, Hashtroudi & Lindsay, 1993) bezieht sich genau auf diese Frage: Warum sind Menschen unter bestimmten Bedingungen in der Lage, den Ursprung bzw. die Quelle von Informationen korrekt zuzuschreiben (z.B. zu entscheiden, ob ein bestimmtes Ereignis geträumt oder erlebt wurde), während dies unter anderen Bedingungen kaum möglich ist (z.B. bzgl. der Frage, ob ein semantisch verwandtes Wort bereits präsentiert wurde oder neu ist)?

Im Rahmen des *source memory framework* wird angenommen, dass die Aktivierung von Gedächtnisrepräsentationen eine ganze Bandbreite von Erinnerungsspuren, angefangen von vagen und globalen Vertrautheitsgefühlen bis hin zu Repräsentationen spezifischer Merkmale, wie sensorische Details, örtliche und zeitliche Informationen, emotionale und kognitive Prozesse, hervorrufen kann. *Vertrautheit* bzw. *Verarbeitungsleichtigkeit* oder *kognitive Zugänglichkeit* (*familiarity, processing fluency, availability*; Jacoby & Dallas, 1981; Kahneman & Tversky, 1982) sollen im Folgenden unter den Oberbegriff *Spurstärke* gefasst werden; spezifische Merkmale werden unter dem Begriff der *Spurqualität* subsumiert. Laut *source memory framework* resultieren unterschiedliche Enkodierprozesse (z.B. Vorstellen, Erleben) in Gedächtnisspuren, die sich charakteristisch voneinander unterscheiden. Erinnerungen an vorgestellte Ereignisse beinhalten bspw. typischerweise weniger lebendige sensorische, zeitliche und örtliche Informationen, jedoch mehr intentionale kognitive Operationen als erlebte Ereignisse (vgl. Mitchell & Johnson, 2000). Dies muss selbstverständlich nicht immer der Fall sein; so kann eine Vorstellung auch einmal deutlich lebhaftere Gedächtnisspuren hinterlassen als ein tatsächlich erlebtes Ereignis. Diese mentalen Repräsentationen stellen aber nur die eine Seite der Medaille dar; hinzu kommen Urteilsprozesse, die heuristischer oder systematischer Natur sein können und wiederum von verschiedenen Situations- bzw. Personmerkmalen abhängen.

Es ist offensichtlich, dass beide Aspekte – Erinnerungsspuren und Urteilsprozesse – zusammengekommen sowohl in korrekten als auch inkorrekten Quellenzuschreibungen münden können. Beide Faktoren wurden daher in einer Reihe von Studien manipuliert, um die Auswirkungen auf die Quellenattribution zu erfassen. Dabei konnten bzgl. der Qualität der Erinnerungsspuren verschiedene Maßnahmen aufgezeigt werden, die die Quellendiskrimination erschweren, wie z.B. die Störung des Enkodierprozesses durch Einführung einer zweiten Aufgabe (z.B. Jacoby, Woloshyn & Kelley, 1989) oder die Reduktion der Quellendistinktheit durch

Steigerung der perzeptuellen Ähnlichkeit von Quellen (z.B. Hashtroudi, Johnson & Chrosniak, 1990; Henkel, Franklin & Johnson, 2000). Bzgl. der Urteilsprozesse wurden ebenfalls Manipulationen untersucht, die Urteilsfehler wahrscheinlicher machten, wie z.B. die Limitierung der für das Urteil zur Verfügung stehenden Zeit (z.B. Zaragoza & Lane, 1998) oder die Induzierung laxer Urteilkriterien (z.B. Dodson & Johnson, 1993). Auch die Art des Gedächtnistests, die einen Einfluss darauf hat, welche Erinnerungsspuren überhaupt abgerufen werden, kann die Erinnerungsleistung beeinflussen. Lindsay und Johnson (1989) konnten z.B. zeigen, dass sich der Falschinformationseffekt deutlich reduzierte, wenn die TeilnehmerInnen anstatt eines Rekognitions- einen Quellengedächtnistest durchliefen, der sie nötigte, für die jeweiligen Quellen diagnostische, qualitative Erinnerungsspuren abzurufen und ihre Erinnerungsurteile nicht auf globale Kriterien, z.B. auf die Spurstärke, zu basieren (vgl. auch Zaragoza & Lane, 1994).

Dass gerade die Spurstärke Erinnerungsurteile stark beeinflussen kann, zeigten schon Jacoby und KollegInnen in den 80er Jahren. In ihren Untersuchungen zum sog. *false-fame-Effekt* fassten TeilnehmerInnen die durch vorherige Exposition erzeugte Vertrautheit von Namen fälschlicherweise als indikativ für die Bekanntheit dieser Personen auf (z.B. Jacoby, Kelley, Brown & Jasechko, 1989; Jacoby, Woloshyn & Kelley, 1989). Weitere Studien haben einen Einfluss von Spurstärke auch innerhalb anderer Paradigmen und Testformate – z.B. auch bei Quellengedächtnisurteilen – aufgezeigt (z.B. Hoffman, 1997).

Das source memory framework stellt insgesamt einen theoretischen Rahmen dar, innerhalb dessen sich Phänomene falscher Erinnerungen als Probleme der Quellendiskrimination rekonstruieren lassen, die durch Manipulationen von Gedächtnisspuren und/ oder Urteilsprozessen zu beeinflussen sind. Ob Erinnerungsurteile eher auf Spurstärke oder Spurqualität basiert werden, und ob laxe oder strikte Antwortkriterien zum Einsatz kommen, hängt also von einer ganzen Bandbreite von Faktoren (der Art des Materials, dem Alter der Erinnerung, dem Testformat usw.) ab und kann experimentell manipuliert werden (z.B. Johnson, 2006).

I.2 Falsche Erinnerungen für Ereignisse und Handlungen: imagination inflation

I.2.1 Imagination inflation für Kindheitsereignisse

I.2.1.1 Pionierstudie

Der Begriff *imagination inflation* geht auf eine Studie von Garry, Manning, Loftus und Sherman (1996) zurück (Tabelle 1 stellt den Ablauf der Untersuchung dar). Die ForscherInnen baten StudienteilnehmerInnen, einzuschätzen, wie sicher sie sich bzgl. des Auftretens bestimmter Ereignisse in ihrer Kindheit seien. Hierbei handelte es sich um durchaus plausible Ereignisse, wie z.B. auf einen Baum zu klettern und nicht mehr herunterzukommen oder auf einem Jahrmarkt ein Spielzeug zu gewinnen. Konkret legten sie ihren ProbandInnen Items des Life Events Inventory (LEI) mit der Bitte vor, diese hinsichtlich ihrer Auftretenswahrscheinlichkeit in der Kindheit auf einer achtstufigen Skala einzuschätzen. In einer nachfolgenden Sitzung, die zwei Wochen später stattfand, wurden dieselben TeilnehmerInnen dann gebeten, sich lebhaft einige der in der ersten Sitzung eingeschätzten Kindheitsereignisse vorzustellen. Dabei handelte es sich um solche Ereignisse, deren Auftretenswahrscheinlichkeit in der ersten Sitzung von den meisten TeilnehmerInnen als gering eingeschätzt worden war. Anschließend wurden dieselben Personen unter dem Vorwand, die Daten aus der ersten Sitzung seien verloren gegangen, erneut gebeten, die Wahrscheinlichkeit für das Eintreffen der LEI-Ereignisse – also auch der vorgestellten – anzugeben. Das Ergebnismuster wurde von den AutorInnen mit dem Begriff *imagination inflation* gekennzeichnet: Nach Vorstellen der Ereignisse stieg die Einschätzung ihrer Auftretenswahrscheinlichkeit an! Der Effekt konnte in mehreren nachfolgenden Untersuchungen repliziert werden. Heaps und Nash (1999) wiesen ihn bspw. anhand eines neuen Sets von LEI-Items nach und ermittelten Korrelationen zwischen dem Ausmaß an falschen Erinnerungen und verschiedenen Persönlichkeitseigenschaften (wie z.B. Suggestibilität).

Tab. 1: Ablauf der Untersuchung von Garry, Manning, Loftus und Sherman (1996).

Phase	Variable	Ausprägungen	Art der Variation
Phase 1: Rating I	AV-1: Rating der Auftretenswahrscheinlichkeit	achtstufige Skala	--
Phase 2: Verarbeitung*	UV: Art der Verarbeitung	keine weitere Verarbeitung Vorstellung	intraindividuell
Phase 3: Rating II**	AV-2: Rating der Auftretenswahrscheinlichkeit	achtstufige Skala	--

Anmerkungen: * 2 Wochen nach Phase 1; ** direkt nach Phase 2. AV/UV: abhängige/ unabhängige Variable.

Die Interpretation dieses spannenden Befundes ist nun allerdings in zweifacher Hinsicht begrenzt. Einerseits lässt sich argumentieren, dass durch das Vorstellen der Ereignisse gar nicht falsche Erinnerungen hervorgerufen, sondern korrekte Erinnerungen wieder zugänglich bzw. abrufbar gemacht worden sein könnten. Damit würde es sich also gar nicht um falsche, sondern um nur zunächst nicht zugängliche richtige Erinnerungen handeln (*reminder hypothesis*; für eine Argumentation gegen ihr Zutreffen siehe Garry et al., 1996). Zweitens liefert die Studie keinen Hinweis darauf, warum das Vorstellen (angeblich) falsche Erinnerungen hervorrufen sollte. Es stellt sich mit anderen Worten die Frage, welcher Mechanismus ursächlich für den imagination-inflation-Effekt ist.

Die erste Limitierung wurde in Folgestudien nicht nur argumentativ ausgeräumt, es wurden auch bekanntermaßen nie aufgetretene Ereignisse genutzt, um das Problem versuchsplanerisch zu umgehen. Ein Beispiel hierfür ist eine bestimmte Operation, die in Großbritannien noch niemals durchgeführt wurde (Mazzoni & Memon, 2003). In anderen Studien wurden neben den TeilnehmerInnen auch deren Angehörige, z.B. die Eltern, bzgl. des Auftretens der Ereignisse befragt (Hyman & Pentland, 1996). Da der Effekt stets repliziert werden konnte, kann insgesamt davon ausgegangen werden, dass es sich tatsächlich um ein Phänomen falscher Erinnerungen handelt. Der Frage nach dem Mechanismus ist in zahlreichen Folgestudien nachgegangen worden, die im Folgenden vorgestellt werden sollen.

I.2.1.2 Nachfolgende Untersuchungen zur Erhellung möglicher Mechanismen

Garry und Polaschek (2000) schlagen zwei Mechanismen zur Erklärung des imagination-inflation-Effekts für Kindheitsereignisse vor: Der Effekt könnte einerseits auf eine Fehlattribution von Spurstärke, d.h. auf die durch das Vorstellen gesteigerte Vertrautheit bzw. Zugänglichkeit der Items, andererseits auf eine Fehlattribution von Spurqualität und somit auf sensorische Eindrücke, die durch das Vorstellen generiert werden, zurückzuführen sein. Demnach ist für den Fall einer Fehlzuschreibung von Spurstärke folgende Heuristik denkbar: „Wenn mir das Ereignis so vertraut vorkommt, dann muss ich es wohl erlebt haben“. Demgegenüber könnte auch die Spurqualität zu falschen Schlüssen verleiten: „Wenn ich mich an Details des Ereignisses erinnere, dann muss ich es wohl erlebt haben“. Diese Heuristiken sind dabei nur illustrativ gedacht; derartige Prozesse müssen nicht bewusst ablaufen. Beide Mechanismen lassen sich Garry und Polaschek zufolge empirisch untermauern. Wurden bspw. ProbandInnen gebeten, sich selbst oder eine andere Person als AkteurIn der Ereignisse vorzustellen, führten beide Perspektiven zu demselben Ausmaß falscher Erinnerungen. Dieser Befund lässt sich – so die AutorInnen – eher mit einer Fehlattribution von Spurstärke vereinbaren (Garry, Frame & Loftus, 1999). Umgekehrt zeigte sich kein imagination-inflation-Effekt für weniger weit zurücklie-

gende Ereignisse – ein Ergebnis, das Garry und Polaschek zufolge eher mit einer Fehlattribution von Spurqualität zu erklären ist (Garry & Hayes, 1999).

Experimente zum sog. *revelation-Effekt* (z.B. Luo, 1993; Westerman & Greene, 1996) legten etwas später nahe, dass die Spurstärke der ausschlaggebende Faktor für das Zustandekommen des imagination-inflation-Effekts sein könnte. Der revelation-Effekt bezeichnet die Tendenz, ein Item in einem Rekognitionstest als „alt“ zu bezeichnen, wenn dieses oder ein vorhergehendes Wort auf eine ungewöhnliche Art und Weise präsentiert wird, z.B. in Anagrammform. Um herauszufinden, ob der revelation-Effekt ebenfalls bei der Erinnerung an Kindheitsereignisse auftritt, verfuhrn Bernstein, Whittlesea und Loftus (2002) folgendermaßen: In einer Übungsphase wurden TeilnehmerInnen Strategien an die Hand gegeben, um schwierige Anagramme – wie „went to the *umoanitsn (mountains)*“ – zu lösen. Später wurden ihnen u.a. LEI-Items vorgegeben, mit der Bitte, diese hinsichtlich ihrer Auftretenswahrscheinlichkeit in der Kindheit einzuschätzen. Dabei enthielten einige dieser LEI-Items ebenfalls Anagramme (z.B. „broke a *dwniwo (window)* playing ball“), andere nicht (z.B. „gave someone a *haircut*“). Die LEI-Anagramme waren jedoch deutlich leichter zu lösen als diejenigen aus der Übungsphase, denn einerseits konnten die eingeübten Strategien genutzt werden, andererseits waren die Items inhaltlich spezifischer formuliert. Bernstein und KollegInnen zeigten, dass TeilnehmerInnen die Auftretenswahrscheinlichkeit derjenigen LEI-Items, die in Anagrammform vorlagen, höher einschätzten als die Auftretenswahrscheinlichkeit derjenigen LEI-Items, die nicht in Anagrammform vorlagen. Offensichtlich hielten sie die (überraschende) Verarbeitungsleichtigkeit der in Anagrammform vorliegenden LEI-Items fälschlicherweise für eine Folge des tatsächlichen Auftretens bzw. Erlebens der Ereignisse (und nicht für eine Folge der besseren Lösbarkeit der Anagramme). Sie führten also die Vertrautheit, die durch die unerwartete Leichtigkeit des Lösens der LEI-Anagramme hervorgerufen wurde, darauf zurück, dass das betreffende Ereignis tatsächlich erlebt wurde; de facto jedoch waren die Anagramme aufgrund von Übung und Gestaltung leichter zu lösen (siehe auch Bernstein, Godfrey, Davison & Loftus, 2004; Whittlesea & Leboe, 2003).

Um beide Mechanismen – Fehlattribution von Spurstärke vs. Spurqualität – innerhalb des imagination-inflation-Paradigmas gegenüberstellen zu können, führten schließlich Sharman, Garry und Beuke (2004) neben dem Vorstellen eine weitere Art der Verarbeitung der LEI-Items ein. Konkret wurden TeilnehmerInnen wieder gebeten, LEI-Items hinsichtlich ihrer Auftretenswahrscheinlichkeit in der Kindheit einzuschätzen. Zwei Wochen später wurden ihnen erneut einige der Items einmal oder mehrmals vorgelegt, von denen die eine Hälfte (wie bisher) mit einer Instruktion zum Vorstellen, die andere Hälfte jedoch mit einer Instruktion zum Paraphra-

sieren präsentiert wurde¹. Um die letztere Aufgabe zu veranschaulichen, seien als Beispiel die Paraphrasierungen des Items „When I was young, I won a toy at the fair“ durch einen Teilnehmer angeführt: „At the fair I won a toy; a toy was won by me when I was young at the fair; I was young when I went to the fair and won a toy; at a fair in my youth I won a toy; I won a toy when I was young and went to the fair; when I went to a fair I won a toy“. Durch die experimentelle Manipulation der Art der Verarbeitung sollte gezeigt werden, ob das Generieren von qualitativen Merkmalen durch das Vorstellen kritisch für den Effekt ist, oder die alleinige Steigerung der Vertrautheit bzw. Verarbeitungsleichtigkeit der Items ohne die Anreicherung mit sensorischen Eindrücken ausreicht. In der Tat schien Letzteres der Fall zu sein: beide Arten der Verarbeitung führten zu einem vergleichbaren inflation-Effekt. Die AutorInnen schlossen, dass der Effekt weniger auf eine Fehlattribution qualitativer Merkmale, sondern vielmehr auf eine Fehlattribution von Vertrautheit bzw. Verarbeitungsleichtigkeit zurückzuführen sei, daher auch besser durch den Begriff der *fluency inflation* (Sharman, Garry & Beuke, 2004) erfasst würde. Die Ergebnisse einer weiteren Studie (Sharman, Manning & Garry, 2005) wiesen in dieselbe Richtung: TeilnehmerInnen waren sicherer, Kindheitsereignisse erlebt zu haben, die sie zuvor erklären sollten, als solche, die sie nicht erklären sollten.

Demgegenüber wiesen Mazzoni und Memon (2003) in einer ähnlich angelegten Studie nach, dass das bloße Lesen von Informationen über das Szenario verglichen mit dem Vorstellen des Szenarios keine falschen Erinnerungen hervorrief. Diese Ergebnisse sprechen demnach eher gegen die Spurstärke- und für die Spurqualitäts-Hypothese.

Auch die Ergebnisse von Libby (2003), die das klassische imagination-inflation-Paradigma nutzte, weisen darauf hin, dass eine Fehlattribution von Vertrautheit nicht den alleinigen Mechanismus darstellen dürfte. Sie variierte die Perspektive beim Vorstellen durch spezifische Instruktionen („1st- vs. 3rd-person imagination“) sowie die Perspektive beim Abruf, die durch den Wortlaut der Abfrage nahe gelegt wurde („remember doing vs. happened to you“). Es zeigte sich, dass der imagination-inflation-Effekt von der Übereinstimmung der Perspektive beim Vorstellen und Abrufen abhing. Offensichtlich wurden – abhängig von der durch das Testformat induzierten Perspektive – qualitativ unterschiedliche Erinnerungsspuren abgerufen und dann hinsichtlich ihrer Diagnostizität für das Erleben des Ereignisses bewertet. Stimmt Enkodier- und Abruf-Perspektive überein (z.B. „3rd-person imagination“ und „happened to you“), fiel der Effekt am höchsten aus.

¹ Diese intraindividuelle Bedingungsvariation ist allerdings auch kritisch zu betrachten, da hier Übertragungseffekte nicht ausgeschlossen werden können; siehe Abschnitt I.5.1.5.

Anknüpfend an die Studie von Libby (2003) untersuchten Sharman, Garry und Hunt (2005) den Einfluss von Quellen- und Vertrauheitscues auf den imagination-inflation-Effekt. Im Rahmen der Standardprozedur variierten sie einerseits die Perspektive beim Vorstellen (1.- vs. 3.-Person-Perspektive); hierdurch wurde der Gruppe, die zur Vorstellung aus der Ich-Perspektive angeleitet wurde, ein zusätzlicher Hinweis auf die Quelle bereitgestellt (*source cue*; bei der Abfrage wurde nämlich konstant die 3.-Person-Perspektive genutzt). Andererseits wurde der Zeitpunkt des Darreichens eines Plausibilitäts-Fragebogens variiert (direkt vor oder nach der Abfrage), der die LEI-Items enthielt. Wurde der Fragebogen vor der Abfrage dargereicht, sollte es den TeilnehmerInnen möglich sein, ihre Vertrauheitsgefühle auf diesen zu attribuieren, anstatt sie fälschlicherweise auf das Erleben der Ereignisse zurückzuführen (*familiarity cue*). Die Ergebnisse zeigten, dass nur beide cues zusammengenommen zu einer Reduzierung des Effekts führten. Demnach scheinen beide Arten von cues beim Abruf genutzt zu werden und zu der Erinnerungstäuschung beizutragen.

I.2.1.3 Résumé und Diskussion

Das von Garry et al. (1996) erstmals berichtete Phänomen falscher Erinnerungen für Kindheitsereignisse, die durch das bloße Vorstellen dieser Ereignisse hervorgerufen wurden, stellt zweifelsohne kein Artefakt dar. Die Kritik, es handele sich nicht um falsche, sondern vielmehr um richtige Erinnerungen, die durch das Vorstellen erst wieder zugänglich geworden seien, wurde argumentativ und methodisch (z.B. Mazzoni & Memmon, 2003) ausgeräumt². Es stellt sich daher die Frage, welcher Mechanismus dem Gedächtnisphänomen zugrunde liegt. Zwei mögliche Prozesse lassen sich empirisch untermauern: Eine Fehlattribution von Spurstärke und eine Fehlattribution von Spurqualität. Auch wenn sich viele Befunde eher mit dem ersten Mechanismus vereinbaren lassen (z.B. Sharman et al., 2004, 2005b), zeigen andere Studien eindeutig, dass auch die Spurqualität im Sinne von vorgestellten sensorischen Details zu der Gedächtnistäuschung beiträgt (z.B. Libby, 2003; Sharman et al., 2005a).

² Es konnte auch ausgeräumt werden, dass es sich um ein bloßes Regressionsphänomen, also ein rein methodologisches Artefakt handeln könnte (Pezdek & Eddy, 2001).

I.2.2 Imagination inflation für einfache Handlungen

I.2.2.1 Pionierstudie

Um die Frage von vornherein zu umgehen, ob es sich beim imagination-inflation-Effekt vielleicht doch um ein Phänomen wieder zugänglich gemachter, richtiger Erinnerungen handeln könnte, entwickelten Goff und Roediger (1998) ein dreiphasiges, vollständig kontrollierbares, experimentelles Labor-Paradigma, das nun eben auch die Phase der Enkodierung umfasste (Tabelle 2, folgende Seite, stellt den Ablauf der Untersuchung dar). Diese Abwandlung machte es notwendig, weit zurückliegende Kindheitserinnerungen als Gegenstand der Untersuchung aufzugeben. Stattdessen arbeiteten die ForscherInnen mit ganz einfachen Handlungen, wie z.B. eine Karte in einen Umschlag zu stecken oder die Faust zu ballen, die im Labor ausgeführt und nach einem gewissen Zeitraum erinnert werden sollten.

Konkret lief die Untersuchung so ab, dass TeilnehmerInnen innerhalb einer ersten Phase (der Enkodierphase) im Abstand von 15 Sekunden Anweisungen zu solch einfachen Handlungen vorgelesen wurden, wie z.B. „Stecken Sie die Karte in den Umschlag!“. Die Handlungsanweisungen, die zum Teil den Gebrauch von Objekten erforderten, zum Teil auch ohne Objekte auszuführen waren, sollten entweder nur angehört, vorgestellt oder tatsächlich ausgeführt werden. In der darauf folgenden Verarbeitungsphase (Phase 2) sollten sich die ProbandInnen die Ausführung einiger alter und einiger neuer Handlungen einmal oder mehrfach für jeweils 12 Sekunden vorstellen. In der letzten, zwei Wochen später stattfindenden Testphase (Phase 3), wurden den TeilnehmerInnen die Handlungsanweisungen erneut vorgelegt und sie wurden gebeten, zu entscheiden, ob diese in der ersten Untersuchungsphase präsentiert worden waren und – falls ja –, ob sie diese nur gehört, sich auch vorgestellt oder auch tatsächlich ausgeführt hätten. Goff und Roediger variierten in ihrer Untersuchung den zeitlichen Abstand von Enkodier- und Verarbeitungsphase: Phase 2 fand entweder direkt, einen Tag, eine Woche oder zwei Wochen nach Phase 1 statt; Phase 3 erfolgte stets 2 Wochen nach Phase 1 (für eine Variation des Intervalls zwischen Phase 1 und 3 siehe Mammarella, Fairfield & Cornoldi, 2007). Auch anhand dieses Paradigmas ließen sich falsche Erinnerungen im Sinne einer „imagination inflation“ nachweisen: Das Vorstellen in der Verarbeitungsphase führte zu falschen Erinnerungen an die Handlungsausführung und zwar umso mehr, je öfter die Handlungsausführung vorgestellt worden war. Für die Auswertung wurden falsche *ausgeführt*-Antworten auf solche Handlungsanweisungen, die innerhalb der ersten Phase nur gehört bzw. vorgestellt oder gar nicht präsentiert worden waren, zusammengefasst. Die Quoten falscher *ausgeführt*-Antworten lagen bspw. für die Präsentationsfrequenzen 0x, 1x, 3x und 5x in Phase 2 bei .02, .05, .08 und .10, wenn zwischen Phase 1 und 2 ein Tag lag. Dabei unterschieden sich die ersten drei

Untersuchungsbedingungen nicht hinsichtlich ihres Potentials zur Induzierung falscher Erinnerungen; fand jedoch das Vorstellen direkt vor der Testung statt (also zwei Wochen nach der Enkodierung), ließ sich kein Anstieg falscher *ausgeführt*-Antworten nachweisen.

Tab. 2: Ablauf der Untersuchung von Goff und Roediger (1998).

Phase	Variable	Ausprägungen	Art der Variation
Phase 1: Enkodierung	UV-A: Art der Enkodierung	ausgeführt vorgestellt gehört nicht enkodiert	intraindividuell
Phase 2: Verarbeitung*	UV-B: Frequenz der Verarbeitung	5x 3x 1x 0x	intraindividuell
Phase 3: Abruf**	AV-1: Rekognition	aus Phase 1 nicht aus Phase 1	--
	AV-2: Quellengedächtnis	ausgeführt vorgestellt gehört	--
	AV-3: Sicherheit	dreistufige Skala	--

Anmerkungen: * direkt, einen Tag, eine Woche oder zwei Wochen nach Phase 1; ** 2 Wochen nach Phase 1.

Durch das von Goff und Roediger entwickelte Untersuchungsdesign, das nun auch eine Enkodierphase umfasst, konnte also letztlich zweifelsfrei nachgewiesen werden, dass Vorstellungen zu falschen Erinnerungen an die Handlungsausführung führen können³. Bei der Erklärung des Phänomens wurde auf die beiden bereits im Rahmen der Studien zu Kindheitserinnerungen postulierten Mechanismen zurückgegriffen. Da durch das wiederholte Vorstellen in der zweiten Phase jedoch sowohl die Vertrautheit mit den Handlungsanweisungen gesteigert wird als auch mehr und mehr sensorische Details der Handlungsausführung generiert werden, lässt das Ergebnismuster der Studie keine Schlussfolgerungen bzgl. des zentralen Mechanismus zu. Gleichzeitig ist eine bloße Übertragung der Befunde aus den Studien zu Kindheitserinnerungen ausgeschlossen, da die von Goff und Roediger genutzten, sehr basalen und alltäglichen Handlungen, die erst kurze Zeit zurückliegen, nicht vergleichbar sind mit den komplexen und besonderen Kindheitsereignissen, die bisher untersucht wurden. Je nach Erinnerungsinhalt und -alter könnte auch die Gewichtung der cues unterschiedlich ausfallen. Nachfolgende Studien

³ Zumindest solange man nicht unterstellt, dass es sich bei dem gefundenen Effekt um das Resultat einer bestimmten Antworttendenz handelt; siehe Abschnitt I.5.4.

haben sich daher v.a. der Frage zugewandt, welcher der beiden Mechanismen zentral für die Erinnerungstäuschung im Goff-und-Roediger-Paradigma ist.

I.2.2.2 Nachfolgende Untersuchungen zur Erhellung möglicher Mechanismen

Um den Einfluss von Vertrautheit vs. qualitativen Merkmalen auf falsche Erinnerungen für die eigene Handlungsausführung im Rahmen des imagination-inflation-Paradigmas für einfache Handlungen direkt gegenüberzustellen, variierten Thomas, Bulevich und Loftus (2003) die Art der Verarbeitung in der zweiten Phase. Konkret baten sie ihre ProbandInnen in der ersten Phase, einfache Handlungsanweisungen tatsächlich auszuführen oder sich deren Ausführung nur vorzustellen; die Verarbeitungszeit betrug pro Handlungsanweisung 15 Sekunden. Im Rahmen einer zweiten Phase, die einen Tag später stattfand, wurden sie dann gebeten, sich die Ausführung einiger alter und einiger neuer Handlungsanweisungen vorzustellen; wiederum für jeweils 15 Sekunden. Dabei wurde eine Gruppe angeleitet, ihre Vorstellungen zu elaborieren, eine andere erhielt keine entsprechenden Instruktionen. Das sog. *elaborierte Vorstellen* unterschied sich dabei vom *einfachen Vorstellen* dadurch, dass neben der Handlungsanweisung zwei weitere Sätze präsentiert wurden, die sich auf den Einschluss von zwei unterschiedlichen Sinnesmodalitäten in die Vorstellung bezogen. Während also bspw. beim einfachen Vorstellen nur die Handlungsanweisung „Stellen Sie sich vor, wie Sie die Karte in den Umschlag stecken!“ präsentiert wurde, wurde diese beim elaborierten Vorstellen durch zwei zusätzliche Aufforderungen ergänzt, wie „Stellen Sie sich das Motiv der Karte vor. Stellen Sie sich vor, wie es sich anhört, wenn Sie die Karte in den Umschlag stecken.“. Diese Ergänzung sollte die TeilnehmerInnen anregen, mehr sensorische Details zu generieren und damit die perzeptuelle Ähnlichkeit tatsächlicher und vorgestellter Handlungsausführung steigern. Sollten TeilnehmerInnen unter dieser Bedingung in dem zwei Wochen später stattfindenden kombinierten Rekognitions- und Quellengedächtnistest mehr falsche Erinnerungen an die eigene Handlungsausführung aufweisen als in der Kontrollbedingung ohne sensorische Elaboration, würde dies den AutorInnen zufolge für eine Fehlattribution von sensorischen Merkmalen als primären Mechanismus der Erinnerungstäuschung sprechen. In der Tat produzierte das elaborierte Vorstellen mehr falsche *ausgeführt*-Antworten als das einfache Vorstellen (siehe Tabelle 3, folgende Seite) Um jedoch die Möglichkeit auszuräumen, dass nicht die sensorische Elaboration, sondern die gesteigerte Vertrautheit durch das wiederholte Erwähnen der Objekte in den zusätzlichen Aufforderungen verantwortlich für diesen Anstieg des Effekts sein könnte, führten die ForscherInnen eine zweite Studie durch. Diesmal wurden drei Gruppen realisiert: Eine Gruppe erhielt die Instruktion, die beiden zusätzlichen Sätze zur Elaboration der Vorstellung

Tab. 3: Ergebnisse des ersten Experiments von Thomas, Bulevich und Loftus (2003). *Ausgeführt*-Antworten und *remember-/ know*-Urteile im Quellengedächtnistest (abgeb. nach Thomas et al., 2003, S. 633).

Session 2 Activity	Simple				Elaborate			
	Familiar		Bizarre		Familiar		Bizarre	
	M	SE	M	SE	M	SE	M	SE
Action not presented in Session 1								
Imagined 0 times								
Did	.01	.01	.00	.00	.01	.01	.01	.01
Remember	.00		.00		.00		.00	
Know	.01		.00		.01		.01	
Imagined 1 time								
Did	.04	.01	.01	.01	.11	.02	.07	.02
Remember	.03		.01		.08		.05	
Know	.01		.00		.03		.02	
Imagined 5 times								
Did	.13	.03	.10	.02	.16	.02	.13	.02
Remember	.10		.06		.14		.10	
Know	.03		.04		.02		.03	
Action imagined in Session 1								
Imagined 0 times								
Did	.04	.01	.03	.03	.02	.01	.02	.01
Remember	.02		.01		.00		.00	
Know	.02		.02		.02		.02	
Imagined 1 time								
Did	.08	.01	.07	.01	.11	.02	.08	.02
Remember	.04		.03		.06		.06	
Know	.04		.04		.05		.02	
Imagined 5 times								
Did	.10	.02	.09	.02	.15	.02	.17	.02
Remember	.09		.08		.13		.10	
Know	.01		.01		.02		.07	

selber zu generieren und anschließend anzugeben, welche beiden Sinnesmodalitäten einbezogen worden waren (in Tabelle 4: „individual“). Diese Gruppe diente als Kontrollgruppe zu einer weiteren Gruppe, die (wie in Experiment 1) zum elaborierten Vorstellen angeleitet wurde. Sollten nun beide Gruppen ein vergleichbares Ausmaß an falschen Erinnerungen aufweisen, wäre der Effekt den AutorInnen zu-

folge eindeutiger der sensorischen Anreicherung zuzuschreiben als noch vor dem Hintergrund der Ergebnisse des ersten Experiments. Die wiederholte Erwähnung würde offensichtlich keine kritische Rolle spielen. Daneben wurde eine weitere Gruppe gebeten, die Handlungsanweisungen nur zu lesen, wodurch die Vertrautheit mit den Handlungsanweisungen gesteigert werden sollte, ohne diese mit sensorischen

Informationen anzureichern (in Tabelle 4: „Text“). Sollte das Lesen keine falschen Erinnerungen verursachen, würde dies ebenfalls für eine kritische Rolle sensorischer Merkmale sprechen. In der Tat sprach das Ergebnismuster Thomas et al. zufolge dafür, dass eine Fehlattribution von sensorischen Merkmalen zentral für falsche Erinnerungen im Goff-und-Roediger-Paradigma ist (siehe Tabelle 4): Die eigenständige Elaboration führte verglichen mit der vorgegebenen Elaboration zu einer vergleichbaren (bzw.

Tab. 4: Ergebnisse des zweiten Experiments von Thomas, Bulevich und Loftus (2003). *Ausgeführt*-Antworten und *remember-/ know*-Urteile im Quellengedächtnistest (abgeb. nach Thomas et al., 2003, S. 635).

Session 2 Activity	Elaborate		Individual		Text	
	M	SE	M	SE	M	SE
Action not presented in Session 1						
Presented 0 times						
Did	.01	.01	.01	.01	.01	.01
Remember	.00		.00		.01	
Know	.01		.01		.00	
Presented 1 time						
Did	.04	.01	.03	.01	.02	.01
Remember	.03		.03		.02	
Know	.01		.00		.00	
Presented 5 times						
Did	.18	.02	.13	.02	.04	.01
Remember	.15		.12		.02	
Know	.03		.01		.02	
Action imagined in Session 1						
Presented 0 times						
Did	.06	.01	.07	.01	.07	.01
Remember	.04		.04		.05	
Know	.02		.03		.02	
Presented 1 time						
Did	.13	.02	.20	.02	.07	.01
Remember	.11		.17		.05	
Know	.02		.03		.02	
Presented 5 times						
Did	.17	.02	.27	.02	.10	.01
Remember	.17		.21		.06	
Know	.00		.06		.04	

sogar leicht erhöhten) Gedächtnisverfälschung. Das bloße Lesen hingegen verzerrte die Erinnerungen fast gar nicht. Thomas et al. baten ihre TeilnehmerInnen ebenfalls, *remember/know*-Urteile hinsichtlich ihrer Erinnerungen abzugeben. Dabei sollte ein *remember*-Urteil

gefällt werden, wenn die Teilnehmerinnen sich lebhaft an Umstände der Enkodierung erinnerten, also offensichtlich qualitative Merkmale in ihre Urteile einbezogen; ein *know*-Urteil hingegen sollte indizieren, dass keine solchen lebhaften Erinnerungen vorlägen, sondern bloß intuitiv, bspw. auf einem Gefühl der Vertrautheit basierend, geurteilt würde. Das Ergebnismuster hinsichtlich der *remember*-Urteile entsprach dabei dem Ergebnismuster der falschen *ausgeführt*-Antworten (siehe Tabellen 3 und 4): Der Anteil an *remember*-Urteilen auf fälschlicherweise als ausgeführt erinnerte Handlungen stieg mit zunehmender Frequenz der Vorstellung. Außerdem wurden nach sensorischer Elaboration mehr *remember*-Urteile vergeben. Alles in allem schlossen die AutorInnen, dass eine Fehlattribution qualitativer Merkmale den zentralen Mechanismus des imagination-inflation-Effekts für einfache Handlungen darstellen würde.

Die Grundlage der soeben referierten Studie von Thomas et al. (2003) stellten sowohl Anweisungen zu ungewöhnlichen als auch gewöhnlichen (Experiment 1; in Tabelle 3: „familiar“ bzw. „bizarre“) bzw. ausschließlich zu ungewöhnlichen (Experiment 2) Handlungen dar. Dabei wurden unter gewöhnlichen Handlungen solche verstanden, die alltäglich und typischerweise mit den eingeschlossenen Objekten ausgeführt werden, z.B. „Spitzen Sie den Bleistift an!“ oder „Öffnen Sie die Sicherheitsnadel!“. Beispiele für ungewöhnliche, nicht mit den Objekten assoziierte, Handlungsanweisungen lauteten: „Balancieren Sie das Lineal auf dem Kopf!“ oder „Setzen Sie sich auf den Würfel!“. Diese letztere Art von Handlungsanweisungen ist bereits 2002 von Thomas und Loftus zur Untersuchung von falschen Erinnerungen im Rahmen des von Goff und Roediger entwickelten Paradigmas eingesetzt worden. Im Zusammenhang mit dem imagination-inflation-Effekt für Kindheitserinnerungen wurde schon früh die Frage aufgeworfen, welche Rolle die Art des Ereignisses – seine Gewöhnlichkeit, Plausibilität, Valenz usw. – spielt (z.B. Mazzoni, Loftus & Kirsch, 2001; Pezdek, Blandon-Gitlin & Gabbay, 2006; Pezdek, Blandon-Gitlin, Lam, Hart & Schooler, 2006; Sharman & Barnier, 2008). Daran anknüpfend untersuchten Thomas und Loftus die Frage nach der Gewöhnlichkeit im Hinblick auf falsche Erinnerungen für einfache Handlungen. Sie konnten zeigen, dass auch das wiederholte Vorstellen ungewöhnlicher Handlungen zu der bekannten Gedächtnistäuschung führte; der Effekt war jedoch etwas schwächer ausgeprägt als nach Vorstellen gewöhnlicher Handlungen (für eine Replikation des imagination-inflation-Effekts für gewöhnliche und ungewöhnliche Handlungen in einem Feld-Setting siehe Seamon, Philbin & Harrison, 2006). Diese Reduktion war nicht unerwartet, konnten doch diverse ForscherInnen zeigen, dass ungewöhnliche Informationen besser behalten werden als gewöhnliche (*bizarrenness-Effekt*; z.B. McDaniel & Einstein, 1986; McDaniel, Einstein, DeLosh, May & Brady, 1995). In der Tat ist es erstaunlich, dass der imagination-inflation-Effekt – trotz des Einsatzes von ungewöhnlichen Handlungsanweisungen – überhaupt noch nachzuweisen war. Thomas und Loftus zufolge weist dieser Befund ebenfalls

darauf hin, dass die falschen Erinnerungen eher durch spezifische als durch globale Erinnerungsqualitäten hervorgerufen werden.

Dieselbe Interpretation wird von Lampinen, Odegard und Bullington (2003) vertreten, die (richtige und falsche) Erinnerungen an die Handlungsausführung im Rahmen des imagination-inflation-Paradigmas hinsichtlich verschiedener Qualitäten einschätzen ließen. Sie nutzten den Memory Characteristics Questionnaire (MCQ; Johnson, Foley, Suengas & Raye, 1988) und erfragten z.B. das Ausmaß erinnerter sensorischer Details, Gefühle oder Gedanken. Zwar wurden richtig erinnerte Handlungen bzgl. aller Kriterien durchschnittlich höher eingeschätzt als falsch erinnerte – dennoch waren falschen Erinnerungen solche Qualitäten überhaupt inhärent, so dass sie den richtigen Erinnerungen qualitativ ähnelten (siehe auch Stern & Rotello, 2000)!

Mammarella (2007) setzte eine zeitlich etwas abgewandelte Prozedur ein (erste und zweite Phase fanden am selben Tag statt, die Abfrage ein oder zwei Tage später) und variierte den Abstand der in Phase 2 wiederholt präsentierten Items zueinander. Der Abstand war entweder variabel (variierte zwischen 5 und 10 Items) oder fix (betrug entweder 0 oder 4 Items vor der wiederholten Präsentation). Nur eine Bedingung rief dabei keinen imagination-inflation-Effekt hervor: Wenn der Abstand fix war und die wiederholten Items direkt aufeinander folgten. Da eine Reihe von Studien zum sog. *spacing-Effekt* (Hintzman, 1974; Whitten & Bjork, 1977) aufgezeigt haben, dass direkt aufeinander folgende Wiederholungen zu einer schlechteren Behaltensleistung führen als nicht direkt aufeinander folgende (d.h. solche, bei denen mindestens ein anderes Item vor der Wiederholung eingeschoben ist), folgert Mammarella, dass auch die Vertrautheit mit diesen direkt wiederholten Items niedriger sein müsse als die Vertrautheit mit den nicht direkt wiederholten. Daher interpretiert er das Ergebnismuster als einen Beleg für die Bedeutung von Vertrautheit für das Zustandekommen des imagination-inflation-Effekts.

I.2.2.3 Résumé und Diskussion

Erinnerungen an Ereignisse bzw. Handlungen stellen episodische Gedächtnisleistungen dar. Vergleicht man die Befunde zu falschen Erinnerungen an Kindheitsereignisse mit den Befunden zu falschen Erinnerungen an einfache Handlungen, so wurden beide auf zwei kognitionspsychologische Mechanismen zurückgeführt: Eine Fehlattribution von Spurstärke bzw. Spurqualität. Dabei wird für das Zustandekommen des imagination-inflation-Effekts für einfache Handlungen fast ausschließlich eine Fehlattribution von qualitativen Merkmalen verantwortlich gemacht. Der imagination-inflation-Effekt für Kindheitsereignisse scheint hingegen vorrangig

auf einer Fehlzuschreibung von Spurstärke zu beruhen; die Bedeutung der fälschlichen Attribution sensorischer Details wird aber zunehmend betont.

Die unterschiedliche, gar gegenläufige, Gewichtung beider Mechanismen ist dabei nicht erstaunlich, denn beide Paradigmen unterscheiden sich grundlegend. So sind lange zurückliegende, komplexe und besondere Kindheitsereignisse, die im Feld erlebt wurden, vs. kurz zurückliegende, basale und alltägliche einfache Handlungen, die im Labor ausgeübt wurden, Gegenstand der Untersuchung, und es werden Sicherheits- vs. Quellengedächtnisurteile erhoben. Allerdings lassen sich auch in der Operationalisierung der Prozesse Unterschiede aufweisen, die eventuell einen Teil der Unterschiede in den Ergebnismustern zu erklären vermögen. So wurde der Einfluss von Spurstärke bei der Beurteilung von Kindheitsereignissen dann nachgewiesen, wenn die entsprechenden Items aktiv-generierend bzw. kontrolliert verarbeitet, wurden – der Effekt fand sich also bspw. nach dem Paraphrasieren, Erklären oder dem Lösen von Anagrammen (Sharman, Garry & Beuke, 2004; Sharman, Manning & Garry, 2005; Bernstein, Whittlesea & Loftus, 2002). Im Gegensatz dazu führte eine Form des bloß automatischen und passiven Verarbeitens (Lesen; Mazzoni & Memon, 2003) zu keinem Effekt. Dies könnte darauf hinweisen, dass die *Tiefe der Verarbeitung* (Craik & Lockhart, 1972) eine kritische Rolle für das Zustandekommen falscher Erinnerungen spielt (vgl. für eine Diskussion innerhalb des imagination-inflation-Paradigmas z.B. Bernstein et al., 2004; Sharman, Manning & Garry, 2005; vgl. für eine Diskussion außerhalb des imagination-inflation-Paradigmas z.B. Drivdahl & Zaragoza, 2001; Lane & Zaragoza, 2007). Einerseits könnte man argumentieren, dass mit der Tiefe der Verarbeitung auch die Vertrautheit steigt – d.h. letztlich wiederum diesen Mechanismus als verursachend ansehen. Man könnte allerdings auch annehmen, dass der *Status der Person* (aktiv oder passiv) und damit der *Ursprung der Gedächtnisspur* (internal oder external) entscheidend ist: Sowohl beim Vorstellen als auch Paraphrasieren, Erklären usw. liegt der Ursprung der Information innerhalb der Person, wobei einerseits Vorstellungsbilder, andererseits verbale Informationen generiert werden. Vielleicht stellt eben diese Selbstreferenz die entscheidende Gemeinsamkeit der Prozesse dar, die beim Abruf als cue verwendet wird (siehe auch Stark & Perfect, 2006). Das Ergebnismuster hinsichtlich des Auftretens falscher Erinnerungen an Kindheitsereignisse lässt sich jedenfalls mit beiden Mechanismen (Vertrautheit und Informationsursprung) vereinbaren.

Wie sieht es nun bei der Erinnerung an einfache Handlungen aus? Auf diesem Gebiet wurden bisher verschiedene Formen der Vorstellung sowie das Lesen eingesetzt, um die unterschiedlichen Prozesse zu operationalisieren. Thomas, Bulevich und Loftus (2003) gingen davon aus, dass durch das elaborierte Vorstellen mit vorgegebener oder selbst zu generierender Anleitung im Vergleich zum einfachen Vorstellen allein das Ausmaß an sensorischen Informationen be-

einflusst würde. Daher sollte ein verstärktes Ausmaß an falschen Erinnerungen unter den beiden Bedingungen elaborierten Vorstellens auf eine Fehlattribution eben dieser vermehrt generierten sensorischen Informationen zurückgeführt werden können. Dabei beachten die AutorInnen nicht, dass durch die vorgegebenen und – in noch stärkerem Ausmaß – durch die selbst generierten Instruktionen zum Elaborieren auch die Verarbeitungstiefe (und damit die Vertrautheit) steigt⁴ (Craik & Lockhart, 1972; siehe auch *generation-Effekt*, Slamecka & Graf, 1978). Aus dieser Sicht kann der Anstieg des imagination-inflation-Effekts unter den beiden Bedingungen elaborierten Vorstellens nicht eindeutig auf das Mehr an sensorischen Informationen zurückgeführt werden, denn er könnte auch das Ergebnis eines Mehr an Verarbeitungstiefe und damit an Vertrautheit sein. Zudem ist zu kritisieren, dass die AutorInnen das Vorstellen dem Lesen gegenüberstellen, denn beide Prozesse unterscheiden sich nicht nur im Ausmaß sensorischer Anreicherung, sondern eben auch in der Verarbeitungstiefe sowie im Ursprung der Erinnerungsspur. Zusammengenommen erscheinen die Operationalisierungen in der Studie von Thomas, Bulevich und Loftus problematisch, und erlauben keine eindeutigen Interpretationen. Um die beteiligten Mechanismen abschließend beurteilen zu können, müssten in weiteren Studien neben dem Ausmaß an sensorischen Informationen die Verarbeitungstiefe und (damit einhergehend) der Ursprung der Erinnerungsspur (bzw. der Status der Person) variiert werden.

I.2.3 Exkurs: Korrekte Erinnerungen für Handlungen: enactment-Effekt

Insbesondere Cohen sowie Engelkamp und MitarbeiterInnen haben seit den 1980er Jahren in einer ganzen Reihe von Studien gezeigt, dass Handlungsanweisungen, die tatsächlich ausgeführt werden (SPT – subject performed tasks), deutlich besser erinnert werden als solche, die nur gelesen werden (VT – verbal tasks) (z.B. Cohen, 1981; Cohen, Sandler & Schroeder, 1987; Engelkamp & Krumnacker, 1980; Engelkamp & Zimmer, 1984; Zimmer & Engelkamp, 1985). Der Erinnerungsvorteil ausgeführter Handlungen liegt dabei zwischen 20% und 30% (Nyberg & Nilsson, 1995). Dieser so genannte *enactment-Effekt* wurde sowohl beim freien Abruf als auch bei der Rekognition nachgewiesen, gilt für Handlungsanweisungen, die sich auf Objekte beziehen, wie auch für solche, die keine Objekte einschließen und für gewöhnliche ebenso wie ungewöhnliche Handlungen (Nilsson, 2000). Er lässt die Tatsache besonders erstaunlich erscheinen, dass Menschen falschen Erinnerungen an Handlungen unterliegen können, da er aufzeigt, dass diese normalerweise besonders gut erinnert werden.

⁴ Demgegenüber spielt m.E. die Wiederholung der Information in der vorgegebenen Instruktion, die von den AutorInnen als kritisch betrachtet wird, eine geringere Rolle. Diese kann allerdings auch nicht durch die individuelle Instruktion ausgeräumt werden, denn auch diese erfordert ein Wiederholen der Informationen (wenn auch nur mental).

Laut Nilsson (2000) lassen sich drei der vier für den enactment-Effekt vorgeschlagenen Erklärungsansätze zusammenführen und dem vierten gegenüberstellen. Das erste Set von Erklärungsansätzen lässt sich dabei insbesondere auf zwei zentrale Faktoren herunterbrechen, denen eine besondere Bedeutung beigemessen wird: physikalische Bewegung und nicht-strategische Verarbeitung (z.B. Cohen, 1981; Bäckman & Nilsson, 1984; Engelkamp & Zimmer, 1984). Demgegenüber zeichnet sich der Ansatz von Kormi-Nouri dadurch aus, dass die Motorik an sich nicht als kritisch angesehen und die Enkodierung als strategisch aufgefasst wird (Kormi-Nouri, 1995; für einen Versuch, die Divergenzen zu überwinden, siehe Nilsson, 2000). Ohne die Unterschiede zwischen beiden Ansätzen näher herausarbeiten bzw. die Fülle an empirischen Daten darstellen zu wollen, soll kurz auf die Rolle von motorischen cues beim Abruf, die insb. von Engelkamp und KollegInnen hervorgehoben wurde (z.B. Engelkamp, 1997; Engelkamp & Jahn, 2003), eingegangen werden, da dies für die Fragestellung der Arbeit wichtig erscheint.

Verhaltensbasierte Studien haben u.a. gezeigt, dass der enactment-Effekt nicht verschwindet und auch nicht reduziert wird, wenn TeilnehmerInnen gebeten wurden, die Handlungen mit geschlossenen Augen auszuführen, also keine visuellen Eindrücke enkodieren konnten (Engelkamp, Zimmer & Biegelmann, 1993). Der Effekt tritt auch unabhängig davon auf, ob die

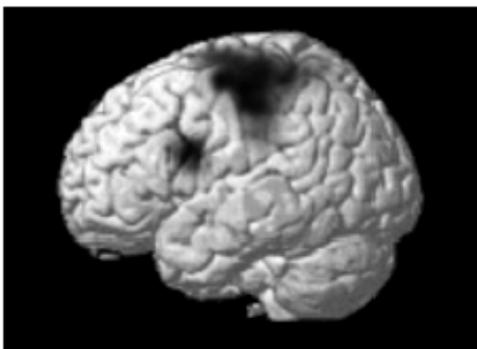


Abb. 2: Gehirnaktivität beim cued-recall-Test. Ausführung kontrastiert mit verbaler Enkodierung in der Studie von Nyberg et al. (2001). Die Aufgabe war vergleichbar mit derjenigen von Nilsson et al. (2000). Erhöhte Aktivität fand sich im kontralateralen (linken) motorischen/somatosensorischen, prämotorischen, inferioren und superioren parietalen Cortex (abgeb. nach Nyberg et al., 2002, S. 836).

Handlung mit einem echten oder nur vorgestellten Objekt ausgeführt wird (Engelkamp & Zimmer, 1983). Wiederholten TeilnehmerInnen die enkodierte Bewegung vor dem Abruf, verbesserte sich die Leistung – aber nur, wenn dieselbe Hand wie bei der Enkodierung genutzt wurde (Engelkamp, Zimmer, Mohr & Sellen, 1994). Diese und weitere Befunde können als Belege für die Wichtigkeit der Bewegungsinformation gewertet werden (siehe auch Engelkamp, 1997; Daprati, Nico, Saimpont, Franck & Sirigu, 2005; Engelkamp & Zimmer, 1994; Hornstein & Mulligan, 2001; Ross, Wang, Kramer, Simons & Crowell, 2007). In jüngerer Zeit wurde anhand von bildgebenden Verfahren zu-

sätzliche Evidenz auf neuronaler Ebene erbracht. So zeigten Nilsson et al. (2000), dass motorische Hirnareale eine erhöhte abrufbezogene Aktivität aufwiesen, wenn die Handlungen zuvor tatsächlich ausgeführt worden waren, verglichen mit nur verbaler Enkodierung (vgl. Abbildung 2). Die AutorInnen schließen, dass eine Ursache für den enactment-Effekt in der Möglichkeit liegen könnte, den Abruf auf motorische Informationen zu basieren (siehe auch Heil, Rolke,

Engelkamp, Rosler, Ozcan & Hennighausen, 1999; Masumoto, Yamaguchi, Sutani, Tsuneto, Fujita & Tonoike, 2006; Nyberg, Petersson, Nilsson, Sandblom & Ingvar, 2001; Senkfor, Van Petten & Kutas, 2002). Nilsson et al. hatten bei der Abfrage einen cued-recall-Test eingesetzt, bei dem das Verb der Handlungsanweisung als cue fungierte. Sie halten es daher auch für denkbar, dass die Aktivierung – anstelle einen Teil der episodischen Gedächtnisspur darzustellen – auf durch den cue getriggerte motorische Repräsentationen zurückzuführen sein könnte, die erst während der Abfrage entstehen. Das Ergebnismuster dieser sowie weiterer Studien (z.B. Heil et al., 1999; Nyberg et al., 2001) lässt sich allerdings insg. besser mit der ersteren Idee – motorische Gedächtnisspur als Basis des Erinnerungsurteils – vereinbaren. In einer kürzlich publizierten Studie zeigten Senkfor, Van Petten und Kutas (2008), dass die Reaktivierung motorischer Areale von der Art der Erinnerungsaufgabe abhängt. Die Forscherinnen leiteten ihre TeilnehmerInnen zu motorischer versus konzeptueller Enkodierung an. Sie baten ihre TeilnehmerInnen entweder, eine typische Handlung mit einem vorgegebenen Objekt auszuführen, oder den Preis des Objektes einzuschätzen. Während eines anschließenden Rekognitions-tests für Objekte aus der Enkodierphase wurde die Gehirnaktivität per EEG aufgezeichnet. Es ergaben sich keine Unterschiede zwischen den beiden Enkodierbedingungen. Danach wurden die TeilnehmerInnen erneut mit Objekten konfrontiert und instruiert, diese motorisch oder konzeptuell zu verarbeiten. Bei der anschließenden Abfrage setzten die Forscherinnen einen Quellengedächtnistest ein. Während dieses Tests zeigten sich nun deutliche Unterschiede in der neuronalen Aktivität zwischen den Enkodierbedingungen. Die Autorinnen schlussfolgern, dass die Art des Erinnerungstests bestimmt, welche Informationen bzw. Gedächtnisspuren abgerufen werden. Soll bspw. – wie bei Nilsson et al. – das Verb der Handlungsanweisung erinnert werden oder – wie bei Senkfor et al. – ein Objekt dahingehend beurteilt werden, ob mit ihm eine Handlung ausgeführt wurde (Quellengedächtnis), ist eine Reaktivierung motorischer Areale funktional. Wenn hingegen ein Objekt daraufhin bewertet werden soll, ob es alt oder neu ist (Rekognition), ist eine Reaktivierung motorischer Informationen hingegen nicht notwendig und erfolgt auch nicht automatisch.

I.3 Vom Vorstellen zum Beobachten: Das Konzept der Simulation

I.3.1 Vorstellen als Simulationsprozess

Beim Vorstellen handelt es sich um die bemerkenswerte menschliche Fähigkeit, mentale Repräsentationen von Objekten, Personen usw. hervorzubringen, die nicht Teil der bestehenden Situation und daher in der Situation nicht wahrnehmbar sind (Kosslyn, Thompson & Ganis, 2006; Harris, 2000; O'Connor & Aardema, 2005). Dabei können Vorstellungen Inhalte zum Gegenstand haben, die tatsächlich (außerhalb der gegebenen Situation) existieren, aber auch vollkommen neue Realitäten erschaffen. Vorstellungen sind an zahlreichen menschlichen Leistungen beteiligt (z.B. Kreativität) und wurden in verschiedenen Zusammenhängen untersucht, z.B. als eine Form der Mnemotechnik (z.B. Bower, 1970; Paivio, 1969; siehe auch Kosslyn, Behrmann & Jeannerod, 1995).

Personen, die gebeten werden, Vorstellungen zu generieren, berichten introspektiv, dass diese Vorstellungen als der Wahrnehmung sehr ähnlich, wenn auch in gewisser Weise „schwächer“ oder „blasser“ erlebt würden (Glaser, 2004). Kosslyn (z.B. 1981; siehe auch Kosslyn, Thompson & Ganis, 2006) vertritt die Meinung, dass Vorstellungen auf der Basis von Wissensbeständen im Langzeitgedächtnis aufgebaut und dann vom Wahrnehmungssystem auf die gleiche Weise verarbeitet werden wie Gegebenheiten der Außenwelt. Die Gegenposition im Rahmen der sog. *imagery-Debatte* wurde von Pylyshyn (z.B. 1981) vertreten, der nicht von einem wahrnehmungsanalogen Prozess, sondern von einem „stillschweigenden Wissen“, einer einheitlichen abstrakten (propositionalen) mentalen Repräsentation als Basis von Vorstellungen ausgeht (vgl. Tye, 1991; Städtler, 2003). Untersuchungen zur mentalen Rotation (z.B. Shepard & Metzler, 1981), aber auch aus anderen Bereichen, zeigen jedoch, dass die Durchmusterungsprozesse beim Vorstellen annähernd denselben Gesetzmäßigkeiten unterliegen wie reale Wahrnehmungen und unterstreichen so die Wahrnehmungsnahe betonende Position Kosslyns (siehe z.B. auch Borst & Kosslyn, 2008; Hostetter & Alibali, 2008; Farah, 1985; Finke, 1980).

Der Akt des Vorstellens weist jedoch nicht nur auf behavioraler, sondern auch auf neuronaler Ebene Ähnlichkeiten mit der Wahrnehmung auf. So konnten bspw. Ganis, Thompson und Kosslyn (2004) zeigen, dass die Gehirnregionen, die beim Vorstellen visueller Reize aktiviert werden, und die Gehirnregionen, die beim Wahrnehmen dieser visuellen Reize aktiviert werden, stark überlappen (siehe auch Kosslyn, 2005; Kensinger & Schacter, 2006).

Analog zu diesen Befunden haben eine ganze Reihe neurowissenschaftlicher Untersuchungen den Akt des *Vorstellens von Handlungen* nicht nur als einen Prozess der perzeptuellen, sondern auch der *motorischen Simulation* oder der „verdeckten“ Handlungsausführung konzipiert, der eben auch zu ähnlichen motorischen Repräsentationen führt, wie die tatsächliche Handlungsausführung (z.B. Jeannerod, 2001). So zeigten z.B. Meister et al. (2004), dass bei Pianisten während des Spielens eines Instrumentes und während des bloßen Vorstellens dieser Handlung ähnliche fronto-parietale Gehirnregionen aktiviert waren.

In einer Metaanalyse stellten Grèzes und Decety (2001) Studien zusammen, im Rahmen derer bildgebende Verfahren (Positronen-Emissions-Tomographie (PET) bzw. funktionelle Magnet-Resonanz-Tomographie (fMRT)) zur Erfassung des cerebralen Blutflusses während des Vorstellens verschiedener Handlungen eingesetzt worden waren (z.B. Ergreifen eines Objektes, Bewegen eines Joysticks; u.a. Decety et al., 1997; Grafton, Arbib, Fadiga, Rizzolatti, 1996; Grèzes & Decety, 2001; Stephan et al., 1995). Es fand sich ein Anstieg der cerebralen Durchblutung im dorsolateralen präfrontalen Cortex, Gyrus praecentralis, supplementär-motorischen Areal (SMA), inferioren Parietallappen, Gyrus cinguli, in den subcorticalen Nuclei und im Cerebellum. Dabei gehören die genannten corticalen und subcorticalen Gehirnregionen einem Netzwerk an, das bekanntermaßen während der tatsächlichen Handlungsausführung aktiviert ist.

I.3.2 Beobachten als Simulationsprozess

I.3.2.1 Die Entdeckung von Spiegelneuronen beim Affen

„Every representation of a movement awakens in some degree the actual movement which is its object“ (James, 1890, S. 1134). Schon in den *Principles of Psychology* legte William James, aufbauend auf den Ausführungen von Lotze (1852), die Idee an, dass motorische Repräsentationen durch *Wahrnehmungen* aktiviert werden könnten. Allerdings ging es James und Lotze v.a. um den Zusammenhang von *Gedanken* an bzw. *Vorstellungen* von Bewegung und tatsächlicher Bewegung. Es dauerte weitere 80 Jahre bis Greenwald (1970, 1972) James' ideomotorische Theorie so erweiterte, dass sie auch durch Wahrnehmung induzierte motorische Repräsentationen einschloss (siehe auch Prinz, 2002; Knuf, Aschersleben & Prinz, 2001). Die Idee eines engen Zusammenspiels beider Komponenten war lange Zeit wenig prominent, denn Wahrnehmung und Motorik galten als strikt getrennte Systeme (Rizzolatti & Sinigaglia, 2008). Erst in jüngerer Zeit wurde diese Trennung überwunden. Die Basis hierfür bildete die Entdeckung von Spiegelneuronen beim Affen und später dann auch beim Menschen.

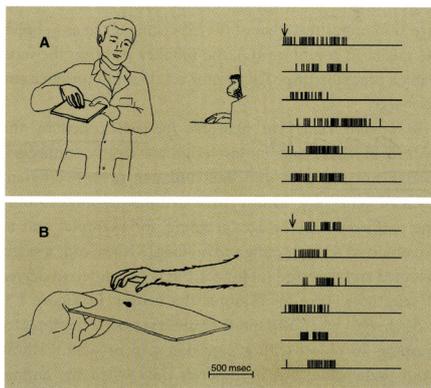


Abb. 3: Antwort eines Spiegelneurons für Greifen auf die Beobachtung des greifenden Experimentators (A) bzw. auf den selbst ausgeführten Greifakt (B) (di Pellegrino et al., 1992; abgeb. nach Rizzolatti & Sinigaglia, 2008, S. 81).

In den 1990er Jahren wurden bei der Untersuchung von Affen *Spiegelneuronen* zunächst im prämotorischen Cortex (Area F5) entdeckt, dann auch im rostralen Teil des inferioren Parietallappens (Di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992; Rizzolatti, Fadiga, Gallese & Fogassi, 1996a; Gallese, Fadiga, Fogassi & Rizzolatti, 1996). Diese Neuronen wurden sowohl bei der Ausführung einer bestimmten Bewegung als auch bei der Beobachtung dieser Bewegung aktiv, d.h. wenn der Affe selbst nach einer Nuss griff oder den Experimentator hierbei beobachtete (siehe Abbildung 3). Spiegelneuronen entladen selektiv bei der Ausführung oder Beobachtung spezi-

fischer Handbewegungen – es gibt z.B. solche, die entweder von dem Akt des Greifens, des Haltens oder des Manipulierens aktiviert werden; manche Spiegelneuronen reagieren aber auch auf zwei oder drei verschiedene Bewegungen. Diese Art von Spiegelneuronen sind im dorsalen Bereich von F5 zu finden; daneben existieren im ventralen Bereich von F5 Neuronen, die durch Mundbewegungen erregt werden. Für die Aktivierung von Spiegelneuronen beim Affen ist eine Interaktion mit einem Objekt erforderlich; ein Akteur, der eine Handlung ohne das entsprechende Objekt nachahmt, das Objekt allein sowie intransitive (nicht objektbezogene) Handlungen rufen keine spiegelneuronalen Antwort hervor (Rizzolatti & Craighero, 2004). Grundsätzlich werden – nach dem Grad der Übereinstimmung von beobachteter und kodierter Handlung – strikt und weniger strikt kongruente („strict and broadly congruent“) Neuronen unterschieden. Die ersteren stellen mit ca. 30% den kleineren Teil dar. Während für ihre Aktivierung die Übereinstimmung zwischen der ausgeführten und beobachteten Handlungen nahezu perfekt sein muss, werden die weniger strikt kongruenten Spiegelneuronen auch von verwandten bzw. ähnlichen Handlungen aktiviert und machen den Großteil (ca. 70%) der Spiegelneuronen des Affen aus. Daneben gibt es einige Neuronen mit spezifischen Reaktionsmustern (z.B. Reaktion nur auf eine bestimmte ausgeführte, aber zwei unterschiedliche beobachtete Handlungen; Rizzolatti & Sinigaglia, 2008). Der Entdeckung von Spiegelneuronen beim Affen schloss sich bald die Frage an, ob ein ähnliches System auch beim Menschen existieren könnte.

I.3.2.2 Neuronale Korrelate des Beobachtens beim Menschen

Im Lichte des Spiegelneuronensystems ließen sich einige zunächst unerklärliche Befunde im Nachhinein verstehen. So hatte Gastaut bereits in den 50er Jahren in EEG-Studien gefunden,

dass der μ -Rhythmus nicht nur desynchronisiert wird, wenn eine Bewegung oder eine somato-sensorische Reizung erfolgt, sondern auch, wenn eine Bewegung einer anderen Person beobachtet wurde (Gastaut & Bert, 1954; Cohen-Seat, Gastaut, Faure & Heuyer, 1954). Dieser Befund wurde später u.a. von der Arbeitsgruppe um Cochin repliziert (Cochin, Barthelemy, Lejeune, Roux & Martineau, 1998; Cochin, Barthelemy, Roux & Martineau, 1999). Ähnliche Ergebnisse wurden auch unter Einsatz von Magnetenzephalographie (MEG) oder motorisch evozierten Potentialen (MEP) mittels transcranieller Magnetstimulation (TMS) erzielt (z.B. Hari, Forss, Avikainen, Kirveskari, Salenius & Rizzolatti, 1998; Fadiga, Fogassi, Pavesi & Rizzolatti, 1995). Fadiga et al. (1995) konnten bspw. zeigen, dass motorisch evozierte Potentiale bei der Beobachtung zunahmen und in ihrem Muster demjenigen entsprachen, das bei der Handlungsausführung gefunden wurde. Im Unterschied zu den an Affen gewonnenen Befunden, wurden die entsprechenden Neuronen beim Menschen allerdings nicht nur durch transitive, sondern auch durch intransitive Handlungen erregt (Rizzolatti & Sinigalia, 2006; siehe auch Maeda, Kleiner-Fisman, Pascual-Leone, 2002). Wichtig war, dass es sich bei den beobachteten Akteuren um natürliche Agenten handelte – künstliche Agenten riefen keine Resonanz hervor (Rizzolatti & Sinigalia, 2008). Außerdem zeigte sich, dass auch die zeitliche Dauer der Bewegungen exakt reproduziert wurde (siehe auch Gangitano, Mottaghy & Pascual-Leone, 2001).

Bildgebende Verfahren haben dazu beigetragen, das Spiegelneuronensystem beim Menschen einzugrenzen: „The areas constantly activated during action observations are the rostral (anterior) portion of the inferior parietal lobule and the lower part of the precentral gyrus plus the posterior part of the inferior frontal gyrus. In certain experimental conditions a more anterior region of the inferior frontal gyrus and the dorsal premotor cortex were also activated.“ (Rizzolatti & Sinigalia, 2008, S. 119; siehe auch Buccino et al., 2001; vgl. Abbildung 4).

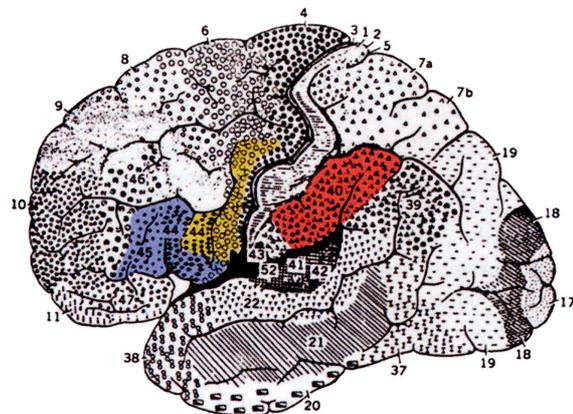


Abb. 4: Das menschliche Spiegelneuronensystem. Rot und gelb: Teile des Parietal- bzw. Frontallappens, die während der Ausführung und Beobachtung aktiviert sind. Manchmal wird auch der dorsale Anteil der Area 6 zum Spiegelneuronensystem gezählt, der allerdings möglicherweise eher zum Zwecke der Handlungsvorbereitung aktiviert ist. Blau: Teil des Frontallappens, der in einigen experimentellen Bedingungen bei der Beobachtung aktiviert war; hierbei könnte es sich möglicherweise aber auch um andere als spiegelneuronale Aktivität handeln (abgeb. nach Rizzolatti & Sinigaglia, 2008, S. 120).

Zu einem ganz ähnlichen Schluss kommen auch Grèzes und Decety (2001), die in die bereits zitierte Metaanalyse auch verschiedene Studien zum Beobachten einschlossen (Bonda, Petrides, Ostry & Evans, 1996; Decety et al., 1997; Grafton et al., 1996; Grèzes, Costes & Decety, 1998; Perani et al., 2001; Rizzolatti, Fadiga, Matelli, Bettinardi, Perani & Fazio, 1996b). Bei der Beobachtung von zielgerichteten Handbewegungen ließ sich in allen Studien ein Anstieg des regionalen cerebralen Blutflusses im prämotorischen Cortex, medialen Gyrus temporalis, inferioren und medialen Gyrus frontalis und parietalen Cortex der linken Hemisphäre beobachten. Demnach könnte man auch davon sprechen, dass eine beobachtende und eine ausführende Person die Bewegungsrepräsentationen „teilen“ (*shared representations*; Grèzes & Decety, 2001; siehe auch Gallese & Goldman, 1998).

Es besteht also ganz offensichtlich eine Überschneidung der Hirnaktivität während der tatsächlichen Ausführung von Handlungen und dem Beobachten derselben. Dies gilt auch für das Vorstellen, so dass alle drei Formen der Verarbeitung Ähnlichkeiten hinsichtlich ihrer neuronalen Repräsentation aufweisen, die v.a. motorische Areale betreffen. Konkret ergab die Metaanalyse von Grèzes und Decety (2001) eine Überschneidung aller drei Prozesse in den supplementär motorischen Arealen (SMA), im dorsalen prämotorischen Cortex, Gyrus supramarginalis und dem superioren Parietallappen (siehe Abbildung 5; siehe auch Jeannerod, 2001; Decety, 2002).

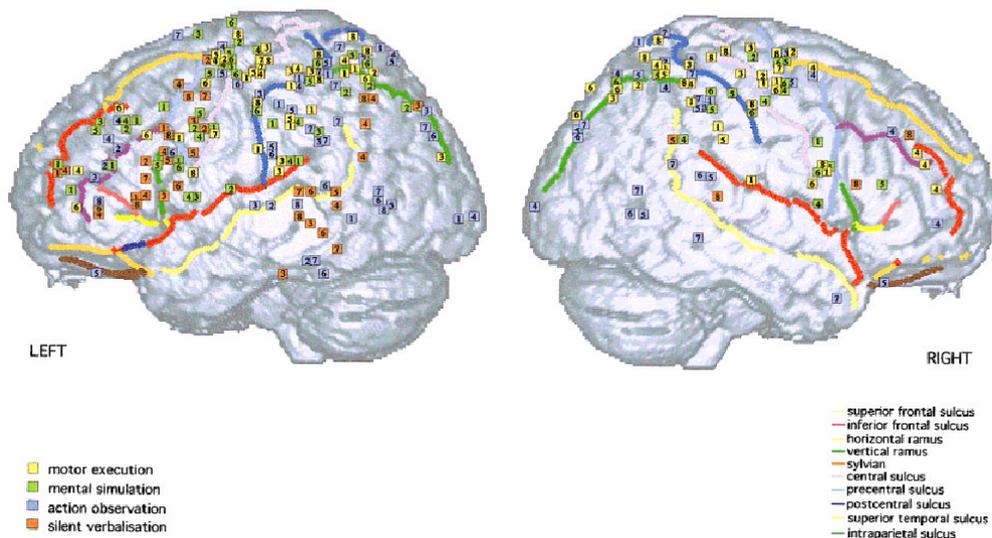


Abb. 5: Lateralansicht der Aktivationsfoki unter den verschiedenen Verarbeitungsbedingungen in der Metaanalyse von Grèzes und Decety (2001). Gelb: Ausführung. Grün: Vorstellung. Blau: Beobachtung. Rot: mentale Verbalisierung. Die Zahlen entsprechen den in die Metaanalyse einbezogenen Studien (abgeb. nach Grèzes & Decety, 2001, S. 5).

Neben Ähnlichkeiten auf Ebene der motorischen Repräsentation beinhalten alle drei Prozesse eine Anreicherung mit sensorischen Informationen. So können z.B. Geräusche der eigenen Ausführung, der Vorstellung, aber auch der Beobachtung entstammen. Allerdings unterscheiden sich Spektrum und Qualität sensorischer Eindrücke je nach Verarbeitungsart: Während es sich bei der Ausführung um eine echte Tätigkeit handelt, in die alle Sinne einbezogen werden können, kann die Vorstellung zwar ebenfalls alle Sinne umfassen, stellt aber nur eine „verdeckte“ Ausführung dar, die „blasser“ als eine tatsächliche Ausführung erlebt wird. Beim Beobachten wiederum werden v.a. visuelle und akustische Eindrücke, also ein sehr begrenztes Spektrum an Sinneseindrücken, enkodiert; diese beruhen jedoch – im Gegensatz zur Vorstellung, aber in Übereinstimmung mit der Ausführung – auf tatsächlichen Wahrnehmungen. Ebenso wie die motorischen spiegeln sich selbstverständlich auch die Ähnlichkeiten sensorischer Repräsentationen auf neuronaler Ebene wider (siehe Abschnitt I.3.1; siehe auch Grèzes & Decety, 2001; Senkfor et al., 2002)

Die beiden Formen der Simulation – Vorstellung und Beobachtung – lassen sich aus kognitionspsychologischer Perspektive jedoch auch entlang weiterer Dimensionen unterscheiden. Hierbei ist insbesondere an die Dimensionen Kontrolliertheit – Automtizität und Aktivität – Passivität, einhergehend mit internalem – externalem Informationsursprung, zu denken (siehe z.B. Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Während das Simulieren beim Vorstellen bewusst initiiert wird und kontrolliert abläuft, also Aktivität von Seiten des Individuums voraussetzt, handelt es sich bei der beobachtungsinduzierten Simulation um einen unbewusst initiierten, automatisch ablaufenden, passiven Prozess. Damit einhergehend unterscheiden sich die Ursprünge der Stimuli – sie sind entweder internal und von außen nicht zugänglich (Vorstellung) oder external und von außen zugänglich (Beobachtung). Als *reality monitoring* bezeichnen Johnson und Raye (1981) eine spezielle Form des source monitorings – nämlich eben die Unterscheidung zwischen dem externalen (auf Wahrnehmung basierenden) oder internalen (auf Vorstellung, Nachdenken o.ä. basierenden) Ursprung einer Information. Beide unterscheiden sich charakteristisch voneinander (siehe Abschnitt I.1.3). Darüber hinaus tritt Simulation durch Beobachtung qua definitionem nur auf, wenn eine soziale Situation gegeben ist, sprich eine Person vorhanden ist, die zu beobachten ist. Vorstellungen können zwar sicherlich auch in sozialen Situationen auftreten, sind aber nicht an diese gebunden. Im Gegenteil handelt es sich um einen sehr individuellen Prozess – die Vorstellung muss kommuniziert werden, um überhaupt interindividuell erfahrbar zu sein.

I.3.2.3 Mögliche Funktionen der Simulation

„Perceiving other people’s behaviors activates imitative motor plans in the perceiver, but there is disagreement as to the function of this activation.” (Wilson & Knoblich, 2005). Im Wesentlichen können fünf Erklärungsansätze unterschieden werden: Simulation durch Beobachtung könnte die Basis für Beobachtungslernen darstellen (Iacoboni et al., 1999; Brass & Heyes, 2005), dazu beitragen, die Handlungen (Rizzolatti, Fogassi & Gallese, 2001) oder sogar Intentionen/ Ziele (Blakemore & Decety, 2001; Iacoboni et al., 2005) anderer zu verstehen, eine Rolle für die Vorhersage von Handlungen spielen (Wilson & Knoblich, 2005), dem Arbeitsgedächtnis dienen (Wilson, 2001) oder für die sprachliche Entwicklung bedeutsam sein (Arbib, 2005). Auf die drei ersten und gleichzeitig prominenteren Ansätze soll im Folgenden kurz eingegangen werden.

Beobachtungslernen stellt eine grundlegende Fähigkeit des Menschen dar, die die Aneignung kulturell adäquaten Verhaltens erlaubt und weitere soziale Funktionen erfüllt (z.B. Empathie empfinden oder Zusammengehörigkeit zeigen; z.B. Wilson & Knoblich, 2005). Dies lässt plausibel erscheinen, dass das Beobachten eines bestimmten Verhaltens dasselbe Verhalten im Beobachter aktiviert, um so die Nachahmung zu erleichtern. In einer Untersuchung von Buccino et al. (2004b) zeigte sich tatsächlich eine Aktivierung des Spiegelneuronensystems während der Imitation von Handlungen.

Auch könnte Simulation dem Verstehen Anderer dienen. Die verschiedenen Varianten dieses Erklärungsstranges unterscheiden sich jedoch in der angenommenen Reichweite des Verstehens (Wilson & Knoblich, 2005). Dieses ist entweder nur auf die beobachtete Handlung beschränkt, umfasst auch die Gründe und Ziele der Handlung oder schließt sogar eine vollständige Repräsentation des geistigen Zustandes der anderen Person ein. Simulation könnte eine Person also in die Lage versetzen, zu erkennen, dass eine andere Person einen Akt des Greifens vollzieht und/ oder dass dieser Akt auf die Aufnahme von Nahrung ausgerichtet ist und/ oder dass diese Person Hunger empfindet. Empirische Evidenz liefert aus Sicht der VertreterInnen dieses Ansatzes bspw. eine Studie von Iacoboni et al. (2005), in der mittels fMRT die Aktivität des Gehirns gemessen wurde, während Personen entweder kurze Videos einer Handlung selbst, des Kontextes der Handlung oder der Handlung im Kontext zu sehen bekamen. Dabei wurde angenommen, dass durch den Kontext die Intention der Handlung zum Ausdruck gebracht würde. Verglichen mit den anderen beiden Bedingungen zeigte sich in der Bedingung „Handlung im Kontext“ eine stärkere Aktivität im posterioren Teil des inferioren Gyrus frontalis und dem ventralen prämotorischen Cortex. Die Forscher schlussfolgerten: „Thus, premotor mirror neuron areas — areas active during the execution and the observation of an action —

previously thought to be involved only in action recognition are actually also involved in understanding the intentions of others." (Jacoboni et al., 2005, S. e79; für eine Kritik siehe z.B. Borg, 2007).

Der dritte Ansatz kehrt die Relation von Wahrnehmung und Motorik um. Durch das Beobachten werden motorische Repräsentationen hervorgerufen, die wiederum die Beobachtung bzw. Wahrnehmung beeinflussen können. Dies geschieht im Sinne eines top-down-Prozesses, der auf die eigenen motorischen Erfahrungen zurückgreift, und dient der Vorhersage des Verhaltens der anderen Person. Graf, Reitzner, Corves, Casile, Giese und Prinz (2007) zeigten TeilnehmerInnen kurze Videos von Bewegungen, die durch Lichtpunkte repräsentiert wurden. Die Bewegungen wurden für ein gewisses Intervall durch eine Maske verdeckt. Dann erschien eine statische Pose und die TeilnehmerInnen wurden gebeten, einzuschätzen, ob es sich um eine Fortsetzung der Bewegung mit der gleichen Tiefenorientierung handelte. Die Leistung der TeilnehmerInnen war am besten, wenn die Testpose eine Fortsetzung der gesehenen Bewegung darstellte, die mit der Maskendauer korrespondierte. Offensichtlich simulierten die TeilnehmerInnen also die Bewegung in Echtzeit.

I.3.2.4 Weitere empirische Befunde

Weitere Evidenz für die Existenz von Simulation liefern verhaltensbasierte Studien, in denen die Effekte von Beobachtung auf die eigene Handlung untersucht wurden. So zeigten bspw. Brass, Bekkering, Wohlschläger und Prinz (2000), dass ihre TeilnehmerInnen eine Fingerbewegung schneller ausführen konnten, wenn sie eine Videoaufnahme derselben Bewegung ansahen als wenn sie eine willkürliche Bewegung sahen. In ähnlicher Weise konnten Kilner, Paulignan und Blakemore (2003) Interferenzeffekte aufzeigen (siehe auch Graf, Schütz-Bosbach & Prinz, in press). Ihre TeilnehmerInnen wurden gebeten, während der Ausführung von Armbewegungen eine andere Person zu beobachten, die entweder mit der eigenen Bewegung kongruente oder inkongruente Bewegungen ausübte. Bei Inkongruenz eigener und fremder Motorik wurden die Bewegungen der TeilnehmerInnen variabler.

Auch Ergebnisse von Studien an PatientInnen mit Schädigungen des Frontallappens liefern Evidenz für simulative Prozesse, da sie manchmal eine automatische und nicht-kontrollierbare Tendenz aufzeigen, beobachtete Handlungen zu imitieren (z.B. Brass, Derrfuss, Matthes-von Cramon & von Cramon, 2003).

Calvo-Merino, Glaser, Grèzes, Passingham und Haggard (2005) untersuchten mittels fMRT die Gehirnaktivität von TänzerInnen bei der Beobachtung des eigenen wie eines fremden Tanzstils und fanden eine stärkere Aktivierung bei der Beobachtung des eigenen Tanzstils. Sie folgerten, dass bei der Simulation beobachteter Handlungen auf das eigene Motorrepertoire zurückgegriffen bzw. dieses aktiviert würde: „Therefore, the parietal and premotor cortex mirror system does not respond simply to visual kinematics of body movement, but transforms visual inputs into the specific motor capabilities of the observer.“ (Calvo-Merino et al., 2005, S. 1248; siehe auch Calvo-Merino, Grèzes, Glaser, Passingham & Haggard, 2006; Buccino et al., 2004a).

Unabhängig von der Uneinigkeit, die bis dato darüber besteht, welche Funktion(en) Simulation erfüllt, herrscht also Einigkeit darüber, dass ein solches Phänomen existiert. Es kann definiert werden als jegliche Art (1) verborgener Aktivität des motorischen Systems, (2) die durch die Wahrnehmung von Ereignissen oder Objekten in der Umgebung des/ der Wahrnehmenden ausgelöst wurde, und (3) auf bedeutungsvolle Weise mit diesen Ereignissen oder Objekten zusammenhängt (siehe auch Graf, Schütz-Bosbach & Prinz, in press).

I.3.3 Hinterlässt Simulation Gedächtnisspuren?

Die bisherigen Befunde bezogen sich ausschließlich auf die Phase der Enkodierung. Die Aktivität verschiedener Hirnregionen wurde also während des Ausführens, Vorstellens oder Beobachtens gemessen. Wie bereits erwähnt, konnte in mehreren Studien nachgewiesen werden, dass auch bei der Erinnerung an de facto ausgeführte Handlungen motorische Areale (re)aktiviert werden – jedenfalls wenn es im Hinblick auf die vorliegende Erinnerungsaufgabe funktional erscheint (Heil et al., 1999; Masumoto et al., 2006; Nyberg et al., 2001; Senkfor, Van Petten & Kutas, 2008). Auf der Basis der dargestellten empirischen Befunde lässt sich fragen, ob auch nach Simulation, also beim Abruf von vorgestellten oder beobachteten Handlungen, eine derartige Aktivierung nachweisbar ist.

Nilsson et al. (2000), die als bildgebendes Verfahren die PET nutzten, untersuchten neben dem Ausführen und Lesen von Handlungsanweisungen auch das Vorstellen der Handlungsausführung. Die während des Abrufs erfasste Aktivität im rechten motorischen Cortex war zwar nach Ausführung am stärksten, aber auch das Vorstellen führte zu einer starken Aktivierung. Das Lesen rief die schwächste Gehirnaktivität in dem Areal hervor.

Senkfor, Van Petten und Kutas (2002) verglichen die Gehirnaktivität – gemessen anhand von ereigniskorrelierten Potentialen (ERP) – für alle drei Prozesse und schlossen zusätzlich eine nichtmotorische Kontrollbedingung ein, die jedoch auch aktives Generieren erforderte: Die TeilnehmerInnen wurden gebeten, den Preis von Objekten einzuschätzen (siehe auch Abschnitt I.2.3). Konkret wurden TeilnehmerInnen Objekte vorgelegt und sie wurden gebeten, mit diesen eine typische Handlung auszuführen, sich diese Handlung vorzustellen, zu beobachten, wie der Versuchsleiter diese Handlung ausführte oder den Preis des Objektes zu schätzen. Beim Abruf wurden den TeilnehmerInnen dann Photographien der Objekte vorgelegt und sie wurden gebeten, anzugeben, unter welcher der vier Bedingungen das Objekt enkodiert worden war (Quellengedächtnistest). Es zeigte sich, dass die Gehirnaktivität beim Abruf stark isomorph zu der Aktivität während des Enkodierens war: Die drei motorischen Prozesse riefen ähnliche ERPs über dem prämotorischen Cortex hervor, die deutlich unterscheidbar von denjenigen der Schätzaufgabe waren. Ausführen und Beobachten unterschieden sich vom Vorstellen und Einschätzen in der Aktivierung parietaler, posterior temporaler und okzipitaler Regionen, was die Forscherinnen darauf zurückführten, dass beim Ausführen und Beobachten Bewegung visuell wahrgenommen wird, beim Vorstellen und Schätzen jedoch nicht⁵. Schließlich wurde beim Abruf von vorgestellten Handlungen im präfrontalen Cortex eine stärkere Positivität gemessen, die möglicherweise mit der Aktivität dieser Areale bei der Enkodierung (also beim Vorstellen) zusammenhängt. Bzgl. des Vergleichs von Ausführen und Beobachten stellen die Autorinnen fest: „Perform and Watch conditions only differ quantitatively: Perform-encoded objects elicit more accurate judgments, faster reaction times, and more positive ERPs between 600 and 800 msec poststimulus onset but are otherwise similar in wave shape and scalp distribution.“ (Senkfor, Van Petten & Kutas, 2002, S. 414).

In ähnlicher Weise haben Stefan et al. (2005) anhand vom TMS nachweisen können, dass das Beobachten einer einfachen Bewegung des Daumens zu einer spezifischen motorischen Gedächtnisspur im primären motorischen Cortex (M1) führte, die noch zwei Tage nach der Enkodierung nachzuweisen war (siehe auch Stefan, Classen, Celnik & Cohen, 2008).

⁵ Dies stellt einen gewissen Widerspruch zu der vorher dargestellten Analogie von Wahrnehmung und Vorstellung dar, der von den AutorInnen jedoch nicht kommentiert wird.

I.4 Résumé und Ableitung der Fragestellung

Neben ihrer praktischen Relevanz und der Faszination, die von ihnen ausgeht, stellen Gedächtnistäuschungen wissenschaftlich ergiebige Phänomene dar, die dabei helfen können, die Funktionsweise des Gedächtnisses zu erhellen. So hat ihre Erforschung bspw. dazu beigetragen, Erinnern als einen (re)konstruktiven Prozess zu konzipieren (z.B. Roediger, 1996). Falsche Erinnerungen daran, ein Ereignis erlebt oder eine Handlung ausgeführt zu haben, sind besonders erstaunlich, da Ereignisse oder Handlungen normalerweise sehr gut erinnert werden (*enactment-Effekt*, z.B. Cohen, 1981; Engelkamp & Krumnacker, 1980). Innerhalb der kognitiven Psychologie wurde diese Form der episodischen Gedächtnistäuschung durch das bloße Vorstellen der Ereignisse bzw. Handlungen hervorgerufen, ein Phänomen, das als *imagination inflation* (Garry, Manning, Loftus & Sherman, 1996; Goff & Roediger, 1998) bekannt geworden ist. Die Verfälschung von Erinnerungen durch Vorstellungen wurde sowohl im Kontext von Kindheitsereignissen, als auch im Kontext von einfachen Handlungen untersucht, die vorteilhafterweise auch eine Kontrolle der Enkodierung ermöglichen und auf die daher im Folgenden fokussiert wird. In beiden Fällen wird durch den Prozess des Vorstellens einerseits das entsprechende Ereignis bzw. die Handlung überhaupt erst aktiviert und damit vertrauter, andererseits werden wahrnehmungsanalogue sensorische Eindrücke generiert.

Der imagination-inflation-Effekt für Kindheitsereignisse wurde vorwiegend auf eine Fehlattribution von Spurstärke zurückgeführt; zunehmend wird auch ein Einfluss qualitativer (sensorischer) Erinnerungsspuren nachgewiesen (z.B. Sharman et al., 2004, 2005a,b; Libby, 2003). Falsche Erinnerungen an die Ausführung einfacher Handlungen wurden hingegen vorwiegend auf eine Fehlattribution von durch die Vorstellung generierten, wahrnehmungsnahen Eindrücken zurückgeführt (z.B. Thomas, Bulevich & Loftus, 2003; Lampinen, Odegard & Bullington, 2003). Einer eindeutigen Interpretation steht jedoch im Wege, dass bisherige Operationalisierungen nicht den Ursprung der Information (internal vs. external) bzw. die Verarbeitungstiefe als weitere Einflussquellen berücksichtigten. So wurden bspw. Vorstellen und Lesen hinsichtlich ihres Potentials zur Induzierung falscher Erinnerungen verglichen – die beiden Prozesse unterscheiden sich jedoch nicht nur im Ausmaß involvierter sensorischer Eindrücke, sondern eben auch im Ursprung der Information – internal im Falle des Vorstellens, external im Falle des Lesens – und damit einhergehend in der Tiefe der Verarbeitung (Craik & Lockhart, 1972) bzw. dem Ausmaß von Vertrautheit. Die Frage, welcher Mechanismus ursächlich für falsche Erinnerungen an die Ausführung einfacher Handlungen ist, kann daher noch nicht endgültig beantwortet werden, zumal die Befundlage zur imagination inflation für Kindheitsereignisse auf eine kritische Rolle des (bei der Untersuchung einfacher Handlungen nicht berücksichtigten) Informationsursprungs bzw. der Verarbeitungstiefe hindeutet. Allerdings sind die Er-

kenntnisse aus beiden Bereichen nur äußerst schwer vergleichbar, da sie in vielerlei Hinsichten divergieren. So handelt es sich z.B. um kurz vs. lang zurückliegende bzw. komplexe vs. einfache Episoden.

Der imagination-inflation-Effekt ist bisher ausschließlich unter kognitionspsychologischer Perspektive betrachtet worden; neurowissenschaftliche Befunde blieben unberücksichtigt. Dabei weisen gerade diese auf einen möglichen weiteren ursächlichen Mechanismus hin: *motorische Simulation* (z.B. Jeannerod, 2001). In verschiedenen Studien wurde nachgewiesen, dass eine Überschneidung von (u.a.) motorischen Hirnarealen, die während der tatsächlichen und der bloß vorgestellten Handlungsausführung aktiviert sind, existiert (z.B. Grèzes & Decety, 2001). Es besteht also auf der Ebene motorischer Repräsentationen eine weitere Ähnlichkeit zwischen Vorstellung und Handlungsausführung, die evtl. ursächlich für den Effekt falscher Erinnerungen an Handlungen sein könnte.

Noch ein weiterer Prozess weist ähnliche Überlappungen mit der Ausführung – und daher auch gleichzeitig mit der Vorstellung – auf: Beobachtung (z.B. Grèzes & Decety, 2001). Ausgehend von der Entdeckung von Spiegelneuronen beim Affen, die sowohl bei der Produktion als auch Rezeption von Bewegung aktiv sind, mehren sich die Befunde, dass auch der Mensch die Handlungen seines Gegenübers automatisch simuliert – d.h. eine korrespondierende motorische Repräsentation entwickelt (z.B. Rizzolatti & Craighero, 2004). Greift also jemand nach einer Tasse, aktiviert die Beobachtung dieser Tätigkeit motorische Areale, die auch bei der Ausführung dieser Handlung aktiviert werden. Dass diese Areale beim Abruf reaktiviert werden und daher als Basis von Erinnerungsurteilen dienen könnten, wurde nicht nur für die eigene Ausführung, sondern auch für die Beobachtung – und im Übrigen auch für die Vorstellung – von Handlungen nachgewiesen (z.B. Nilsson et al., 2000; Senkfor, Van Petten & Kutas, 2002). Aber nicht nur die motorischen Repräsentationen von Ausführung, Vorstellung und Beobachtung ähneln sich. Beim Beobachten einer anderen Person werden – wie beim Ausführen und Vorstellen – sensorische Details enkodiert. Darüber hinaus wird natürlich durch Beobachtung auch die Handlung an sich aktiviert und damit vertrauter, d.h. die Spurstärke wird gesteigert.

Da die drei Prozesse des Ausführens, Vorstellens und Beobachtens von Handlungen also starke neurokognitive Überlappungen aufweisen, stellt sich die Frage, ob nicht – ebenso wie das Vorstellen – auch das Beobachten von Handlungen falsche Erinnerungen an die Ausführung dieser Handlungen hervorrufen könnte. Können also fremde Handlungen zu eigenen werden? Bzw.: Gibt es ein Phänomen, das man analog zur imagination inflation als *observation inflation* bezeichnen könnte?

Diese Frage ist vor dem geschilderten theoretisch-empirischen Hintergrund nicht eindeutig zu beantworten, denn den Ähnlichkeiten von Ausführung und Beobachtung stehen auch Unterschiede gegenüber. Vermutlich machen die allermeisten Menschen täglich die Erfahrung, dass es ihnen keinerlei Probleme bereitet, zu entscheiden, ob sie selbst oder andere Personen Akteure eines Ereignisses oder einer Handlung sind oder waren. Wir scheinen also über ausgeprägte Mechanismen der Unterscheidung von Selbst und Anderem zu verfügen. Dies ist erstaunlich, wenn man die zitierten neurowissenschaftlichen Befunde betrachtet. So finden bspw. Senkfor, Van Petten und Kutas (2002) im EEG rein quantitative Unterschiede bei der Erinnerung an die eigene vs. beobachtete Ausführung. Der Frage, wie Menschen Selbst und Andere(n) unterscheiden, wird in einer zunehmenden Anzahl an Studien nachgegangen. Schütz-Bosbach, Mancini, Aglioti und Haggard (2006) setzten die *rubber-hand-Illusion* ein, bei der ein Proband seine Hand unter eine Verdeckung platziert und eine Gummihand so vor sich liegen sieht als wäre sie seine eigene. Dann werden beide Hände simultan stimuliert – z.B. mit einem Pinsel gestreichelt. Dabei ruft synchrone Stimulation in der Person das Gefühl hervor, die fremde Hand sei die eigene – es wird also die Illusion einer Zugehörigkeit zum eigenen Körper erzeugt. Asynchrone Stimulation vermag dieses Erleben nicht zu erzeugen. Schütz-Bosbach et al. (2006) nutzten anstatt einer Gummihand die Hand einer anderen Person. Zunächst wurde entweder durch synchrone simultane Stimulation die Illusion hervorgerufen, dass die fremde Hand zum eigenen Körper gehören würde, oder diesem Erleben wurde durch asynchrone Stimulation entgegen gewirkt. Dann vollführte die fremde Hand Bewegungen mit dem Zeigefinger, die zu beobachten waren. Diese riefen die bekannte motorische Spiegelung nur dann hervor, wenn die ProbandInnen glaubten, die fremde Hand sei auch tatsächlich fremd – also nach vorhergegangener asynchroner Stimulation. Erlagen die ProbandInnen jedoch aufgrund synchroner Stimulation der Illusion, die fremde Hand sei de facto ihre eigene, trat nahezu der entgegengesetzte Effekt auf – es trat eine Inhibition ein. Demnach scheinen Selbst- und Fremdattribution bzw. Ausführung und Beobachtung doch auch auf qualitativ distinkter neuronaler Aktivität zu beruhen. Die Studie von Schütz-Bosbach et al. zeigt dies besonders eindrücklich, da durch die Nutzung der rubber-hand-Illusion eine hohe Vergleichbarkeit der Selbst- und Fremdwahrnehmung hergestellt wird (gleiche Perspektive, Ausschalten der Propriozeption usw.). Ob die Elimination der Unterschiede die Prozesse zu erhellen vermag, die während der tatsächlichen (bzw. simulierten) Ausführung außerhalb der hergestellten, artifiziellen Laborsituation ablaufen, bleibt jedoch fraglich, da die genannten Unterschiede ja eben charakteristisch für die Selbst-Fremd-Differenzierung unter gewöhnlichen Umständen sind. Ebenso müssen weitere Studien zeigen, inwieweit die sensomotorisch basierte, soziale Differenzierung zur bewussten Unterscheidung von Selbst und Anderem beiträgt. Andere ForscherInnen führen die Selbst-Fremd-Unterscheidung auf den Beitrag spezifischer Hirnstrukturen zurück, wie z.B. der extrastriate body area des okzipito-temporalen Cortex (EBA; David et

al., 2007) oder der rechten temporoparietalen Hirnregion/ den inferioren parietalen Cortex (Decety & Grèzes, 2006) (vgl. für eine Übersicht David, Newen & Vokeley, 2008).

Es lässt sich insgesamt festhalten, dass die Unterscheidung von Selbst und Anderem bzw. eigener und fremder Ausführung (*sense of agency*; Gallagher, 2000), deren neuronale Grundlage noch einzugrenzen ist, zu den basalen menschlichen Fähigkeiten gehört und sich auf mehreren Ebenen vollzieht. Die zentrale Dimension ist dabei sicherlich diejenige des Informationsursprungs – internal im Falle der Ausführung und external im Falle der Beobachtung. Diese Ebene ist identisch mit derjenigen des aktiven vs. passiven Status' der Person. Damit einhergehend können bei der eigenen Ausführung stets alle Sinne involviert sein, insb. ist sie durch propriozeptive Prozesse bzw. allgemeiner den Vergleich vorhergesagter und tatsächlicher sensorischer Konsequenzen gekennzeichnet. Derartige Vergleichsprozesse stellen sicherlich ein Kernstück der Selbst-Fremd-Differenzierung dar (siehe *comparator-Modell*, z.B. Blake-more & Frith, 2003; David, Newen & Vokeley, 2008; Farrer, Franck, Paillard & Jeannerod, 2003; Wolpert, Ghahramani & Jordan, 1995). Beim Beobachten werden jedoch nur visuelle und akustische, evtl. olfaktorische, Eindrücke enkodiert, wobei auf visueller Ebene die Perspektive variiert. Die Handlung wird bei eigener Ausführung aus der 1.- bei fremder Ausführung aus der 2.- oder 3.-Person-Perspektive beobachtet. Dabei soll unter 2.-Person-Perspektive verstanden werden, dass die Person ein frontales Gegenüber darstellt, während 3.-Person-Perspektive meint, dass die Person nicht als Gegenüber, sondern z.B. von der Seite beobachtet wird. Zudem stellt Simulation aufgrund von Beobachtung einen automatisch ablaufenden Prozess dar; die tatsächliche Ausführung kann entweder kontrolliert oder automatisch erfolgen.

Auch das Vorstellen, das ja bekanntermaßen zu Erinnerungstäuschungen führen kann, und das Beobachten unterscheiden sich in verschiedenen Hinsichten. Wiederum können beim Vorstellen einer Handlung alle Sinne einbezogen werden, während – wie bereits ausgeführt – die Beobachtung auf einem eingeschränkten Spektrum von Sinneseindrücken beruht, die dafür allerdings deutlich klarer sein dürften als die vorgestellten Eindrücke. Simulation durch Vorstellung ist ein bewusst initiiertes und kontrollierter Prozess des aktiven internalen Generierens sensorischer und motorischer Eindrücke – beobachtungsinduzierte Simulation läuft unbewusst und automatisch ab; die sensorischen Eindrücke werden passiv perzipiert, haben also einen externalen Ursprung, und auch die motorische Simulation wird external angestoßen. Daher handelt es sich bei der Vorstellung auch um den Prozess mit der größeren Verarbeitungstiefe, der gleichzeitig die im Vergleich zum Beobachten intensivere Form der Simulation darstellen könnte. Auf diesen Aspekt wird in Abschnitt I.5.4 noch näher eingegangen.

Insgesamt lassen sich Ausführung, Vorstellung und Beobachtung also hinsichtlich der folgenden Dimensionen klassifizieren und weichen auf diesen mehr oder weniger stark voneinander ab: Vertrautheit/ Verarbeitungstiefe, Informationsursprung, sensorische und motorische Eindrücke. Hierbei sind insb. die ersten beiden und die letzten beiden Merkmale nur schwer voneinander abzugrenzen.

Vor dem Hintergrund der dargestellten Gemeinsamkeiten und Unterschiede der drei Prozesse stellt sich zunächst die Frage, ob ein zum imagination-inflation- analoger *observation-inflation-Effekt* überhaupt existiert. Aufgrund der perfekten Kontrollierbarkeit und Manipulierbarkeit aller Phasen sollen im Rahmen dieser Arbeit falsche Erinnerungen an einfache Handlungen anhand des von Goff und Roediger (1998) entwickelten, dreistufigen Paradigmas untersucht werden. Um die Fragestellung beantworten zu können, soll im Rahmen der ersten Untersuchung in Anlehnung an Thomas, Bulevich und Loftus (2003) die Art der Verarbeitung in der zweiten Untersuchungsphase variiert werden – neben das Vorstellen soll das Beobachten treten. Gleichzeitig soll der mögliche Einfluss von Vertrautheit/ Verarbeitungstiefe und Informationsursprung kontrolliert werden. Daher sollen neben dem Vorstellen und Beobachten zwei weitere Verarbeitungsprozesse realisiert werden: Lesen und Generieren der Handlungsanweisungen. Durch das Lesen wird Vertrautheit mit den Handlungsanweisungen hervorgerufen. Da das Generieren – im Gegensatz zum Lesen – eine aktive Form der Verarbeitung darstellt, wird die Vertrautheit in dieser Bedingung sogar noch gesteigert; gleichzeitig weisen die generierten Handlungsanweisungen – analog der Vorstellung – einen internalen Ursprung auf. Dabei werden die Handlungsanweisungen durch beide Verarbeitungsprozesse weder mit sensorischen noch mit motorischen Repräsentationen angereichert – im Gegensatz zum Vorstellen und Beobachten⁶. Diese Überlegungen werden anhand der folgenden Tabelle systematisiert und zusammengefasst. Die getroffenen Einstufungen sind dabei nur als Anhaltspunkte gedacht.

⁶ Diese Vorstellung ist vereinfachend, wie Studien zur *embodied* bzw. *grounded cognition* zeigen, eine amodale Repräsentation wird hier gerade abgelehnt (z.B. Wilson, 2002; Barsalou, 2008). Dennoch sollen die Prozesse Vorstellen/ Beobachten vs. Generieren/ Lesen insofern als qualitativ distinkt angesehen werden, als sensorische bzw. motorische Informationen bei letzteren Verarbeitungsformen in deutlich geringerem Ausmaß und nur implizit aktiviert werden.

Tab. 5: Charakterisierung der Prozesse Ausführen, Vorstellen, Beobachten, Generieren und Lesen hinsichtlich der vier Dimensionen Vertrautheit, Informationsursprung, sensorische und motorische Eindrücke.

	Ausführen	Vorstellen	Beobachten	Generieren	Lesen
Vertrautheit/ Verarbeitungstiefe	++/--	++	+-	++	+-
Informations- ursprung	internal	internal	external	internal	external
Sensorische Eindrücke	+++	++	++	--	--
Motorische Eindrücke	+++	++	++	--	--

Anmerkungen: Die Dimension des internalen vs. externalen Informationsursprungs wird als identisch mit dem aktiven vs. passiven Status der Person angesehen. Die Einschätzung der Vertrautheit/ Verarbeitungstiefe bezieht sich ausschließlich auf den paarweisen Vergleich von Vorstellen und Beobachten bzw. Generieren und Lesen. Bzgl. der Ausführung sind beide Ausprägungen (++/--) angegeben, da diese sowohl kontrolliert als auch automatisiert ablaufen kann.

Die dieser Untersuchung zugrundeliegende Frage lautet anders formuliert, welche Informationen oder Erinnerungsqualitäten Menschen als diagnostisch für die Ausführung einer Handlung ansehen. Dies wird besonders dann salient, wenn sie fälschlicherweise erinnern, Handlungen ausgeführt zu haben, die sie tatsächlich niemals ausführten, sondern anderweitig verarbeitet, sich bspw. vorstellten. Aus dem theoretisch-empirischen Hintergrund ergeben sich insgesamt vier mögliche Mechanismen, die einer Verwechslung von Ausführung und Vorstellung und damit dem Zustandekommen von falschen Erinnerungen an die Handlungsausführung zugrunde liegen könnten: Eine Fehlattribution der Vertrautheit, die wiederum durch die Verarbeitungstiefe beeinflusst wird, des (internalen) Informationsursprungs, der sensorischen oder motorischen Repräsentationen. Diese vier möglichen Mechanismen sollen durch die Realisierung von vier verschiedenen Prozessen der Verarbeitung nicht ausgeführter Handlungsanweisungen auf ihr Zutreffen geprüft werden: Vorstellen, Beobachten, Generieren und Lesen. Je nachdem, welcher Mechanismus als (primär) ursächlich für die falschen Erinnerungen angesehen wird, ergeben sich unterschiedliche Vorhersagen des Ergebnismusters, die im Folgenden erläutert werden sollen. Allein die Annahme einer Fehlattribution von sensorischen vs. motorischen Merkmalen führt zu der gleichen Vorhersage, da sich beide Ebenen anhand der untersuchten Arten der Verarbeitung nicht trennen lassen.

Wäre eine Fehlattribution von Vertrautheit alleinig verantwortlich für den Effekt, sollte jegliche Art der Verarbeitung den Effekt hervorrufen, wobei tiefere Formen der Verarbeitung (Vorstellen, Generieren) zu einem größeren Effekt führen sollten als oberflächlichere (Beobachten, Lesen). Da aufgrund der Unterschiedlichkeit der Prozesse keine Anhaltspunkte dafür existieren, Vorstellen und Generieren bzw. Beobachten und Lesen in ihrer Verarbeitungstiefe voneinander abzugrenzen, kann kein differenzierteres Ergebnismuster vorhergesagt werden. Daher bezieht

sich auch der Vergleich bzgl. der Vertrautheit/ Verarbeitungstiefe in der vorausgegangenen Tabelle jeweils nur auf die beiden Prozesse des Vorstellens vs. Beobachtens und Generierens vs. Lesens. So ist bspw. zu erwarten, dass aufgrund der Anreicherung mit sensorischen/ motorischen Eindrücken Vorstellung und Beobachtung unabhängig von der Tiefe der Verarbeitung eine stärkere Gedächtnisspur erzeugen werden als Generieren und Lesen (*pictorial-superiority-Effekt*; Nelson, Reed & Walling, 1976).

Würde hingegen der interne Ursprung der Information den entscheidenden Faktor darstellen, sollten Vorstellen und Generieren zu falschen Erinnerungen führen, Beobachten und Lesen hingegen nicht.

Würde die Gedächtnistäuschung jedoch auf einer fälschlichen Zuschreibung von sensorischen Informationen beruhen, sollten wiederum Vorstellen und Beobachten einen inflation-Effekt hervorbringen, Lesen und Generieren jedoch nicht. Die gleiche Vorhersage resultiert, wenn motorische Repräsentationen ursächlich für die falschen Erinnerungen wären. Eine differenziertere Vorhersage dieser beiden Ansätze erscheint auf der Grundlage der berichteten Literatur nicht möglich – es ist also nicht vorherzusagen, ob Vorstellen oder Beobachten zu einem stärkeren Fehler führen sollte. Einerseits können beim Vorstellen mehr Sinne eingeschlossen werden, so dass die sensorische Repräsentation der tatsächlichen Handlungsausführung hinsichtlich des Spektrums ähnlicher sein könnte. Andererseits sind visuelle und akustische Eindrücke durch Beobachtung zweifelsohne klarer und lebendiger, und damit in ihrer Qualität der Ausführung ähnlicher als vorgestellte Eindrücke, wobei einschränkend anzumerken ist, dass sich Ausführung und Beobachtung in der visuellen Perspektive unterscheiden. Auch die Frage, ob bewusstes oder automatisch initiiertes Simulieren stärkere motorische Repräsentationen hervorruft, lässt sich vor dem skizzierten theoretisch-empirischen Hintergrund nicht beantworten. Daher kann – unabhängig davon, welcher Mechanismus als primär verantwortlich für den Effekt angesehen würde – keine Vorhersage über das Verhältnis beider Prozesse zueinander getätigt werden. Die Untersuchung würde somit darüber Auskunft geben, ob ein zum imagination-inflation- analoger observation-inflation-Effekt existiert bzw. ob falschen Erinnerungen an die Handlungsausführung eine Fehlattribution von sensorischen und/ oder motorischen Repräsentationen zugrunde liegen könnte. Sie würde aber keine Unterscheidung der letztgenannten Mechanismen erlauben.

Wie bereits dargelegt, weisen bisherige Untersuchungen zum imagination-inflation-Effekt auf die kritische Bedeutung sensorischer Informationen für das Zustandekommen von falschen Erinnerungen an die Handlungsausführung hin (wobei der Einfluss des Informationsursprungs nicht kontrolliert und die mögliche Bedeutung motorischer Repräsentationen nicht berücksich-

tigt wurde). Vor diesem Hintergrund wird das Zutreffen des letzten Ergebnismusters und damit als Mechanismus eine Fehlattribution sensorischer und/ oder (bisher nicht berücksichtigter) motorischer Repräsentationen für wahrscheinlich gehalten. Gegenstand der ersten Studie ist demnach die Überprüfung folgender Hypothese:

Falsche Erinnerungen an die Ausführung einfacher Handlungen basieren zumeist auf einer Fehlattribution von sensorischen und/ oder motorischen Repräsentationen, weshalb analog zum imagination-inflation ein observation-inflation-Effekt existiert. Sie werden zumeist nicht durch Vertrautheit oder den (internen) Ursprung der Information hervorgerufen.

Sollte das vorgefundene nicht dem vorhergesagten Ergebnismuster entsprechen, können die anderen Erklärungsvarianten in Betracht gezogen und a posteriori auf ihr Zutreffen überprüft werden. Selbstverständlich ist auch denkbar, dass die postulierten Mechanismen interagieren – z.B. könnten sowohl internaler Informationsursprung als auch sensorische Merkmale eine Rolle spielen und zusammengenommen besonders starke Erinnerungsverfälschungen produzieren. Auch diese Interaktionen führen zu eindeutigen Aussagen hinsichtlich der erwarteten Ergebnismuster und können ggf. überprüft werden.

I.5 Experiment I

I.5.1 Versuchsplanung

I.5.1.1 Operationalisierung und Variablenvalidität

Bei der Operationalisierung der unabhängigen und abhängigen Variablen wurde besonderes Augenmerk auf die Sicherung der Variablen- bzw. Ableitungsvalidität sowie auf die Präzision der Hypothesenprüfung gelegt.

Im Rahmen des ersten Experiments sollte geprüft werden, ob falsche Erinnerungen an die Ausführung einfacher Handlungen durch eine Fehlattribution von sensorischen und/ oder motorischen Repräsentationen hervorgerufen werden und demnach analog zum bekannten imagination-inflation- auch ein observation-inflation-Effekt existiert. Gleichzeitig sollte ausgeschlossen werden, dass der Ursprung der Information oder die gesteigerte Verarbeitungsleistung nach mehrfacher Darbietung ursächlich für die falschen Erinnerungen sind.

Um diese Ziele verfolgen zu können, wurde der Untersuchung das von Goff und Roediger (1998) entwickelte, experimentelle Paradigma zugrunde gelegt, das sich bereits in mehreren Studien, insbesondere der Vorläuferstudie von Thomas, Bulevich und Loftus (2003), als valide erwiesen hat. Das Paradigma umfasst die drei bereits erörterten Phasen der Enkodierung, Verarbeitung und des Abrufs, die in den bisherigen Studien durch unterschiedlich lange Intervalle voneinander getrennt wurden, wobei Enkodierung und Verarbeitung häufig zwei Tage, Enkodierung und Abruf häufig zwei Wochen auseinander lagen. Essentiell sind zwei intraindividuell variierte, unabhängige Variablen: die Art der Enkodierung der Handlungsanweisungen in der ersten Phase (ausgeführt vs. nicht ausgeführt) und die Frequenz der Verarbeitung in der zweiten Phase, die in der ursprünglichen Untersuchung von 0 bis 5 Mal mit verschiedenen Zwischenstufen reichte.

Als Alternative zum Ausführen kam als Art der Verarbeitung in der ersten Phase häufig schon das Vorstellen zum Einsatz, d.h. TeilnehmerInnen wurden gebeten, Handlungen entweder auszuführen oder sich die Ausführung vorzustellen (z.B. Goff & Roediger, 1998; Thomas et al., 2003). Da nun aber gerade das Vorstellen der eigenen Handlungsausführung nachweislich zu einer Verzerrung der Erinnerung führen kann, sollte es durch einen anderen, der Handlungsausführung möglichst unähnlichen, Prozess ersetzt werden. Den Ergebnissen von Thomas et al. (2003) zufolge stellt das Lesen einen solchen, zum Ausführen stark distinkten, Prozess dar: So wurden Handlungsanweisungen, die in der ersten Phase nicht enkodiert und in der zweiten Phase nur gelesen wurden, fast nie als ausgeführt erinnert. Bei

Phase nur gelesen wurden, fast nie als ausgeführt erinnert. Bei Handlungsanweisungen, die in der ersten Phase vorgestellt und in der zweiten gelesen wurden, lag diese Quote schon deutlich höher. Auch wird das Lesen in Studien zum enactment-Effekt stets der Ausführung gegenübergestellt (vgl. Abschnitt I.2.3). Anstatt TeilnehmerInnen in der ersten Phase zu bitten, Handlungen entweder auszuführen oder sich deren Ausführung vorzustellen, sollten sie also gebeten werden, Handlungsanweisungen zu lesen und anschließend tatsächlich umzusetzen oder nicht umzusetzen. Diese zwei äußerst unterschiedlichen Formen der Enkodierung sollten es ermöglichen, den Einfluss der verschiedenen Verarbeitungsprozesse der zweiten Phase möglichst unverzerrt bzw. rein zu erfassen. Zum selben Zwecke sollte – wie in den bisherigen imagination-inflation-Studien – eine dritte Art von Items nicht in der ersten, sondern erst in der zweiten Phase präsentiert werden, so dass der Faktor „Art der Enkodierung in Phase 1“ insgesamt die drei Stufen ausgeführt, gelesen und nicht-enkodiert umfasste. Als Frequenzen der Verarbeitung in der zweiten Phase sollten in Anlehnung an Thomas, Bulevich und Loftus (2003) neben der Baseline 0x die Stufen 1x und 5x realisiert werden, wobei der Vergleich zwischen den Stufen 0x und 5x zentral ist (Näheres hierzu ist dem Abschnitt I.5.1.3 zur Stichprobengröße zu entnehmen). Insgesamt umfasste also die Basis-Versuchsplananlage 9 Itemarten bzw. Zellen – Items, die in der ersten Phase ausgeführt bzw. gelesen bzw. nicht enkodiert und in der zweiten Phase 5x bzw. 1x bzw. 0x dargeboten wurden.

Zwischen Enkodierung und Verarbeitung vergingen in den Studien von Thomas et al. (2002, 2003) zwei Tage, zwischen Enkodierung und Abruf zwei Wochen. Aufgrund der ausgesprochen guten Erinnerbarkeit von Handlungen wären sicherlich keine oder kaum Fehler zu erwarten, wenn die Abfrage direkt oder auch nur kurz nach dem Enkodieren erfolgen würde (siehe z.B. Thomas & Bulevich, 2006). Daher sollte das bewährte Intervall von zwei Wochen beibehalten werden. Die deutliche zeitliche Trennung von erster und zweiter Phase bzw. Enkodierung und Verarbeitung wurde jedoch als kritisch betrachtet, da diese zusätzliche Hinweisreize bzw. Merkmale einbringt, die beim Erinnerungstest eine Rolle spielen könnten. So könnte eine Person sich z.B. ganz sicher sein, dass eine Handlungsanweisung überhaupt erst in der zweiten Sitzung präsentiert wurde, und daraus schließen, dass sie die Handlung gar nicht ausgeführt haben konnte – da in der zweiten Sitzung keine einzige Handlung ausgeführt wurde. Dadurch, dass die Sitzungen derart stark separiert sind, wird es also ermöglicht bzw. erleichtert, das Merkmal „Zeitpunkt der Präsentation“ beim Erinnerungstest zu berücksichtigen. Wie das obige Beispiel zeigt, kann dadurch eine Person verleitet werden, gar nicht aufgrund von Erinnerungsspuren zu urteilen, die für die Handlungsausführung diagnostisch sind, sondern aufgrund anderer Kriterien. Da aber im Rahmen der anvisierten Untersuchung gerade die Frage im Mittelpunkt stand, welche Erinnerungsqualitäten als diagnostisch für die Handlungsausführung herangezogen werden, sollten Enkodierung und Verarbeitung möglichst zeitnah aufeinander folgen. Hierdurch sollte die Möglichkeit des Einflusses weiterer Hinweisreize (z.B.

einander folgen. Hierdurch sollte die Möglichkeit des Einflusses weiterer Hinweisreize (z.B. des Zeitpunktes der Informationsdarbietung) beim Abruf minimiert und gleichzeitig gewährleistet werden, dass Erinnerungsurteile überhaupt auf den fraglichen Merkmalen (Vertrautheit, Informationsursprung, sensorische und motorische Repräsentationen) basiert würden. Daher sollten beide Phasen innerhalb einer Sitzung stattfinden und nur durch eine kurze Pause getrennt werden. Dieses Vorgehen ist gleichzeitig deutlich ökonomischer, da TeilnehmerInnen so nicht drei-, sondern „nur“ zweimal zur Untersuchungsteilnahme erscheinen müssen. Darüber hinaus hat die Untersuchung von Goff und Roediger (1998) gezeigt, dass durch das Zusammenlegen der Phasen 1 und 2 keine bedeutsamen Unterschiede im Ergebnismuster hervorgerufen werden.

Neben den beiden bereits benannten, intraindividuell variierten unabhängigen Variablen war für die anvisierte Untersuchung die Variation der Art der Verarbeitung der Handlungsanweisungen in der zweiten Phase zentral. Wie bereits beschrieben, sollte neben dem Vorstellen der Einfluss des Beobachtens, Generierens und Lesens auf die Erinnerungsurteile erfasst werden. Eine intraindividuelle Variation der Verarbeitungsart schloss sich allein schon aufgrund des Materialumfangs aus – würde bspw. die Zellgröße um ein Item erhöht, würde dies aufgrund der bis zu fünffachen Wiederholung einen Zuwachs um 12 Präsentationen in Phase 2 bedeuten. Wenn also bei intraindividuelle Bedingungsvariation eine angemessene Anzahl von Beobachtungen pro Zelle gewährleistet werden sollte, würde dies eine Vervielfachung der Präsentationen in Phase 2 zur Folge haben, die schließlich eine nicht mehr zumutbare Belastung darstellen würde. Darüber hinaus erschien es fraglich, ob die Prozesse bei intraindividuelle Präsentation überhaupt noch unabhängig voneinander bzw. rein erfasst werden könnten. So ist z.B. denkbar, dass sich eine Teilnehmerin trotz der Aufforderung, die nächste Handlungsanweisung nur zu lesen, die entsprechende Ausführung ebenfalls vorstellt, da sie dies mit einer vorherigen Handlungsanweisung ebenfalls getan hat. Daher wurde insgesamt die Lösung favorisiert, die Art der Verarbeitung interindividuell zu variieren (siehe auch Thomas et al., 2003).

Von den vier Formen der Verarbeitung in Phase 2 – Vorstellen, Beobachten, Generieren, Lesen – sind in bisherigen Untersuchungen nur zwei zum Einsatz gekommen, Vorstellen und Lesen. Dabei hat sich gezeigt, dass bei möglichst lebendiger *Vorstellung* – d.h. Vorstellung, die möglichst viele Sinnesmodalitäten einschließt – die Erinnerungsurteile fehleranfälliger sind (Thomas et al., 2003). Daher sollten TeilnehmerInnen auch in der anvisierten Untersuchung aufgefordert werden, sich die Handlungsausführung möglichst lebhaft vorzustellen. Da für das *Lesen* keine Freiheitsgrade in der Umsetzung existieren, muss an dieser Stelle nicht näher auf diese Art der Verarbeitung eingegangen werden. In der konkreten Umsetzung war allerdings

sicherzustellen, dass die TeilnehmerInnen keine Gelegenheit hatten, die Handlungsanweisungen tiefer zu verarbeiten und/ oder mit sensorischen/ motorischen Informationen anzureichern. Für das *Beobachten* wären grundsätzlich zwei verschiedene Formen der Umsetzung in Betracht gekommen: Das Beobachten einer tatsächlich während der Untersuchung anwesenden Person, die die Handlungen direkt vor den Augen der TeilnehmerInnen ausführt, oder das Beobachten von Videos einer Person, die zuvor bei der Handlungsausführung gefilmt wurde. Im Vergleich spricht für die erste Variante, dass die Handlungsausführung in „Lebensgröße“ und dreidimensional beobachtet würde, also der eigenen Handlungsausführung ähnlicher wäre als eine Präsentation per Video, die das Ganze in verkleinertem Maßstab und zweidimensional wiedergibt. Darüber hinaus wäre durch die Präsenz des Modells gewährleistet, dass die TeilnehmerInnen der Handlungsausführung auch wirklich Aufmerksamkeit schenken, während die Aufmerksamkeit bei der Verfolgung der Videos nicht ständig kontrolliert werden könnte. Als nachteilig bei der Beobachtung in realiter wird jedoch erachtet, dass keine perfekte Standardisierung möglich erscheint. So könnte die Versuchsleiterin natürlich trainiert werden, die Handlungen möglichst auf die gleiche Art und Weise und in gleichem Tempo so auszuführen, dass alle TeilnehmerInnen möglichst exakt dasselbe beobachten – allerdings würde es immer Abweichungen geben, eine perfekte Vergleichbarkeit könnte nicht gewährleistet werden. Durch das Zeigen eines Videos jedoch wäre absolut standardisiert, was TeilnehmerInnen beobachten. Ein weiteres Problem der Beobachtung einer tatsächlich anwesenden Person ist, dass es zu einer Interaktion zwischen Versuchsleiterin und TeilnehmerInnen kommen könnte, so dass TeilnehmerInnen von der Handlungsausführung abgelenkt sein und z.B. die Mimik der Versuchsleiterin beobachten oder sich weitergehende Gedanken machen könnten (z.B. ob die Handlung typisch für die Person ist usw.). Würden Videoaufzeichnungen genutzt, könnte der Ausschnitt ausschließlich auf die Handlung fokussieren. Gleichzeitig könnten andere AkteurInnen als die Versuchsleiterin gewählt werden, um durch diese Anonymität zu vermeiden, dass derartige Gedanken, die eine tiefere Form der Verarbeitung bedeuten bzw. zusätzliche Merkmale darstellen, die beim Abruf eine Rolle spielen könnten, überhaupt angestoßen bzw. weiterverfolgt würden. Auch vor dem Hintergrund der Tatsache, dass es für die Versuchsleiterin eine kaum zumutbare Belastung dargestellt hätte, die Handlungen (immerhin über einen Zeitraum von jeweils ca. 40 Minuten) stets selbst auszuführen, wurde die zweite Variante der Videodarbietung gewählt. Bzgl. des *Generierens* der Handlungsanweisungen sind – verglichen mit den anderen Formen der Verarbeitung – die meisten Alternativen denkbar. In der Literatur werden z.B. das Lösen von Anagrammen oder Umstellen von Sätzen sowie das Paraphrasieren genutzt (Bernstein, Whittlesea & Loftus, 2002; Sharman, Garry & Beuke, 2004). Aufgrund der mehrfachen Präsentation in Phase 2 kam allerdings nur eine Variante in Frage, die eine Steigerung des Schwierigkeitsgrades erlaubte. Hiermit sollte sichergestellt werden, dass die Handlungsanweisungen auch tatsächlich jedes Mal erneut aktiv generiert und nicht automatisch reprodu-

ziert würden. Somit schied – insbesondere vor dem Hintergrund der sehr einfachen Handlungsanweisungen – das Paraphrasieren aus. Am geeignetsten erschienen umgestellte Handlungsanweisungen bzw. Satzanagramme, deren Schwierigkeitsgrad steigerbar ist.

Die zur Beantwortung der Untersuchungsfragen hauptsächlich interessante abhängige Variable stellten falsche *ausgeführt*-Antworten im Quellengedächtnistest dar. Zur Erhebung dieser Urteile waren zwei Varianten denkbar. Einerseits hätte ein sogenannter „forced-choice“-Quellengedächtnistest eingesetzt werden können, bei dem ausschließlich präsentierte Items vorgelegt werden, die entlang der Quellendimension (hier: *ausgeführt* – nicht *ausgeführt*) beurteilt werden sollen. Daneben war denkbar, den Quellen- mit einem Rekognitionstest für die erste Phase zu verbinden; d.h. alte und neue Handlungsanweisungen zu präsentieren und nur als alt eingestufte Items auf ihre Quelle hin beurteilen zu lassen (z.B. Glisky, Polster & Routhieaux, 1995). Da die Rekognitionsdaten für die geplante Untersuchung in zweierlei Hinsicht relevant waren – einerseits zum Nachweis des enactment-Effekts, andererseits zum Nachweis des Einflusses von Vertrautheit bzw. Verarbeitungstiefe, wobei beides als Manipulation-Check verstanden wird, – wurde letztere Option gewählt. Diese Variante wurde bisher in allen einschlägigen Untersuchungen genutzt, wobei der Rekognitions- dem Quellengedächtnistest vorgeschaltet war, so dass ein Quellenurteil nur nach erfolgter Rekognition abgegeben wurde (z.B. Goff & Roediger, 1998; Thomas et al., 2003). In der Untersuchung von Thomas und Bulevich (2006) wurde dagegen erstmals eine kombinierte Variante eingesetzt, bei der alle drei Antwortoptionen (alt und *ausgeführt*, alt und nicht *ausgeführt*, neu) gleichzeitig dargeboten wurden. Da beide Vorgehensweisen zum selben Ergebnismuster zu führen scheinen, erstere jedoch die häufiger genutzte und informationshaltigere Variante darstellt, die zudem durch die Separierung der Erinnerungsurteile für TeilnehmerInnen leichter und eindeutiger zu verstehen sein sollte, sollte diese Variante gewählt, d.h. Rekognitions- und Quellengedächtnisurteile separat erfasst werden. Dabei sollte der Tatsache Rechnung getragen werden, dass diese Urteile auf unterschiedlich klaren Erinnerungsspuren basieren und daher mit unterschiedlich starker Sicherheit abgegeben werden können. Daher sollte zusätzlich zu den Erinnerungsurteilen auch die Sicherheit, mit der diese Urteile gefällt wurden, erhoben werden.

Ein Ablaufschema ist der folgenden Tabelle zu entnehmen.

Tab. 6: Ablauf des ersten Experiments.

Phase	Variable	Ausprägungen	Art der Variation
Phase 1: Enkodierung	UV-A: Art der Enkodierung	ausgeführt	intraindividuell
		gelesen nicht enkodiert	
Phase 2: Verarbeitung*	UV-B: Frequenz der Verarbeitung	5x	intraindividuell
		1x 0x	
Phase 2: Verarbeitung*	UV-C: Art der Verarbeitung	vorstellen	interindividuell
		beobachten generieren lesen	
Phase 3: Abruf**	AV-1: Rekognition	aus Phase 1	--
		nicht aus Phase 1	
	AV-2: Quellen- gedächtnis	ausgeführt gelesen	--
	AV-3: Sicherheit	siebenstufige Skala	--

Anmerkungen: * 5 Minuten nach Phase 1; ** 2 Wochen nach Phase 2.

Ein inflation-Effekt manifestiert sich im Rahmen des vorgestellten Paradigmas in einem anhand des Quellengedächtnistests erfassten, erhöhten Ausmaß an *ausgeführt*-Antworten auf in der ersten Phase de facto nicht ausgeführte (gelesene oder nicht enkodierte) Handlungsanweisungen, die in der zweiten Phase fünfmal verarbeitet wurden, verglichen mit Handlungsanweisungen, die in der zweiten Phase nicht verarbeitet wurden. Demnach lässt sich die aufgestellte Hypothese überprüfen, indem für jede Form der Verarbeitung separat geprüft wird, ob nach fünfmaliger im Vergleich zu nullmaliger Verarbeitung in Phase 2 im Durchschnitt mehr falsche *ausgeführt*-Antworten gegeben werden. Angenommen wurde, dass dies nach Vorstellen und Beobachten, nicht jedoch nach Generieren oder Lesen der Fall ist. Um eine strenge Hypothesenprüfung zu gewährleisten, sollte jedoch zusätzlich der Einfluss des Informationsursprungs und der Verarbeitungstiefe/ Vertrautheit kontrolliert werden. Hierfür sollte das Ausmaß falscher Erinnerungen nach Vorstellen mit demjenigen nach aktivem Generieren verglichen werden, ebenso wie das Ausmaß falscher Erinnerungen nach Beobachten demjenigen nach bloßem Lesen der Handlungsanweisungen gegenübergestellt werden sollte. Beide Vergleiche sollten signifikante Interaktionen erbringen⁷.

⁷ Ein Vergleich von Beobachten und Lesen war u.a. notwendig, da auch im Rahmen des Beobachtens die Handlungsanweisungen bei jeder Präsentation gelesen werden sollten; siehe Abschnitt I.5.2.4. Dem Vorgehen liegt also die idealisierte Vorstellung zugrunde, dass der Einfluss sensorischer und motorischer Repräsentationen auf den inflation-Effekt erfasst werden könnte, indem eine „Bereinigung“ um Informationsursprung bzw. Vertrautheit in Form einer signifikanten Interaktion zwischen der Frequenz der Präsentation (5x vs. 0x) und

Demnach müssen für das Zutreffen der Hypothese

Falsche Erinnerungen an die Ausführung einfacher Handlungen basieren zu-
meist auf einer Fehlattribution von sensorischen und/ oder motorischen Reprä-
sentationen, weshalb analog zum imagination-inflation- ein observation-
inflation-Effekt existiert. Sie werden zumeist nicht durch Vertrautheit oder den
internalen Ursprung der Information hervorgerufen.

alle im Folgenden aufgeführten 6 Bedingungen erfüllt sein (die Notation bezieht sich auf den
im folgenden Abschnitt präsentierten Versuchsplan):

- (1) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu
nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen sich in der zwei-
ten Phase die Handlungsausführung vorstellten.

$$H_1: \mu_{(A_2B_1C_1+A_3B_1C_1)} > \mu_{(A_2B_3C_1+A_3B_3C_1)}, H_0: \mu_{(A_2B_1C_1+A_3B_1C_1)} \leq \mu_{(A_2B_3C_1+A_3B_3C_1)}$$

- (2) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu
nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der zweiten
Phase die Handlungsausführung beobachteten.

$$H_1: \mu_{(A_2B_1C_2+A_3B_1C_2)} > \mu_{(A_2B_3C_2+A_3B_3C_2)}, H_0: \mu_{(A_2B_1C_2+A_3B_1C_2)} \leq \mu_{(A_2B_3C_2+A_3B_3C_2)}$$

- (3) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu
nullmaliger Präsentation nicht signifikant höher, wenn die TeilnehmerInnen in der
zweiten Phase die Handlungsanweisungen generierten.

$$H_0: \mu_{(A_2B_1C_3+A_3B_1C_3)} \leq \mu_{(A_2B_3C_3+A_3B_3C_3)}, H_1: \mu_{(A_2B_1C_3+A_3B_1C_3)} > \mu_{(A_2B_3C_3+A_3B_3C_3)}$$

- (4) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu
nullmaliger Präsentation nicht signifikant höher, wenn die TeilnehmerInnen in der
zweiten Phase die Handlungsanweisungen lasen.

$$H_0: \mu_{(A_2B_1C_4+A_3B_1C_4)} \leq \mu_{(A_2B_3C_4+A_3B_3C_4)}, H_1: \mu_{(A_2B_1C_4+A_3B_1C_4)} > \mu_{(A_2B_3C_4+A_3B_3C_4)}$$

- (5) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und
nullmaliger Präsentation ist signifikant höher, wenn die TeilnehmerInnen sich in der
zweiten Phase die Handlungsausführung vorstellten, als wenn sie die Handlungsanwei-
sungen generierten.

$$H_1: \mu_{(A_2B_1C_1+A_3B_1C_1)} - \mu_{(A_2B_3C_1+A_3B_3C_1)} > \mu_{(A_2B_1C_3+A_3B_1C_3)} - \mu_{(A_2B_3C_3+A_3B_3C_3)},$$

$$H_0: \mu_{(A_2B_1C_1+A_3B_1C_1)} - \mu_{(A_2B_3C_1+A_3B_3C_1)} \leq \mu_{(A_2B_1C_3+A_3B_1C_3)} - \mu_{(A_2B_3C_3+A_3B_3C_3)}$$

- (6) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und
nullmaliger Präsentation ist signifikant höher, wenn die TeilnehmerInnen in der zwei-
ten Phase die Handlungsausführung beobachteten, als wenn sie die Handlungsanwei-
sungen lasen.

$$H_1: \mu_{(A_2B_1C_2+A_3B_1C_2)} - \mu_{(A_2B_3C_2+A_3B_3C_2)} > \mu_{(A_2B_1C_4+A_3B_1C_4)} - \mu_{(A_2B_3C_4+A_3B_3C_4)},$$

der Art der Verarbeitung (Vorstellen vs. Generieren bzw. Beobachten vs. Lesen) hinsichtlich falscher *ausge-
führt*-Antworten gefordert wird, um die Hypothese annehmen zu können. Diese Annahme ist sicherlich vereinfachend (zumal das Vorstellen und Generieren sehr unterschiedliche Prozesse involvieren), stellt aber im Ver-
gleich zu vorhergehenden Untersuchungen eine Verbesserung dar.

$$H_0: \mu_{(A2B1C2+A3B1C2)} - \mu_{(A2B3C2+A3B3C2)} \leq \mu_{(A2B1C4+A3B1C4)} - \mu_{(A2B3C4+A3B3C4)}$$

Insgesamt wurden also – sofern möglich – bewährte Operationalisierungen aufgegriffen und entweder direkt übernommen oder modifiziert, wenn dies hinsichtlich der untersuchungsleitenden Hypothese sinnvoller bzw. valider erschien. Auch die Überlegungen zur Operationalisierung von Konstrukten, die in der bisherigen Forschung zum imagination-inflation-Effekt noch gar nicht eingesetzt worden sind, wurden sorgfältig hinsichtlich ihrer validitäts- und präzisionsbezogenen Vor- und Nachteile abgewogen. Stets wurde eine möglichst strenge Hypothesenprüfung angestrebt.

In den folgenden Abschnitten werden die Überlegungen zur Operationalisierung noch konkretisiert.

I.5.1.2 Versuchsplananlage und Versuchsplan

Insgesamt lassen sich die zur Operationalisierung und Variablenvalidität angestellten Überlegungen folgendermaßen zusammenfassen: Der Untersuchung lag ein 3 (UV-A: Art der Enkodierung in der 1. Phase: ausgeführt, gelesen, nicht enkodiert) x 3 (UV-B: Frequenz der Präsentation in der 2. Phase: 5x, 1x, 0x) x 4 (UV-C: Art der Verarbeitung in der 2. Phase: Vorstellen, Beobachten, Generieren, Lesen) Versuchsplan zugrunde, wobei die ersten beiden Variablen intra- und die dritte interindividuell variiert wurden. Als abhängige Variablen wurden Reaktions- und Quellengedächtnisurteile sowie Sicherheitsratings bzgl. der Erinnerungsurteile erhoben. Es resultierte also eine dreifaktorielle Versuchsplananlage (VPL-A3, Hussy & Jain, 2002; siehe Tabelle 7, folgende Seite).

Anhand der folgenden vier Entscheidungen ließ sich diese noch abstrakte und auf strukturelle Aspekte beschränkte Versuchsplananlage in einen konkreten Versuchsplan überführen, der nun auch strategische Aspekte umfasste und somit einer validen, präzisen und ökonomischen Versuchsplanung dienen sollte (z.B. Hager, 1987; Hussy & Jain, 2002).

Um den vollen Informationsgehalt der Versuchsplananlage auszuschöpfen, sollten alle angelegten Zellen auch realisiert werden, so dass sich ein vollständig gekreuzter Versuchsplan ergab.

In allen Zellen sollte die gleiche Anzahl an Versuchspersonen beobachtet werden, so dass sich ein balancierter (und orthogonaler) Versuchsplan ergab.

Wie bereits dargelegt, mussten UV-A und UV-B intraindividuell variiert werden, um überhaupt den inflation-Effekt, d.h. mehr *ausgeführt*-Antworten auf de facto nicht ausgeführte Handlungsanweisungen nach fünfmaliger verglichen mit nullmaliger Präsentation, nachweisen zu

können. Die UV-C sollte aufgrund von praktischen und inhaltlichen Überlegungen interindividuell variiert werden, so dass sich insgesamt ein gemischter Versuchsplan ergab.

Bei dem interindividuell variierten Faktor sollte eine randomisierte Zuweisung der ProbandInnen zu den Versuchsbedingungen erfolgen, so dass es sich um einen experimentellen Faktor handelt. Bei den intraindividuell variierten Faktoren handelt es sich per definitionem zunächst um quasi-experimentelle Faktoren; im Folgenden wird jedoch dafür argumentiert, diese als experimentelle Faktoren im eigentlichen Sinne aufzufassen, so dass in der Nomenklatur von Hussy und Jain (2002) folgendes Design zugrunde liegt: Vpl3-RRR.

Tab. 7: Versuchsplan des ersten Experiments.

				UV-C: Art der Verarbeitung in Phase 2			
				Vorstellen	Beobachten	Generieren	Lesen
UV-A: Art der Enkodierung in Phase 1	ausgeführt	UV-B: Frequenz der Präsentation in Phase 2	5x	A ₁ B ₁ C ₁	A ₁ B ₁ C ₂	A ₁ B ₁ C ₃	A ₁ B ₁ C ₄
			1x	A ₁ B ₂ C ₁	A ₁ B ₂ C ₂	A ₁ B ₂ C ₃	A ₁ B ₂ C ₄
			0x	A ₁ B ₃ C ₁	A ₁ B ₃ C ₂	A ₁ B ₃ C ₃	A ₁ B ₃ C ₄
	gelesen		5x	A ₂ B ₁ C ₁	A ₂ B ₁ C ₂	A ₂ B ₁ C ₃	A ₂ B ₁ C ₄
			1x	A ₂ B ₂ C ₁	A ₂ B ₂ C ₂	A ₂ B ₂ C ₃	A ₂ B ₂ C ₄
			0x	A ₂ B ₃ C ₁	A ₂ B ₃ C ₂	A ₂ B ₃ C ₃	A ₂ B ₃ C ₄
	nicht enkodiert		5x	A ₃ B ₁ C ₁	A ₃ B ₁ C ₂	A ₃ B ₁ C ₃	A ₃ B ₁ C ₄
			1x	A ₃ B ₂ C ₁	A ₃ B ₂ C ₂	A ₃ B ₂ C ₃	A ₃ B ₂ C ₄
			0x	A ₃ B ₃ C ₁	A ₃ B ₃ C ₂	A ₃ B ₃ C ₃	A ₃ B ₃ C ₄

I.5.1.3 Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang

Um die optimale Stichprobengröße berechnen zu können, mussten Signifikanzniveau, Betafehler und zu erwartende Effektgröße vor der Untersuchungsdurchführung spezifiziert werden. In der Literatur zum imagination-inflation-Effekt werden jedoch keine Effektgrößen berichtet. Es ließ sich aber vor dem Hintergrund der berichteten Parameter schätzen, dass nach in der zweiten Phase stattfindender fünfmaliger verglichen mit nullmaliger Präsentation von Handlungsanweisungen, die in der ersten Phase nicht ausgeführt wurden, der Effekt als groß zu klassifizieren ist. Beim Vergleich der Frequenzen 1x und 0x fällt er hingegen deutlich schwächer aus. Da keine Anhaltspunkte dazu vorlagen, welche Effektgröße ein möglicherweise vorliegender observation-inflation-Effekt aufweisen könnte, wurde bei der Versuchsplanung von der Größe des imagination-inflation-Effekts ausgegangen. Hierauf wiesen auch die Ergebnisse der Voruntersuchung hin. Da die Planung auf einem Vergleich der Frequenzen 5x und 0x und

somit auf dem Nachweis eines großen Effekts basierte, wird die Frequenz 1x bei der Hypothesenprüfung nicht berücksichtigt, soll aber v.a. aus versuchsplanerischen Gründen (siehe weiter unten) beibehalten werden. Des Weiteren fällt der Effekt in der vorliegenden Literatur stets etwas kleiner aus für solche Handlungsanweisungen, die in der 1. Phase gar nicht präsentiert (enkodiert) wurden im Vergleich zu jenen, die präsentiert, aber nicht ausgeführt wurden. Die zu erwartende Effektgröße wurde daher auf das Mittel von beiden Itemarten bezogen (siehe auch Goff & Roediger, 1998).

Zur Ermittlung der optimalen Stichprobengröße wurden Alpha- und Betafehler – der Konvention folgend – in ein Verhältnis von 1:4 gebracht, so dass die Wahrscheinlichkeit eines Alphafehlers auf 5% und die des Betafehlers auf 20% beziffert wurde (siehe z.B. Hussy & Jain, 2000). Dieses Vorgehen ist für die Hypothesenprüfung insofern problematisch, als auch Nullhypothesen zu testen waren, deren Beibehaltung bzw. Nachweis dann mit einer vergleichsweise großen Fehlerwahrscheinlichkeit behaftet wäre. Dennoch wurde aus zwei Gründen an dieser Vorgehensweise festgehalten: Die zu überprüfende theoretische Hypothese wurde durch sechs Unterhypothesen operationalisiert, die mehrheitlich Alternativhypothesen darstellten. Bei der Prüfung jeder dieser Unterhypothesen sollte jedoch an einer einheitlichen Festlegung von Alpha- und Betafehler festgehalten werden. Außerdem sollte die theoretische Hypothese letztlich nur im Zusammenspiel der Entscheidungen über die Null- und (gemeint ist die logische und-Verknüpfung) Alternativhypothesen bewertet werden. Es sollten also alle 6 Hypothesen zutreffen, um die theoretische Hypothese annehmen zu können. Auch dies legte nahe, an den gewählten Fehlerniveaus festzuhalten.

Unter Zugrundelegung dieser Überlegungen ließ sich der optimale Stichprobenumfang ermitteln. Dies geschah unter Zuhilfenahme des Programms G-Power 3.0 (Faul, Erdfelder, Lang & Buchner, 2007). Vor dem Hintergrund der Parameter $\eta^2=.14$, $\alpha=.05$ und $\beta=.20$ wurde für die Überprüfung der Hypothesen ein optimaler Umfang von 18 Personen pro Zelle angestrebt, so dass insgesamt Daten von 72 TeilnehmerInnen in die Auswertung einfließen sollten⁸.

I.5.1.4 Kontrolle von allgemeinen Störeffekten

Zur Kontrolle der allgemeinen Störeffekte bzw. der Situations-, VersuchsleiterInnen- und ProbandInnenmerkmale, die im Konfundierungsfall zu eingeschränkter interner Validität und damit Interpretierbarkeit der Daten führen können, wurden in der experimentalpsychologischen

⁸ Unter Zugrundelegung der angegebenen Parameter resultierte ein Umfang von 21 Personen pro Zelle für den interindividuellen Vergleich. Da die Effektgröße jedoch nur eine Schätzung darstellte und zudem 9 Materialkombinationen existierten, wurde der Umfang auf die genannte Anzahl von 18 Personen korrigiert.

Literatur mehrere Maßnahmen vorgeschlagen (z.B. Hussy & Jain, 2002). Im Rahmen der vorliegenden Untersuchung wurden Situations- und Versuchsleitermerkmale durch den Einsatz von Konstanthaltung und Elimination, ProbandInnenmerkmale durch Randomisierung kontrolliert. Dies soll im Folgenden näher erläutert werden.

Unter Situationsmerkmalen versteht man „alle Merkmale der Untersuchungssituation im raumzeitlichen Untersuchungskontext: Räumlichkeiten, Tageszeit, Lichtverhältnisse, Geräuschkulisse, Untersuchungsmaterial usw.“ (Hussy & Jain, 2002, S. 103). Angestrebt werden sollte die Realisierung möglichst vergleichbarer situationaler Bedingungen für alle TeilnehmerInnen. Daher wurden alle ProbandInnen im selben Raum unter durch Abdunklung möglichst konstant gehaltenen visuellen sowie durch Einzelsitzungen und die Anwesenheit nur der Versuchsleiterin im Raum möglichst ähnlichen auditiven Bedingungen untersucht. Auch die Präsentation des Untersuchungsmaterials war aufgrund der computergestützten Darbietung für alle ProbandInnen vergleichbar. Da ebenfalls die Instruktionen per Computer vermittelt wurden und die TeilnehmerInnen vollständig autonom arbeiteten, kann auch ausgeschlossen werden, dass VersuchsleiterInnenmerkmale die interne Validität der Untersuchung minderten. Zudem wurden alle Daten von derselben Versuchsleiterin erhoben. Um ebenfalls einen systematischen Einfluss von ProbandInnenmerkmalen auf die abhängigen Variablen auszuschließen, wurden die TeilnehmerInnen dem interindividuell variierten Faktor „Art der Verarbeitung in Phase 2“ sowie einer der 9 Materialkombinationen, deren Zustandekommen noch näher erläutert werden wird, zufällig zugeordnet. Da auch zwei intraindividuell variierte Faktoren existierten, soll im nächsten Abschnitt auf die Kontrolle spezieller Störeffekte bei Messwiederholung eingegangen werden.

An dieser Stelle bleibt zu resümieren, dass durch den Einsatz der Kontrolltechniken Elimination, Konstanthaltung und Randomisieren die interne Validität der Untersuchung sowie ihre Präzision entscheidend erhöht werden konnte (für die Präzision gilt dies allerdings nur im Hinblick auf die Elimination und Konstanthaltung).

I.5.1.5 Kontrolle von speziellen Stör- und von Materialeffekten

Spezielle Stör- bzw. sog. Sequenzeffekte können auftreten und die interne Validität der Untersuchung beeinträchtigen, wenn sich aufgrund von Messwiederholung die Position in Form von Übung, Erinnerung, Ermüdung oder Übertragung von zeitlich früheren auf spätere Bedingungen auf die abhängige Variable auswirkt. Solchen Sequenzeffekten kann jedoch immer dann entgegengewirkt werden, wenn die Reihenfolge in der Position nicht fixiert ist – und dies war

in der vorliegenden Untersuchung der Fall. Daher wird im Folgenden dafür plädiert, die intraindividuell variierten Faktoren als genuin experimentelle aufzufassen.

Da sich der zentrale Vergleich auf die Stufen 5x vs. 0x der intraindividuell variierten unabhängigen Variable „Frequenz der Verarbeitung“ bezog, waren im vorliegenden Fall besonders auch Materialeffekte zu kontrollieren. Hierdurch sollte ausgeschlossen werden, dass ein möglicherweise zu findender Unterschied zwischen beiden Frequenzen lediglich auf die unterschiedliche Beschaffenheit des Materials – seine Erinnerbarkeit, Vorstellbarkeit o.ä. – zurückgeführt werden könnte. Im Folgenden soll jedoch zunächst auf die Kontrolle möglicher Sequenzeffekte eingegangen werden.

Im Rahmen dieser Untersuchung existierten insg. zwei intraindividuell variierte Faktoren: Die Art der Enkodierung in Phase 1 (ausgeführt, gelesen, nicht enkodiert) sowie die Frequenz der Verarbeitung in Phase 2 (5x, 1x, 0x). Eine Diskussion möglicher Sequenzeffekte bzgl. des erstgenannten Faktors – der Art der Enkodierung der Items – erübrigte sich im Grunde, da sich die zu prüfenden Hypothesen gar nicht auf den Vergleich der Stufen dieser unabhängigen Variablen, sondern nur auf den Vergleich innerhalb einer Stufe bezogen (genauer gesagt auf den Mittelwert zweier Stufen, nämlich in Phase 1 gelesener und nicht enkodierter Handlungsanweisungen). Dennoch soll hier kurz dafür argumentiert werden, diesen Faktor als experimentell aufzufassen. Dabei betrifft die experimentelle Variation die Enkodierung, während die eigentliche Messung der abhängigen Variablen zwei Wochen später erfolgt. Während dieser Messung wurden alle Itemarten in zufälliger und nicht in fixierter Reihenfolge – z.B. in Form einer geblockten Darbietung der verschiedenen Itemarten – präsentiert. Da die Messung 55 Items umfasste (Näheres hierzu später), waren Ermüdungs- oder Sensibilisierungseffekte nicht auszuschließen. Da allerdings für jeden Teilnehmer eine neue, zufällige Reihenfolge der Präsentation gewählt und 5 Items jeder Itemart erfasst wurden, sollten sich derartige Effekte inter- und intraindividuell ausgleichen. Davon zu unterscheiden ist die experimentelle Variation des Faktors in der ersten Phase der Enkodierung. Um diese maximal valide und damit interpretierbar zu gestalten, wurden die Items und Itemarten auch in dieser Phase für jede Teilnehmerin in zufälliger Reihenfolge präsentiert, so dass Ermüdungseffekte ausgeglichen werden sollten. Da in der ersten Phase allerdings insg. nur 30 Handlungsanweisungen präsentiert wurden und diese Phase aufgrund der auftretenden Handlungsausführungen als relativ kurzweilig erlebt wurde, waren ohnehin keine bzw. kaum Ermüdungseffekte zu erwarten. Aufgrund der Einfachheit des Materials und der Untersuchungsanlage konnte ein Einfluss von Übungs- und Übertragungseffekten im Rahmen von Enkodierung und Abruf ausgeschlossen werden. Erinnerungseffekte und Sensibilisierung im klassischen Sinne dürften ebenfalls keine Rolle spielen.

Die zweite intraindividuell variierte Variable – die Frequenz der Präsentation – stellte allerdings den zentralen Faktor bei der Hypothesenprüfung dar und war von daher bzgl. möglicher Sequenzeffekte genauestens zu beleuchten. Aufgrund der Einfachheit der Aufgaben und der Anlage der Untersuchung dürften auch hier Übung und Sensibilisierung von nur geringer Bedeutung sein. Insbesondere aufgrund der bis zu fünfmaligen Präsentation hätte es jedoch zu Ermüdungs-, Erinnerungs- und teilweise auch zu Übertragungseffekten kommen können. Um dem entgegenzusteuern, wurden Maßnahmen ergriffen, die sicherstellen sollten, dass die präsentierten Handlungsanweisungen auch wirklich wie instruiert gehandhabt und nicht einfach aufgrund von Ermüdung ignoriert bzw. allein durch Erinnerung oder bloße Übertragung verarbeitet werden konnten (ein Beispiel hierfür ist die Steigerung des Schwierigkeitsgrades beim Generieren). Diese Maßnahmen sind im Detail dem nächsten Abschnitt zu entnehmen. Des Weiteren wurden die Handlungsanweisungen wiederum in (quasi-)zufälliger Reihenfolge präsentiert, um sicherzustellen, dass keine geblockten Präsentationen vorkommen und sich systematisch auswirken könnten. Auch diese Maßnahmen sind im Abschnitt I.5.2.4 näher erläutert. An dieser Stelle soll nur darauf hingewiesen werden, dass die beiden intraindividuell variierten Variablen mit großer Sorgfalt eingeführt wurden, um eine Minderung der internen Validität zu vermeiden. Dazu gehört insbesondere auch, dass Materialeffekte durch Ausbalancieren verhindert wurden. Zu diesem Zwecke wurden Materialpools gebildet, um (der Logik des lateinischen Quadrats folgend) sicherzustellen, dass alle Items gleich häufig in jeder Stufe der intraindividuell variierten Faktoren vorkamen (Näheres hierzu siehe Abschnitt I.5.2.2). Die resultierenden neun Materialkombinationen wurden fortlaufend eingesetzt, so dass von einer zufälligen Zuordnung gesprochen werden kann. Neben der Sicherung der internen Validität trug diese Maßnahme auch zur Generalisierbarkeit über das Itemmaterial (externe Validität) bei. Abschließend bleibt zu erwähnen, dass der Einfluss von Positions-, Übertragungs- und Materialeffekten aufgrund des zweiwöchigen Abstandes zwischen Enkodierung/ Verarbeitung und Abfrage ohnehin deutlich abgeschwächt sein sollte. Gleichzeitig können aber Einflüsse, die nicht auf die eigentliche experimentelle Manipulation zurückgehen, sog. zwischenzeitliches Geschehen (Hussy & Jain, 2002), nicht ausgeschlossen werden. Allerdings sollte dies alle TeilnehmerInnen in gleichem Maße betreffen und daher aufgrund der randomisierten Zuordnung der TeilnehmerInnen zu den experimentellen Bedingungen keinen systematischen Einfluss auf die Ergebnisse ausüben.

Insgesamt wird hier also die Ansicht vertreten, dass beide intraindividuell variierten Variablen als genuin experimentelle Faktoren aufzufassen sind, deren Auswirkungen auf die abhängige Variable uneingeschränkt kausal interpretiert werden können.

I.5.1.6 Ethische Gesichtspunkte

Im Rahmen der Untersuchungsplanung sollten ethische Gesichtspunkte stets Berücksichtigung finden, um negative Konsequenzen für TeilnehmerInnen wissenschaftlicher Untersuchungen zu vermeiden (z.B. Bortz & Döring, 2002; Hussy & Jain, 2002).

Ethisch problematisch könnte sich im Rahmen der vorliegenden Studie bereits das Untersuchungsziel darstellen: Durch experimentelle Manipulationen sollen Gedächtnistäuschungen, falsche Erinnerungen, hervorgerufen werden. Die Verfolgung dieses Ziels wird jedoch als legitim und verantwortbar angesehen, da Gegenstand der Gedächtnistäuschungen einfache und alltägliche Handlungen ohne persönliche Relevanz sind, die keinerlei negative Konsequenzen für die TeilnehmerInnen nach sich ziehen dürften. Problematischer stellt sich jedoch die Strategie dar, die TeilnehmerInnen nicht schon zu Beginn der Untersuchung über den eigentlichen Gegenstand und den konkreten Ablauf selbiger zu informieren. Durch eine sehr schwammige Darstellung des Untersuchungsthemas sollte eine voreingenommene und selektive Enkodierung vermieden werden. D.h., dass zwar kein Gebrauch von einer sog. coverstory gemacht bzw. die TeilnehmerInnen über den eigentlichen Sinn und Zweck der Untersuchung getäuscht wurden, dieser jedoch auch nicht explizit genannt wurde. Diese mangelnde Transparenz war aber Gegenstand eines Vertrages, der mit allen TeilnehmerInnen zu Beginn der Untersuchung geschlossen wurde (siehe Anhang A.1). Der Vertrag gab Aufschluss über die Rechte der TeilnehmerInnen, insb. das Abbruchsrecht, das aufgrund der mangelnden Aufklärung im Vorhinein besonderes Gewicht erhielt, aber auch über ihre Pflichten, z.B. den Instruktionen so gut wie möglich nachzukommen. Eine Teilnahme war nur möglich, wenn die TeilnehmerInnen sich mit diesen Bedingungen einverstanden erklärten. Die Untersuchung selbst, insb. die zweite Phase, erforderte von den TeilnehmerInnen ein hohes Maß an Ausdauer. Diese Phase mit ihren zahlreichen, wiederholten Präsentationen wurde zwar als sehr anstrengend, aber machbar empfunden. Im Rahmen der Voruntersuchung wurden TeilnehmerInnen hierzu und auch zur angesprochenen Intransparenz befragt, empfanden beides jedoch als zumutbar. Nach Abschluss der Studie wurden die TeilnehmerInnen detailliert über den Hintergrund der Untersuchung informiert. Auch wird eine Zusammenfassung der Dissertationsdaten denjenigen TeilnehmerInnen zugehen, die diesen Wunsch äußerten.

I.5.2 Versuchsdurchführung

I.5.2.1 Stichprobe

Die TeilnehmerInnen wurden über die üblichen Kanäle (Aushänge, Mailinglisten usw.) auf die Untersuchung aufmerksam gemacht. An einer Voruntersuchung, auf die im Folgenden kurz eingegangen werden wird, nahmen 18 an der Universität zu Köln eingeschriebene Psychologiestudierende im Grundstudium teil, die damit einen Teil der curricular vorgeschriebenen Versuchspersonenstunden ableisteten. 17 Teilnehmerinnen waren weiblich, ein Teilnehmer männlich. Das Alter reichte von 19 bis 51 Jahren ($M=28.83$; $SD=10.10$).

Nachdem vor dem Hintergrund der Voruntersuchung einige kleinere Änderungen vorgenommen worden waren, konnten im Rahmen der Hauptuntersuchung Daten von 66 Psychologiestudierenden im Grundstudium gesammelt werden. Von diesen nahmen zwei den Termin zur zweiten Sitzung nicht wahr; ein Mal verhinderten technische Gründe die Datenerhebung. Drei TeilnehmerInnen schienen die Instruktionen in der ersten Phase teilweise missverstanden zu haben bzw. bemühten sich nicht, ernsthaft mitzuarbeiten, so dass ihre Daten aus der zweiten Sitzung unbesehen von der Analyse ausgeschlossen wurden.

Dies resultierte in einer Stichprobe von 60 TeilnehmerInnen – 15 pro Gruppe – mit den folgenden Charakteristika: Das Alter umfasste eine Spannweite von 19 bis 49 Jahren mit einem Mittelwert von 23.07 Jahren ($SD=5.40$), wobei 49 Personen weiblichen und 11 männlichen Geschlechts waren. Die Verteilung der Merkmale Alter und Geschlecht war zwischen den Gruppen vergleichbar.

I.5.2.2 Material

Grundlegend für die Untersuchung war die Zusammenstellung eines Pools von Handlungsanweisungen. Dabei sollten diese drei Bedingungen genügen: Sie sollten alltäglich sein, sie sollten den Gebrauch eines Objektes einschließen und sie sollten allein mit den Händen auszuführen sein.

Es lassen sich mehrere Gründe für die erste Bedingung anführen. Einerseits haben Thomas und Loftus (2002) gezeigt, dass der imagination-inflation-Effekt in stärkerem Maße auftritt, wenn alltägliche anstelle von ungewöhnlichen Handlungen genutzt werden; somit beeinflusst die Wahl des Materials die Präzision der Hypothesenprüfung. Zweitens sollte theoretisch das Beobachten einer bekannten, also im Motorrepertoire bereits vorhandenen, Handlung zu einer

stärkeren Simulation führen als das Beobachten einer unbekannteren Handlung (vgl. Abschnitt I.3.2.4). Drittens war von Anfang an anvisiert, das Design auch zur Untersuchung von älteren gesunden (und evtl. auch dementen) TeilnehmerInnen einzusetzen. Es wurde jedoch angenommen, dass es gerade diesen Gruppen schwerlich nahe gebracht bzw. von diesen auch als unangenehm empfunden werden könnte, ungewöhnliche Handlungen mit den Materialien auszuführen (wie „Balancieren Sie das Lineal auf dem Kopf!“). Im Gegenteil interessierte in diesem Falle ja auch gerade die Erinnerung an alltägliche Handlungen (siehe Abschnitt II), gegen deren Einsatz man jedoch ins Feld führen könnte, dass es beim Abruf leicht zu Verwechslungen mit anderen, außerhalb des Experiments liegenden, Kontexten kommen könnte. So könnte ein Teilnehmer bspw. erinnern, eine Karte in den Umschlag gesteckt zu haben – dabei aber nicht berücksichtigen, dass er die Karte am Vortag zu Hause in den Umschlag steckte und nicht während der Untersuchung. Mit anderen Worten könnte eine weitere Quelle der Fehlattribution eingeführt werden. Allerdings ist hier keine systematische Variation mit den Frequenzstufen oder den verschiedenen Arten der Verarbeitung zu erwarten. Sollten sich Hinweise auf ein derartiges Problem in den Daten finden, wäre dies für die darauf folgenden Untersuchungen noch einmal zu überdenken. Insgesamt wurde daher zunächst an der Wahl von alltäglichen Handlungen festgehalten, insb. um dem Auffinden eines observation-inflation-Effekts eine möglichst große Chance einzuräumen. Die zweite und dritte Bedingung – die Handlungen sollten mit den Händen auszuführen sein und den Gebrauch eines Objektes einschließen – sollten v.a. die Vergleichbarkeit der Handlungsanweisungen sicherstellen. Der Einschluss von Objekten diente dabei der sensorischen Anreicherung der Handlungen. Darüber hinaus wurde anvisiert, die Videoaufzeichnungen auf einen sehr selektiven Ausschnitt zu begrenzen (es sollten nur der Tisch, ein Teil des Torsos und die Arme bzw. Hände der AkteurInnen sichtbar sein), was ebenfalls das Spektrum der Handlungen auf solche, die mit den Händen auszuführen sind, restringierte.

Als angemessene Masse an Handlungsanweisungen pro Zelle, die nicht nur eine reliable Messung, sondern auch einen zumutbaren Umfang gewährleisten sollte, wurde eine Anzahl von fünf angesehen (siehe auch Thomas & Bulevich, 2006). Dies bedeutet, dass in der ersten Phase insg. 30 Handlungsanweisungen präsentiert wurden, von denen 15 auszuführen und 15 zu lesen waren; 15 Handlungsanweisungen wurden in dieser Phase gar nicht präsentiert bzw. enkodiert, kamen jedoch in der zweiten Phase zum Einsatz. Jede Handlungsanweisung erschien für 15 Sekunden auf dem Bildschirm – eine Dauer, die schon in den Vorgängerstudien eingesetzt worden war (Goff & Roediger, 1998; Thomas et al., 2002, 2003) und sich in der Voruntersuchung als ausreichend erwiesen hatte. Die Präsentation erfolgte randomisiert.

Es wurde somit ein Gesamt von 45 Handlungsanweisungen benötigt (plus 10 Distraktoren, die lediglich bei der Abfrage zum Einsatz kommen sollten). Da der vorhandenen Literatur nicht genügend Handlungsanweisungen entnommen werden konnten, die alle drei aufgestellten Bedingungen erfüllten, wurden zusätzliche Handlungsanweisungen analog formuliert.

Es war aus technischen Gründen nicht möglich, für jede Teilnehmerin zufällig zu bestimmen, welche Handlungen ausgeführt werden sollten und welche nicht – denn dies hätte wiederum bei der Präsentation in der zweiten Phase berücksichtigt werden müssen. Außerdem erschien dieses Vorgehen aufgrund der kleinen Stichprobengröße und der großen Anzahl an Handlungsanweisungen nicht sinnvoll. Daher wurde eine Poolbildung angestrebt. Um die Handlungsanweisungen für die beiden dreistufigen, intraindividuell variierten Faktoren „Art der Enkodierung der Handlungsanweisungen in Phase 1“ und „Frequenz der Präsentation der Handlungsanweisungen in Phase 2“ auszubalancieren, war die Bildung von drei Pools mit jeweils drei Unterpools notwendig. Dies bedeutete, dass die 45 Handlungsanweisungen in drei Pools à 15 Items und dann in Subpools à 5 Items aufgeteilt werden mussten. Hierfür wurden die Anweisungen zunächst in semantisch verwandte Gruppen eingeteilt – so gab es bspw. Objekte, die man eher dem Bereich Lebensmittel, andere die man dem Bereich Büroartikel oder Drogerieartikel zuordnen konnte usw. Um sicherzustellen, dass beim Abruf nicht von der Zugehörigkeit eines Gegenstandes bzw. einer Handlung zu einer bestimmten inhaltlichen Kategorie auf die Art der Verarbeitung geschlussfolgert werden könnte („Bei der Präsentation waren doch gar keine Drogerieartikel dabei – dann kann ich die Handlung ja nicht ausgeführt haben.“), wurden diese semantischen Gruppen gleichmäßig und zufällig auf die Pools verteilt.

Dies resultierte in der in der folgenden Tabelle dargestellten Auflistung.

Tab. 8: Materialpool des ersten Experiments.

Pool	Nr.	Handlungsanweisung	Darst. (Video)
Pool 1			
A	1	Würfeln Sie mit den Würfeln!	m
	2	Schütteln Sie die Flasche!	m
	3	Wickeln Sie das Band auf!	m
	4	Falten Sie das Blatt Papier!	w
	5	Betätigen Sie die Fernbedienung!	w
B	6	Mischen Sie die Karten!	w
	7	Tun Sie das Ei in den Eierbecher!	w
	8	Ziehen Sie die Socke auf links!	m
	9	Spitzen Sie den Bleistift an!	w
C	10	Reißen Sie einen Zettel von dem Block ab!	m
	11	Legen Sie die Seife in die Dose!	m
	12	Öffnen Sie die Sicherheitsnadel!	w
	13	Lochen Sie das Papier!	m
	14	Biegen Sie eine Büroklammer auseinander!	w
	15	Nehmen Sie ein Streichholz aus der Schachtel!	m
Pool 2			
A	16	Falten Sie das Handtuch!	w
	17	Dehnen Sie das Gummiband!	w
	18	Zerschneiden Sie das Papier!	m
	19	Schalten Sie die Taschenlampe an!	w
	20	Nehmen Sie den Deckel von der Schachtel!	m
B	21	Betätigen Sie die Klingel!	m
	22	Wickeln Sie ein Bonbon aus!	w
	23	Stellen Sie das Teelicht in das Glas!	m
	24	Nehmen Sie ein Stück Tesaband von der Rolle!	m
	25	Schlagen Sie das Buch auf!	w
C	26	Stoßen Sie das Spielzeugauto an!	w
	27	Tun Sie den Teebeutel in die Tasse!	m
	28	Stellen Sie die Uhr eine Stunde vor!	w
	29	Nehmen Sie eine Tintenpatrone aus der Schachtel!	m
	30	Klappen Sie die Sonnenbrille zusammen!	m
Pool 3			
A	31	Stempeln Sie auf das Papier!	m
	32	Legen Sie die Nudeln in die Tasche!	m
	33	Betätigen Sie das Feuerzeug!	w
	34	Öffnen Sie das Mäppchen!	w
	35	Nehmen Sie ein Taschentuch heraus!	m
B	36	Schrauben Sie den Deckel vom Deo!	w
	37	Schneiden Sie ein Stück Garn von der Rolle ab!	w
	38	Zerreißen Sie das Papier!	w
	39	Nehmen Sie die CD aus der Hülle!	m
	40	Schließen Sie das Schloss auf!	m
C	41	Werfen Sie den Ball von der einen in die andere Hand!	w
	42	Nehmen Sie einen Becher von dem Stapel!	w
	43	Entfernen Sie die Batterie aus dem Wecker!	m
	44	Schrauben Sie den Kugelschreiber auf!	w
	45	Stecken Sie die Karte in den Umschlag!	m
Pool 4 (Distraktoren)			
A	46	Werfen Sie die Münze!	
	47	Unterschreiben Sie auf dem Papier!	
	48	Stechen Sie eine Nadel in das Nadelkissen!	
	49	Nehmen Sie die Kappe von dem Textmarker!	
	50	Schließen Sie den Reißverschluss!	
B	51	Entfernen Sie den Anhänger vom Schlüssel!	
	52	Machen Sie eine Schleife in das Band!	
	53	Nehmen Sie die Zahnbürste aus dem Becher!	
	54	Betätigen Sie den Taschenrechner!	
	55	Klappen Sie das Taschenmesser auseinander!	

Wollte man nun alle möglichen Kombinationen der drei Ober- (1,2,3) und Unterpools (A,B,C) realisieren, würden am Ende 36 Kombinationsmöglichkeiten resultieren – das Doppelte der benötigten Stichprobengröße pro Gruppe. Daher wurde entschieden, das Material nach der Methode des lateinischen Quadrates auszubalancieren, so dass sich pro Faktor insg. drei mögliche und für die Kombination beider Faktoren neun Materialkombinationen ergaben:

Tab. 9: Materialkombinationen des ersten Experiments.

		Art der Enkodierung in Phase 1: ausgeführt, gelesen, nicht enkodiert		
		Pools 1,2,3	Pools 3,1,2	Pools 2,3,1
Frequenz der Präsentation in Phase 2: 5x, 1x, 0x	Pools A,B,C	K1	K2	K3
	Pools C,A,B	K4	K5	K6
	Pools B,C,A	K7	K8	K9

Da die Handlungsanweisungen alle gleichermaßen alltäglich und einfach erschienen und den Pools zufällig zugeteilt wurden, schien es auch nicht notwendig, alle 36 Kombinationsmöglichkeiten zu realisieren – es war schlicht nicht zu erwarten, dass kombinationsspezifische Materialeffekte auftreten würden. Vielmehr ging es darum, sicherzustellen, dass nicht nur eine einzige Materialzusammenstellung genutzt würde, die zwar für den interindividuellen Vergleich aufschlussreich gewesen wäre, beim intraindividuellen Vergleich jedoch stets hätte zweifelhaft erscheinen lassen, ob der Effekt nun auf das Material oder die experimentelle Manipulation zurückzuführen sei. Die Möglichkeit eines Materialeffekts ließ sich durch das genannte Vorgehen ausschließen, da jede Handlungsanweisung in allen Stufen der intraindividuell variierten Faktoren einmal vorkam. Die Distraktoren kamen stets nur bei der Abfrage zum Einsatz, wurden also nicht ausbalanciert. Sie dienten primär dem Zweck, möglichen Antworttendenzen dadurch entgegenzuwirken, dass bei der Abfrage etwa gleich viele Items jeder Itemart präsentiert werden konnten.

Die zur Ausführung der Handlungsanweisungen notwendigen Objekte wurden so gewählt, dass sie möglichst eindeutig identifizierbar waren (vgl. Anhang B.3, in dem beispielhaft die Objekte des zweiten Experiments abgebildet sind). Dies wurde durch die TeilnehmerInnen der Voruntersuchung bestätigt. Somit existierte ein Fundus von Handlungsanweisungen und zugehörigen Objekten, der die erste Phase der Untersuchung, in der Handlungsanweisungen entweder nur gelesen oder auch ausgeführt werden sollten, bereits realisierbar machte. Ebenso

waren für die beiden Bedingungen des Vorstellens und Lesens in Phase 2 wie auch für den Abruf in Phase 3 mit den Handlungsanweisungen bereits alle notwendigen Materialien vorhanden, da diese keiner zusätzlichen Bearbeitung oder Ergänzung bedurften. Für die Bedingungen „Beobachten“ und „Generieren“ mussten jedoch weitere Materialien konstruiert werden.

Wie bereits erläutert, wurde als Form des Generierens das Lösen umgestellter Sätze bzw. von Satzanagrammen gewählt, da hierbei der Schwierigkeitsgrad gesteigert werden konnte. Da keine Vorlagen in der Literatur gefunden werden konnten, wurden eigene Regeln aufgestellt, um die Handlungsanweisungen in die gewünschte Form zu bringen. Maßgebend sollte dabei sein, dass diese Vorlagen so schwierig sein sollten, dass eine Lösung nicht auf den ersten Blick gefunden werden konnte, aber dennoch so einfach, dass die richtige Lösung in sehr kurzer Zeit mit an Sicherheit grenzender Wahrscheinlichkeit erfolgen konnte. Die im Rahmen der Voruntersuchung eingesetzten umgestellten Sätze stellten sich dabei als allzu einfach und durchsichtig heraus. So wurde die Schwierigkeit der „Satz-Anagramme“ auf einer Skala von 1 bis 7 im Mittel über alle Anagramme und alle TeilnehmerInnen hinweg mit 1.45 ($SD=.32$) eingeschätzt. Daher wurden die Regeln im Nachgang zur Voruntersuchung noch einmal überarbeitet und die Vorlagen erschwert, was letztlich in folgendem Regelwerk resultierte:

Tab. 10: Regeln zur Erstellung umgestellter Sätze des ersten Experiments.

Schwierigkeitsstufe	Definition
	Beispiel: Schütteln Sie die Flasche!
1	Eine Vertauschung von zwei Wörtern, die direkt nebeneinander stehen; Zahlen anstatt Leerzeichen zwischen den Wörtern. Bsp.: Sie3Schütteln9die5Flasche!
2	Zwei Vertauschungen von zwei Wörtern, die nicht direkt nebeneinander stehen (bei nur 4 Wörtern auch direkt nebeneinander stehend); ein Konsonant fehlend in zwei Wörtern; keine Leerzeichen zwischen den Wörtern. Bsp.: Flas_hedieSieSchüt_eln!
3	Eine Vertauschung von zwei Wörtern, die nicht direkt nebeneinander stehen; ein Vokal fehlend in zwei Wörtern; Buchstaben anstatt Leerzeichen zwischen den Wörtern. Bsp.: d_eMSieWSch_ttelnAFlasche!
4	Zwei Vertauschungen von zwei Wörtern, die nicht direkt nebeneinander stehen (bei nur 4 Wörtern auch direkt nebeneinander stehend); ein Vokal fehlend in drei Wörtern; Buchstaben und Zahlen anstatt Leerzeichen zwischen den Wörtern. Bsp.: S_eMSchütt_In7Flasch_Gdie!
5	Zwei Vertauschungen von zwei Wörtern, die nicht direkt nebeneinander stehen (bei nur 4 Wörtern auch direkt nebeneinander stehend); ein Vokal fehlend in zwei Wörtern, zwei Vokale fehlend in zwei Wörtern; Kleinschreibung . Bsp.: d__fl_sche sch_tt_In s_e!

Diese neu konstruierten Anagramme (vgl. Anhang A.2) wurden im Rahmen der Hauptuntersuchung mit einem Mittelwert von 2.76 ($SD=.95$) dann auch insgesamt als deutlich schwieriger – wenn auch (wie gewollt) als gut lösbar – eingeschätzt.

Für die Bedingung „Beobachten“ wurde die Handlungsausführung aufgezeichnet. Die Maßgaben sind dabei schon mehrfach genannt worden: Es sollten auf dem Ausschnitt nur der Tisch, das Objekt/ die Objekte, die Hände und Teile der Arme bzw. des Torsos der AkteurInnen zu sehen sein, um eine volle Konzentration auf die Handlungsausführung sowie keine Ablenkung durch etwaige Überlegungen zur Person zu gewährleisten. Dabei wurde aus einer für die Beobachtung gewöhnlichen Perspektive gefilmt, d.h. so, als säßen die AkteurInnen den BeobachterInnen gegenüber (2.-Person-Perspektive; vgl. Abschnitt I.7). Es kamen sowohl eine weibliche Akteurin als auch ein männlicher Akteur zum Einsatz (siehe Abbildung 6); das Geschlecht der DarstellerInnen wurde innerhalb der Pools und Subpools ausbalanciert, die Verteilung ist der Tabelle 8 zu entnehmen. Aus den Videoaufnahmen wurde pro Handlungsanweisung eine Sequenz geschnitten, die zeigte, wie die Handlung (entweder von der Akteurin oder von dem Akteur) einmal ausgeführt wurde (z.B. einmaliges Anschalten der Taschenlampe). Dabei konnte die Dauer der Handlungsausführungen aufgrund der unterschiedlichen Anforderungen nicht für alle Handlungen konstant gehalten werden. Dies wurde allerdings als unproblematisch erachtet, da auch für die anderen Formen der Verarbeitung keine Konstanz erzielt werden konnte (so kann z.B. auch das Vorstellen einiger Handlungen länger dauern als das anderer bzw. hängt das Lesen von der Länge der Handlungsanweisungen ab usw.). Jede dieser einzelnen Sequenzen wurde anschließend vervielfacht und so zu einer Schleife zusammengesetzt, dass sich exakt dieselbe Sequenz für ein Intervall von mind. 15 Sekunden immer wieder wiederholte (während der Untersuchungsdurchführung wurden die Videodateien nach 15 Sekunden Abspielzeit durch das Programm gestoppt). So zeigte bspw. eine Sequenz wiederholt die immer gleiche Szene des Anschaltens einer Taschenlampe.



Abb. 6: Ausschnitte aus den Videos zum ersten Experiment. Die Handlungsausführung erfolgte durch eine Akteurin (links) bzw. einen Akteur (rechts).

Jeweils ein Drittel der Handlungsanweisungen, die in der Enkodierphase ausgeführt, gelesen oder gar nicht präsentiert worden waren, sollte in der zweiten Phase 5x, 1x bzw. 0x dargeboten werden. Demnach sollten in der zweiten Phase 5 der 15 in Phase 1 ausgeführten, 5 der 15 in Phase 1 gelesenen und 5 der 15 in Phase 1 nicht enkodierten Handlungsanweisungen 5x dargeboten, weitere 5 Handlungsanweisungen jeder Itemart 1x und die restlichen 5 gar nicht mehr präsentiert werden (0x). Somit ergab sich eine Anzahl von insgesamt $(15 \times 5) + (15 \times 1) + (15 \times 0) = 90$ Präsentationen in der zweiten Phase. Das Vorgehen lässt sich anhand der folgenden Darstellung nachvollziehen. Abgebildet sind nur die intraindividuell variierten unabhängigen und die abhängigen Variablen.

Tab. 11: Darbietung der intraindividuell variierten unabhängigen und der abhängigen Variablen des ersten Experiments (am Beispiel der Materialkombination 1).

Phase 1: Enkodierung	UV-A	Art der Enkodierung			
	Stufe d. UV-A	ausgeführt	gelesen	nicht enkodiert	
	Anzahl	15	15	(15)	Σ 30
	Pools	1A,1B,1C	2A,2B,2C	3A,3B,3C	
	Dauer	15 Sekunden/Handlungsanweisung			
	Reihenfolge	vollständig randomisiert			
Phase 2: Verarbeitung*	UV-B	Frequenz der Verarbeitung			
	Stufe d. UV-B	5x	1x	0x	
	Anzahl	15	15	(15)	Σ 90
	Pools	1A,2A,3A	1B,2B,3B	1C,2C,3C	
	Dauer	15 Sekunden/Handlungsanweisung			
	Reihenfolge	fixiert randomisiert			
Phase 3: Abruf**	AVn	Rekognitions-/ Quellengedächtnis-/ Sicherheitsurteile			
	Anzahl	45 + 10 Distraktoren			Σ 55
	Pools	1A,1B,1C,2A,2B,2C,3A,3B,3C,4A,4B			
	Dauer	nicht vorgegeben			
	Reihenfolge	vollständig randomisiert			

Anmerkungen: * 5 Minuten nach Phase 1; ** 2 Wochen nach Phase 1.

Um sicherzustellen, dass im Rahmen der zweiten Phase nicht alle mehrmaligen Präsentationen zu Beginn und die einmaligen gegen Ende erfolgten, wenn die Konzentration schon deutlich nachgelassen hätte, wurde folgende Regel aufgestellt: Auf jeweils 5 Items, die fünfmal dargeboten werden, folgt ein Item, das einmal präsentiert wird. Die kontinuierliche Unterbrechung der wiederholt präsentierten durch die nur einmalig präsentierten Handlungsanweisungen sollte gewährleisten, dass die ProbandInnen aufmerksam blieben. Zusätzlich wurde eine weitere Regel aufgestellt: Die Reihenfolge der Präsentation erfolgt innerhalb des genannten

Schemas zufällig, wobei fünf Zufallsreihenfolgen der 15 fünfmalig präsentierten Items gebildet werden sollten, die dann nacheinander zum Einsatz gelangen sollten (alternativ hätte auch für das Gesamt der 15 fünfmalig präsentierten Items (also für alle 75 Präsentationen) eine Zufallsreihenfolge gebildet werden können; hierbei hätten aber (zufällige) Blockpräsentationen gleicher, wiederholt präsentierter Items nicht ausgeschlossen werden können). Da sich die an solche Regeln gebundenen Zufallsreihenfolgen aus technischen Gründen nicht für jede Teilnehmerin eigens generieren ließen, wurden sie zuvor fixiert. Konkret existierten also für alle neun Materialkombinationen unterschiedliche fixierte Zufallsreihenfolgen für Phase 2. Für Materialkombination 1 sah diese bspw. folgendermaßen aus (die fett markierten Items stellen die wiederholt präsentierten dar, jede Zeile entspricht der gesamten Liste der 15 fünfmal präsentierten Items):

Tab. 12: Beispiel einer fixierten Zufallsreihenfolge des ersten Experiments (Materialkombination 1).

Frequenz	Beispiel: Pools der Materialkombination 1																	
5x	Pool 1A (01,02,03,04,05), Pool 2A (16,17,18,19,20), Pool 3A (31,32,33,34,35)																	
1x	Pool 1B (06,07,08,09,10), Pool 2B (21,22,23,24,25), Pool 3B (36,37,38,39,40)																	
0x	Pool 1C (11,12,13,14,15), Pool 2C (26,27,28,29,30), Pool 3C (41,42,43,44,45)																	
Beispiel einer fixierten Zufallsreihenfolge für Materialkombination 1																		
	31	03	20	02	32	25	35	17	18	34	19	10	05	16	04	33	01	07
	34	02	05	04	17	22	31	18	32	03	20	40	16	35	33	01	19	37
	35	32	31	02	03	08	04	16	19	34	17	24	18	20	01	33	05	23
	17	05	32	33	18	39	19	16	34	35	20	06	02	31	01	04	03	09
	20	35	32	34	05	36	17	02	33	01	16	21	03	18	04	19	31	38

Als Fillertask zwischen erster und zweiter Phase kam der Need-for-Cognition/ Faith-in-Intuition- (NFC/FII-) Fragebogen (Keller, Bohner & Erb, 2000) zum Einsatz. Dieser misst die Tendenz zu einer analytisch-rationalen und aufwändigen (need for cognition) respektive intuitiv-experientiellen und sparsamen (faith in intuition) Informationsverarbeitung (Epstein, Pacini, Denes-Ray & Heider, 1996). Für beide Modi wurden in der Vergangenheit Zusammenhänge zu Gedächtnisleistungen berichtet (vgl. z.B. Whittlesea & Price, 2001). Explorativ soll daher überprüft werden, ob sich Zusammenhänge zwischen der Kognitions-/ Intuitionsneigung und der Gedächtnisleistung finden lassen.

Des Weiteren wurden beim Abruf – im Anschluss an den Gedächtnistest – ebenfalls explorativ einige Items aus dem Vividness of Visual Imagery Questionnaire (VIVIQ; Marks, 1973; die Items wurden frei ins Deutsche übersetzt und z.T. modifiziert) und aus der Dissociative Experiences Scale (DES; Bernstein & Putnam, 1986; die Items wurden der deutschen Fassung (FDS) entnommen und entstammten den Skalen Absorption und Amnesie; Spitzer, Stieglitz &

Freyberger, 2005) aufgenommen. Beide wurden bereits in verschiedenen Studien eingesetzt – allerdings zeigten sich keine konsistenten Zusammenhänge (z.B. Heaps & Nash, 1999; Mazzoni & Memmon, 2003). Zusätzlich wurden einige Items des Hamburger Zwangsinventars (HZI; Zaworka, Hand, Jauernig & Lünenschloß, 1983) modifiziert und in die Untersuchung einbezogen. Insgesamt wurden zunächst 12 VIVIQ-Items, dann 4 HZI-Items und schließlich 9 FDS-Items zur Einschätzung auf einer siebenstufigen Ratingskala vorgegeben.

I.5.2.3 Hilfsmittel und Geräte

Die Untersuchung lief computergestützt ab. Dabei kam ein eMAC (1 GHz PowerPC mit G4 Prozessor und 17" Monitor) zum Einsatz. Die Bildschirmauflösung (1024x768), Helligkeit und Kontrast wurden konstant gehalten.

Als Experimentiersoftware wurde Variotest (Mutz, 2002) eingesetzt. Hierbei handelt es sich um ein Programm, das speziell zum Zwecke der Konstruktion psychologischer Experimente und Testverfahren entwickelt wurde und mit einer eigenen Programmiersprache (TCL, Test Construction Language) versehen ist. Das Programm ermöglichte es, alle Phasen einer Sitzung unterbrechungslos aufeinanderfolgend zu präsentieren, so dass die Intervention der Versuchsleiterin nur zu Beginn der Untersuchung notwendig war. Beim Starten des Programms werden soziodemographische Daten sowie ein Code für jeden Teilnehmer eingegeben.

Die Handlungsausführung wurde per Camcorder aufgenommen, das Band in den Computer eingelesen und dann mit entsprechender Filmbearbeitungssoftware (v.a. Ulead VideoStudio 10/11, Corel) in einzelne Sequenzen geschnitten, die anschließend vervielfacht und zu einer mindestens fünfzehnsekündigen Schleife zusammengesetzt wurden.

I.5.2.4 Ablauf

Die Untersuchung lief vollständig computergestützt ab; neben dem experimentellen Paradigma wurden also auch die Fragebögen bzw. Ratings per Computer erhoben. Sie firmierte unter der vagen Bezeichnung „Studie zu mentalen Repräsentationen alltäglicher Handlungen“. Den TeilnehmerInnen wurde nach der Begrüßung dann auch nur mitgeteilt, dass sie im Folgenden – zwecks Erfassung verschiedener Formen mentaler Repräsentationen – gebeten werden würden, auf unterschiedliche Art und Weise mit Anweisungen zu alltäglichen Handlungen umzugehen, die auf dem Computermonitor präsentiert werden würden. So könnten sie bspw. aufgefordert werden, einige Handlungsanweisungen tatsächlich auszuführen und andere nur zu lesen. Anschließend wurde mit ihnen ein Vertrag abgeschlossen, in dem sie über ihre Rechte

und Pflichten als VersuchsteilnehmerInnen aufgeklärt wurden (vgl. Abschnitt I.5.1.6). Dann wurde das Computerprogramm gestartet.

Anhand einer am Computer dargebotenen Instruktion (siehe Anhang A.3) wurde den TeilnehmerInnen mitgeteilt, dass ihnen in der ersten Untersuchungsphase einfache Handlungsanweisungen wie „Zeichnen Sie einen Kreis auf das Papier!“ präsentiert werden würden, die entweder zu lesen oder auszuführen seien. Der Ablauf wurde dann sehr detailliert beschrieben, um sicherzustellen, dass die TeilnehmerInnen – die autonom arbeiten sollten – ganz genau wussten, wie sie vorgehen sollten:

Der Ablauf wird konkret so aussehen, dass Ihnen zuerst mitgeteilt wird, welche Utensilien im Zusammenhang mit einer folgenden Handlungsanweisung gebraucht werden könnten. Die erste Aufforderung könnte also z.B. lauten: `Bitte holen: Stift und Blatt Papier`. Sie würden sich dann beide Utensilien heraussuchen und vor sich auf den Tisch legen. Sobald Sie dies getan haben, drücken Sie den Start-Button. Es folgt die Handlungsanweisung zusammen mit dem Hinweis, ob Sie diese nun ausführen oder nur lesen sollen. Ganz wichtig für die Durchführung ist, dass Sie sich bei beiden Vorgehensweisen (Lesen bzw. Ausführen) gleich lang mit der Handlungsanweisung beschäftigen. Um dies sicherzustellen, ist für jede Handlungsanweisung die gleiche Zeit vorgesehen, die nach Erscheinen der Handlungsanweisung abläuft. Ihnen steht so viel Zeit zu Verfügung, dass Sie die Handlungsanweisungen mehrmals lesen bzw. ausführen können. Bitte wiederholen Sie das Lesen bzw. Ausführen einfach so lange, bis ein Endsignal ertönt und die Handlungsanweisung ausgeblendet wird.

Dies wurde im Folgenden anhand eines Beispiels veranschaulicht:

Noch einmal ganz konkret: Hätten Sie sich z.B. einen Stift und ein Blatt Papier holen sollen und läge Ihnen nun die Handlungsanweisung 'Zeichnen Sie einen Kreis auf das Papier!' zusammen mit der Aufforderung, diese zu lesen, vor, würden Sie Stift und Papier einfach nur vor sich liegen lassen, d.h. die Handlung NICHT ausführen, sondern die Anweisung so lange lesen, bis das abschließende Signal ertönt und die Handlungsanweisung ausgeblendet wird.

Läge Ihnen dieselbe Handlungsanweisung mit der Bitte vor, diese auszuführen, würden Sie so lange Kreise auf das Papier zeichnen, bis das Endsignal ertönt ist. Bitte legen Sie die Utensilien jeweils wieder an ihren Platz zurück! So geht es dann immer weiter, bis die nächste Untersuchungsphase startet. [...] Bei den meisten Handlungsanweisungen ist es im Übrigen möglich, sie auf verschiedene Arten und Weisen auszuführen bzw. zu wiederholen (z.B. jeweils ein neues Blatt Papier nehmen, jeweils einen neuen Kreis auf das alte Blatt Papier zeichnen oder denselben Kreis immer wieder nachzeichnen). Die genaue Umsetzung liegt aber ganz bei Ihnen. Ich möchte Sie nur bitten, nicht allzu verschwenderisch mit den Materialien umzugehen.

Des Weiteren wurden die TeilnehmerInnen darüber informiert, dass sie – zum Zwecke der Kontrolle – nach dem Ausblenden jeder Handlungsanweisung gebeten werden würden, ehrlich anzugeben, ob sie die jeweilige Anweisung gelesen oder ausgeführt hätten, und wie häufig sie dies vollzogen hätten. Abschließend wurde darauf hingewiesen, dass sie bei Unklarheiten unbedingt die Versuchsleiterin kontaktieren sollten.

Im Rahmen dieser ersten Phase wurden insgesamt 30 Handlungsanweisungen in randomisierter Reihenfolge für jeweils 15 Sekunden präsentiert; diese waren zur Hälfte zu lesen und zur Hälfte auszuführen. 15 Handlungsanweisungen des Gesamtpools wurden in der ersten Phase gar nicht präsentiert bzw. enkodiert. Welche Handlungsanweisungen wie zu handhaben waren, wurde durch die Wahl der Materialkombination festgelegt, die fortlaufend erfolgte, um sicherzustellen, dass in jeder Gruppe jede Materialkombination gleich häufig gewählt wurde. Der Ablauf wird anhand von Abbildung 7 veranschaulicht.

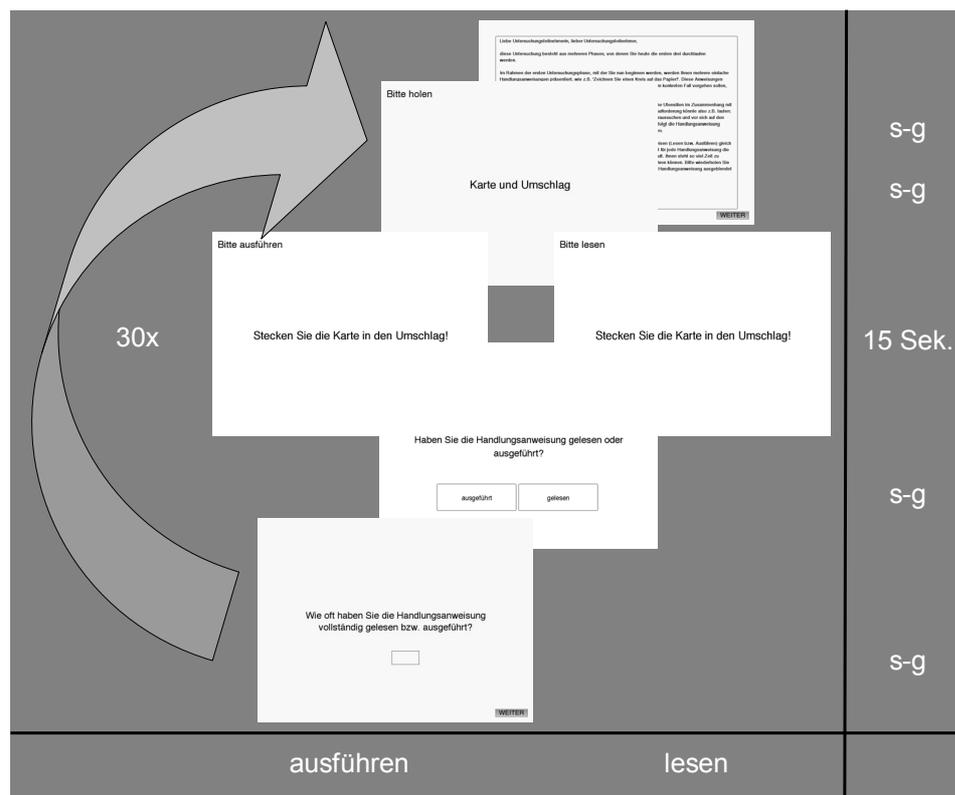


Abb. 7: Ablauf der ersten Phase des ersten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.

Anschließend wurden die TeilnehmerInnen gebeten, die Items des NFC/FII-Fragebogens anhand einer siebenstufigen Skala hinsichtlich des Zutreffens auf die eigene Person zu beurteilen.

In der Verarbeitungsphase (Phase 2) nun unterschieden sich die Instruktionen je nach Gruppenzugehörigkeit der TeilnehmerInnen. Allen Gruppen gemeinsam war, dass die Frequenz in drei Stufen variiert (5x, 1x, 0x) und die Handlungsanweisungen für 15 Sekunden in fixierten Zufallsreihenfolgen präsentiert wurden. Zudem folgten auf jede Präsentation zwei Ratings, von denen das erste für alle Gruppen gleich war und sich auf die Häufigkeit der Verarbeitung bezog; das zweite unterschied sich je nach Gruppe.

Den TeilnehmerInnen der Bedingung *Vorstellen* wurde jeweils eine Handlungsanweisung präsentiert, die zu lesen und deren Umsetzung anschließend vorzustellen war. In der generellen Instruktion wurden sie gebeten, sich die Handlungsausführung möglichst lebhaft – d.h. unter Einbezug möglichst vieler sensorischer Details – vorzustellen. Das Vorstellen sollte bis zum Ende des Intervalls von 15 Sekunden fortgesetzt bzw. wiederholt werden. Anschließend wurde erfasst, wie häufig sich die ProbandInnen die Ausführung vorgestellt hatten. Daneben wurden die TeilnehmerInnen gebeten, die Lebendigkeit ihrer Vorstellung auf einer siebenstufigen Skala einzuschätzen. Dieses Rating sollte die TeilnehmerInnen bei jeder Präsentation erneut daran erinnern, dass sie möglichst lebhaftere Vorstellungen generieren sollten, und der Aufrechterhaltung der Konzentration dienen.

Den TeilnehmerInnen, die die Handlungsausführung *beobachten* sollten, wurden die Handlungsanweisungen zusammen mit den entsprechenden Videos präsentiert. Sie wurden instruiert, jede Handlungsanweisung einmal zu lesen und dann aufmerksam die Handlungsausführung zu beobachten. Um sicherzustellen, dass sie die Handlung auch jedes Mal erneut aufmerksam verfolgten, wurden sie gebeten, jeweils die Anzahl vollständiger Handlungsausführungen zu zählen und nach dem Ausblenden der Präsentation einzugeben. Danach sollte das Ausmaß an Konzentration, das der Beobachtung der Videos geschenkt wurde, auf einer siebenstufigen Skala angegeben werden. Dieses Rating sollte die TeilnehmerInnen fortwährend anregen, sich so intensiv wie möglich auf die Videos zu konzentrieren.

Den TeilnehmerInnen in der *Generieren*-Bedingung wurden die umgestellten bzw. in Anagrammform gebrachten Sätze präsentiert und sie wurden gebeten, die entsprechende Handlungsanweisungen zu generieren, diese per Tastatur in die dafür vorgesehene freie Fläche einzugeben und dann so oft wie möglich zu lesen. Der letzte Arbeitsschritt wurde vor dem Hintergrund der Erfahrungen aus der Voruntersuchung eingebracht, im Rahmen derer sich gezeigt hatte, dass ein wiederholtes Eingeben der korrekten Handlungsanweisung (anstatt des wiederholten Lesens) zu deutlich niedrigeren Häufigkeiten der Verarbeitung im Vergleich zu allen anderen Gruppen führte. Um hier eine Vergleichbarkeit zu gewährleisten, wurde das Lesen der generierten Handlungsanweisung eingeführt. Nach Ablauf der fünfzehnsekündigen Präsentationszeit wurde den TeilnehmerInnen die richtige Lösung gezeigt und sie wurden gebeten, anzugeben, wie häufig sie diese gelesen hatten. In dem Falle, dass sie eine andere, d.h. falsche Lösung gefunden hatten, wurden sie gebeten, in das entsprechende Feld eine 99 einzugeben. Die Darbietung der richtigen Lösung sollte sicherstellen, dass die eigentlich interessierende Handlungsanweisung (wenigstens) einmal verarbeitet wurde. Des Weiteren wurden die ProbandInnen gebeten, die Schwierigkeit des Anagramms auf einer siebenstufigen Skala einzuschätzen. Dieses Rating sollte die Bedeutung des Prozesses unterstreichen.

Die TeilnehmerInnen, die die Handlungsanweisungen nur *lesen* sollten, wurden instruiert, dies einmal zu tun, dann einmal alle Konsonanten zu zählen und danach – bis zum Ende der Präsentation – jede Handlungsanweisung so häufig wie möglich zu lesen. Durch das Zählen der Konsonanten sollte sichergestellt werden, dass den TeilnehmerInnen nicht genug Zeit für eine Elaboration der Handlungsanweisungen (z.B. durch Vorstellen) blieb, sondern diese sehr oberflächlich verarbeitet wurden. Des Weiteren sollte die Zusatzaufgabe verhindern, dass die Frequenz des Lesens deutlich über der Frequenz der Verarbeitung in den anderen Gruppen lag. Außerdem sollte hierdurch sichergestellt werden, dass die TeilnehmerInnen – um sich instruktionsgemäß zu verhalten – die Handlungsanweisung wenigstens einmal überfliegen mussten. Denn während die Häufigkeit des Lesens der Handlungsanweisungen nicht nachprüfbar war, galt dies für die Anzahl an Konsonanten schon. Der Grund für die Wahl dieser speziellen Aufgabe lag darin, dass es relativ unwahrscheinlich erschien, dass TeilnehmerInnen sich die Anzahl an Konsonanten über die verschiedenen Wiederholungen hinweg würden merken können; d.h. sie waren jedes Mal erneut gezwungen, die Handlungsanweisung wenigstens einmal zu überfliegen. Thomas et al. (2003) setzten zum gleichen Zwecke die Aufgabe ein, die Position des längsten Wortes zu bestimmen. Diese Aufgabe erschien allerdings weniger geeignet, da die Position ohne Überfliegen des Satzes, sozusagen auf den ersten Blick, ersichtlich und zudem besser zu merken ist als die Anzahl an Konsonanten pro Handlungsanweisung. Um den Instruktionen Nachdruck zu verleihen, sollten die TeilnehmerInnen nach jeder Präsentation die Häufigkeit des Lesens der Handlungsanweisungen sowie die Anzahl an enthaltenen Konsonanten angeben.

Abbildung 8 zeigt den Ablauf der zweiten Phase.



Abb. 8: Ablauf der zweiten Phase des ersten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.

Die erste Sitzung dauerte insg. ca. 1.5 Stunden.

Die zweite Sitzung fand zwei Wochen nach der ersten statt und war mit einer Dauer von ca. 30 Minuten wesentlich kürzer.

Zunächst wurden TeilnehmerInnen die 45 Handlungsanweisungen der 3 Pools sowie die 10 Distraktor-Handlungsanweisungen in zufälliger Reihenfolge mit der folgenden Aufgabe präsentiert:

Sie sollen bitte jeweils angeben, ob diese Handlungsanweisungen aus der ERSTEN Untersuchungsphase [der letzten Sitzung] stammen, ob Sie diese Handlungsanweisungen in dieser ersten Untersuchungsphase gelesen oder ausgeführt haben, und wie sicher Sie sich bei diesen Einschätzungen sind. Beziehen Sie sich bitte nur auf die ALLERERSTE UNTERSUCHUNGSPHASE!

Wurde also eine Handlungsanweisung als aus der ersten Phase stammend rekonstruiert, sollte zusätzlich spezifiziert werden, ob diese ausgeführt oder nur gelesen worden war. Zudem sollte die Sicherheit des Erinnerungsurteils auf einer siebenstufigen Skala mit den Polen „gar nicht sicher“ und „absolut sicher“ angegeben werden. Wurde die Zugehörigkeit einer Handlungsanweisung zur ersten Phase negiert, folgte direkt das Sicherheitsrating (siehe Abbildung 9).

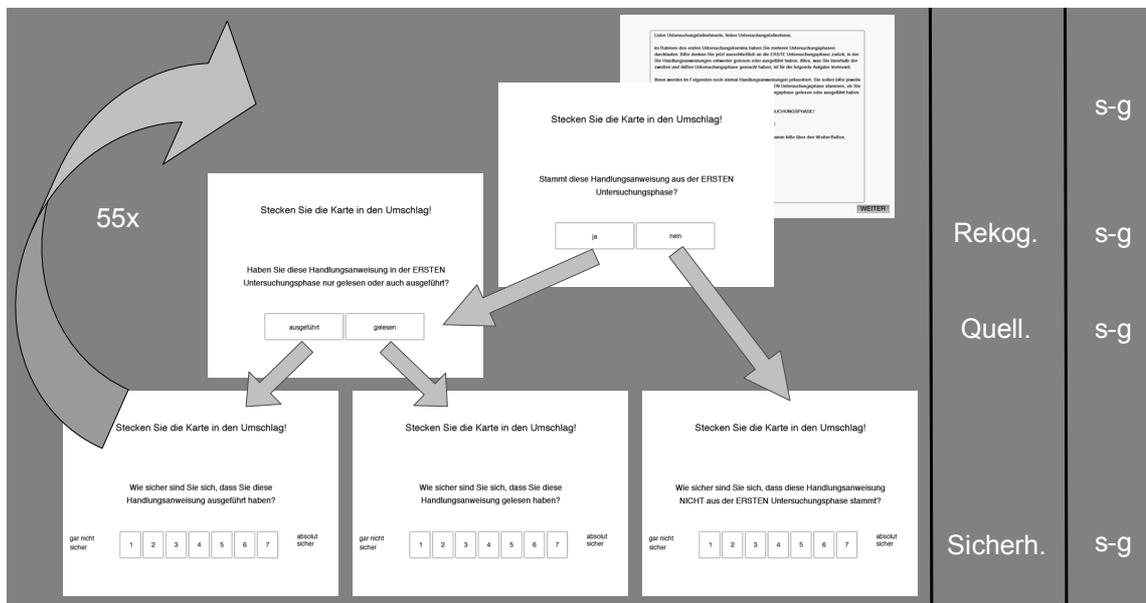


Abb. 9: Ablauf der dritten Phase des ersten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung. Rekog.: Rekonognitionstest. Quell.: Quellengedächtnistest. Sicherh.: Sicherheitsrating.

Anschließend wurden die TeilnehmerInnen gebeten, die insgesamt 25 Items des VIVIQ, HZI und FDS auf einer ebenfalls siebenstufigen Skala hinsichtlich ihres Zutreffens auf die eigene Person einzuschätzen.

Abschließend folgten zwei Ratings, die die Untersuchung betrafen und erfassen sollten, wie gut die TeilnehmerInnen die Instruktionen während der beiden Untersuchungsphasen befolgt hatten.

I.5.3 Ergebnisse

I.5.3.1 Deskriptive Statistiken

Die folgenden Tabellen beinhalten die relativen Häufigkeiten von *Phase 1-* bzw. *ausgeführt-* Antworten, d.h. sowohl Rekognitions- als auch Quellengedächtnisurteile. Für die inferenzstatistische Auswertung wurden die beiden Arten nicht ausgeführter Handlungsanweisungen zusammengefasst (Handlungsanweisungen, die in Phase 1 überhaupt nicht enkodiert wurden, bzw. Handlungsanweisungen, die in Phase 1 nur gelesen wurden; vgl. Abschnitt I.5.1.3). Um gleiche Zellbesetzungen der verglichenen Itemarten zu gewährleisten, wurden die Distraktoren, d.h. diejenigen Items, die weder in Phase 1 noch in Phase 2, sondern ausschließlich in Phase 3 präsentiert wurden, nicht in die Tabelle und in die Auswertung aufgenommen. Diese wurden aber mit in allen Gruppen vergleichbar niedrigen Quoten beantwortet (*Phase 1* (M (SD)): Vorstellen: .05 (.13), Beobachten: .05 (.06), Generieren: .03 (.06), Lesen: .05 (.06); *ausgeführt* (M (SD)): Vorstellen: .01 (.03), Beobachten: .01 (.03), Generieren: .01 (.03), Lesen: .01 (.04)).

Die *gelesen-*Antworten im Quellengedächtnistest, die nicht einfach das Komplement der *ausgeführt-*Antworten darstellen, sondern die Differenz zwischen *ausgeführt-* und *Phase 1-*Antworten (siehe auch Abschnitt I.5.4.1), können an dieser Stelle nicht dargestellt bzw. analysiert werden.

Tab. 13: Relative Häufigkeiten von *Phase 1*-Antworten im Rekognitionstest des ersten Experiments.

		Art der Verarbeitung in Phase 2:							
		Vorstellen		Beobachten		Generieren		Lesen	
		Art der Enkodierung in Phase 1: Ausgeführt							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.99	.05	.99	.05	.96	.08	.92	.13
	1x	.96	.11	.92	.17	.93	.12	.97	.07
	0x	.83	.17	.85	.21	.91	.13	.92	.13
		Art der Enkodierung in Phase 1: Gelesen							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.84	.19	.87	.16	.81	.26	.79	.18
	1x	.80	.21	.80	.19	.72	.18	.65	.23
	0x	.60	.25	.56	.28	.63	.21	.65	.27
		Art der Enkodierung in Phase 1: Nicht enkodiert							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.67	.23	.56	.28	.52	.25	.28	.24
	1x	.31	.21	.35	.26	.16	.20	.11	.13
	0x	.07	.10	.12	.15	.07	.12	.04	.11

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *Phase 1*-Antworten relativiert an der Anzahl aller (gültigen) Handlungsanweisungen in der entsprechenden Zelle. $N=60$, d.h. 15 pro Gruppe.

Tab. 14: Relative Häufigkeiten von *ausgeführt*-Antworten im Quellengedächtnistest des ersten Experiments.

		Art der Verarbeitung in Phase 2:							
		Vorstellen		Beobachten		Generieren		Lesen	
		Art der Enkodierung in Phase 1: Ausgeführt							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.91	.20	.91	.10	.80	.17	.69	.20
	1x	.81	.23	.81	.28	.79	.24	.79	.19
	0x	.70	.25	.61	.19	.81	.18	.69	.16
		Art der Enkodierung in Phase 1: Gelesen							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.33	.26	.23	.29	.08	.11	.01	.05
	1x	.20	.19	.05	.09	.08	.15	.05	.09
	0x	.05	.09	.06	.10	.09	.13	.00	.00
		Art der Enkodierung in Phase 1: Nicht enkodiert							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.21	.24	.12	.13	.01	.05	.03	.07
	1x	.07	.10	.01	.05	.03	.10	.00	.00
	0x	.01	.05	.04	.08	.01	.05	.00	.00

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *ausgeführt*-Antworten relativiert an der Anzahl aller (gültigen) Handlungsanweisungen in der entsprechenden Zelle. $N=60$, d.h. 15 pro Gruppe.

Da es sich bei den a priori aufgestellten Hypothesen um einfache (paarweise) Einzelvergleiche handelte, die zudem gerichtet formuliert worden waren, stellte der t-Test das Verfahren der Wahl dar (vgl. Hussy & Jain, 2002). Je nach zugrundeliegender Hypothese wurde dabei entweder auf den t-Test für abhängige (intraindividuelle Variation des Faktors „Frequenz der Präsentation in Phase 2“) oder für unabhängige Stichproben (interindividuelle Variation des Faktors „Art der Verarbeitung in Phase 2“) zurückgegriffen. Die Prüfung auf Signifikanz erfolgte einseitig. Auch im Rahmen der explorativen Analysen wurden für die Berechnung einfacher Einzelvergleiche t-Tests, beim Vorliegen mehrerer Faktoren bzw. Faktorstufen Varianzanalysen eingesetzt; zudem wurden Produkt-Moment-Korrelationen und χ^2 -Tests berechnet. Die Signifikanzprüfung im Rahmen der explorativen Analysen erfolgte zweiseitig.

Wie bereits aus den Tabellen ersichtlich, traten im Rahmen der Untersuchung sowohl Deckeneffekte (siehe z.B. Rekognitionsleistung, korrekte *Phase 1*-Antworten, für in der ersten Phase ausgeführte Handlungsanweisungen) als auch Bodeneffekte (siehe z.B. Quellengedächtnisleistung, falsche *ausgeführt*-Antworten, für in der ersten Phase nicht ausgeführte und in der zweiten Phase nullmalig präsentierte Handlungsanweisungen) auf. An dieser Tatsache ist bereits erkennbar, dass die Anwendungsvoraussetzungen des t-Tests, insb. Normalverteilung und Varianzhomogenität (z.B. Bortz, 2005; Diehl & Arbinger, 1992; Nachtigall & Wirtz, 1998), teilweise verletzt waren. Obwohl sowohl der t-Test für abhängige als auch der t-Test für unabhängige Stichproben auf Verletzungen der Anwendungsvoraussetzungen robust reagieren (z.B. Bortz, 2005), wurden die betroffenen Daten zusätzlich anhand von non-parametrischen Verfahren analysiert. Diese werden jedoch nicht weiter erwähnt, da sie in allen Fällen zu derselben Hypothesenentscheidung führten wie die parametrischen Analysen. Da bei der Hypothesenprüfung ein strenges Kriterium angesetzt wurde (logische und-Verknüpfungen aller Unterhypothesen) wird eine α -Fehler-Adjustierung als nicht notwendig erachtet (vgl. Hussy & Jain, 2002).

I.5.3.2 Manipulation-Checks

Wie bereits erläutert, wurden die TeilnehmerInnen in der ersten Phase gebeten, autonom zu arbeiten – d.h. die Gegenstände selbst auszuwählen und diese (je nach Instruktion) zur Handlungsausführung zu nutzen oder sie zu ignorieren und die Handlungsanweisungen nur zu lesen. Dabei stand ihnen genug Zeit zur Verfügung, eine Handlung mehrmals auszuführen bzw. eine Handlungsanweisung mehrmals zu lesen. Nach jeder Präsentation wurden die TeilnehmerInnen daher gebeten, anzugeben, ob und wie häufig sie die vorausgegangene Handlungsanweisung ausgeführt bzw. gelesen hatten. Stimmt die per Computer vermittelte Anweisung und die Angabe der TeilnehmerInnen bzgl. der Handhabung einer Handlungsanweisung nicht überein, wurde das entsprechende Item aus der Analyse der Quellengedächtnisdaten ausgeschlossen. Da dies nur selten der Fall war (die Quote an Handlungsanweisungen, bei denen die Angabe der TeilnehmerInnen und die computervermittelte Anweisung nicht übereinstimmten, betrug 1.16%) und die TeilnehmerInnen zudem augenscheinlich gut mit dem Prozedere zurecht kamen, kann angenommen werden, dass die in die Analyse einbezogenen Handlungsanweisungen auch tatsächlich so gehandhabt wurden wie vorgesehen. Dennoch sollte dies zusätzlich anhand von drei Maßen abgesichert werden. Erstens war davon auszugehen, dass das Lesen der Handlungsanweisungen weniger Zeit kosten und somit häufiger getätigt werden würde als das Ausführen der Handlungen. Daher sollte – bei korrekter Handhabung durch die TeilnehmerInnen – die mittlere Frequenz des Lesens höher liegen als dieje-

nige des Ausführens. Zweitens wurde angenommen, dass zwar aufgrund der unterschiedlichen Länge der Handlungsanweisungen auch die benötigte Zeit und damit die Häufigkeit des Lesens der verschiedenen Handlungsanweisungen variieren würde, jedoch sollte dies nicht so stark der Fall sein, wie beim Ausführen der unterschiedlich aufwändigen Handlungen. Beim Ausführen waren aufgrund der unterschiedlichen Komplexität der Handlungen (z.B. eine Batterie aus dem Wecker entfernen vs. den Deckel von einer Schachtel nehmen) größere Zeitunterschiede und somit eine größere Variabilität der Frequenzen zu erwarten. Drittens sollten ausgeführte im Vergleich zu nur gelesenen Handlungsanweisungen deutlich besser erinnert werden (enactment-Effekt, siehe Abschnitt I.2.3). Alle drei Effekte sollten unabhängig von der Art der Verarbeitung in der zweiten Phase auftreten, wurden also für die Gesamtstichprobe von $N=60$ TeilnehmerInnen geprüft.

Um die Validität der Daten abzusichern, wurde vor diesem Hintergrund zunächst überprüft, ob die TeilnehmerInnen angegeben hatten, die Handlungsanweisungen in Phase 1 im Schnitt häufiger gelesen als ausgeführt zu haben. Hierzu wurden für jeden Teilnehmer bzw. jede Teilnehmerin gesondert Mittelwert und Standardabweichung der Häufigkeit des Ausführens bzw. Lesens bestimmt, indem die entsprechenden Angaben für die 15 ausgeführten bzw. gelesenen Handlungsanweisungen zugrunde gelegt wurden. Die beiden resultierenden Parameter für Ausführen resp. Lesen wurden dann über alle TeilnehmerInnen hinweg gemittelt und hinsichtlich eines Einflusses der Art der Enkodierung analysiert. Es zeigte sich das erwartete Ergebnismuster: Die durchschnittliche Frequenz des Lesens betrug 7.04 ($SD=2.36$) und war damit deutlich höher als die durchschnittliche Frequenz des Ausführens ($M=4.78$, $SD=1.50$); $t(59)=9.79$, $p<.05$, $\eta^2=.619$. Des Weiteren war die Variabilität der Häufigkeitsangaben bei den ausgeführten größer als bei den gelesenen Handlungsanweisungen: Die durchschnittliche Standardabweichung der Frequenz des Ausführens betrug 3.39 ($SD=1.62$), die des Lesens nur 1.58 ($SD=1.09$); $t(59)=7.61$, $p<.05$, $\eta^2=.495^9$. Drittens war zur Sicherung der Datenvalidität ein enactment-Effekt nachzuweisen, d.h. bessere Rekognitionsraten von in Phase 1 ausgeführten im Vergleich zu in Phase 1 nicht ausgeführten Handlungsanweisungen. Die in Phase 1 nicht enkodierten Handlungsanweisungen wurden hierbei nicht berücksichtigt. Es fand sich ein deutlicher enactment-Effekt: Die Rekognitionsquote der ausgeführten Handlungsanweisungen lag bei .93 ($SD=.08$) und war damit deutlich höher als diejenige der gelesenen, die bei .73 ($SD=.16$) lag; $t(59)=10.71$, $p<.05$, $\eta^2=.660$.

⁹ Bei der – mathematisch korrekteren – Analyse der Varianzen ergibt sich das identische Bild.

In der zweiten Phase der ersten Sitzung wurden nach jeder der 90 Präsentationen zwei Ratings erhoben, um eine instruktionsgemäße Handhabung durch die ProbandInnen zu begünstigen. Dabei sollte zum einen die Häufigkeit des Vorstellens, Beobachtens usw. angegeben werden. Auf diese Daten wird an späterer Stelle eingegangen. Zum anderen wurden verarbeitungsspezifische Einschätzungen erhoben. Beim Lesen wurde für jede Handlungsanweisung die Anzahl an Konsonanten abgefragt; diese wurde jedoch nicht auf Korrektheit überprüft bzw. näher analysiert. Beim Generieren sollte die Schwierigkeit der Anagramme eingeschätzt werden; die Ergebnisse wurden bereits in Abschnitt I.5.2.2 berichtet. Beim Beobachten sollte das Ausmaß der auf die Videos gerichteten Konzentration, beim Vorstellen die Lebhaftigkeit der Vorstellung eingeschätzt werden. Die Konzentration über alle Präsentationen und alle TeilnehmerInnen hinweg wurde mit 6.08 ($SD=.74$), die Lebhaftigkeit der Vorstellung mit 5.48 ($SD=.79$) eingeschätzt (beides auf einer siebenstufigen Skala, höhere Zahl entspricht höherer Ausprägung). Beide Werte sind – als Mittelwerte über 90 Präsentationen und jeweils 15 Personen – als sehr hoch anzusehen und können als Beleg dafür gelten, dass die TeilnehmerInnen (zumindest in den Bedingungen Beobachten und Vorstellen) auch in Phase 2 instruktionsgemäß vorgingen.

Im Rahmen der zweiten Sitzung wurde explizit und für beide Phasen getrennt nach dem Befolgen der Instruktionen gefragt. Das erste diesbzgl. Item lautete: „Denken Sie nun bitte noch einmal an den ersten Untersuchungstermin zurück und antworten Sie bitte ganz ehrlich. Die Beantwortung hat keinerlei Konsequenzen für Sie, ist aber für die Auswertbarkeit und Interpretierbarkeit meiner Daten sehr wichtig. Wie sorgfältig haben Sie während der ERSTEN Untersuchungsphase die Instruktionen befolgt, d.h. Handlungsanweisungen ausgeführt, wenn diese ausgeführt werden sollten, und gelesen, wenn diese gelesen werden sollten?“ Im Durchschnitt gaben die TeilnehmerInnen auf der siebenstufigen Antwortskala mit den Polen gar nicht bzw. absolut sorgfältig den Wert 6.69 ($SD=.50$) an, wobei es keinen Unterschied zwischen den Bedingungen gab; $F(3,55)=1.20$, $p>.05$. Auf ein ähnliches Item bzgl. Phase 2 („Wie sorgfältig haben Sie während der LETZTEN Untersuchungsphase die Instruktionen befolgt? D.h. wie sorgfältig haben Sie (je nach Bedingung) die Handlungsanweisungen gelesen BZW. geschrieben und gelesen BZW. gelesen und die Videos verfolgt BZW. gelesen und sich die Handlungsausführung vorgestellt? Es geht mir bei dieser Frage nicht darum, wie leicht oder schwer Ihnen das Befolgen der Instruktionen gefallen ist, sondern ausschließlich darum, wie sehr Sie sich bemüht haben, die Instruktionen zu befolgen.“) fielen die Einschätzungen – wie aufgrund der zahlreichen Wiederholungen zu erwarten – etwas niedriger und variabler aus ($M=6.27$; $SD=.94$), jedoch ergab sich auch hier kein Unterschied zwischen den Bedingungen; $F(3,55)=.91$, $p>.05$.

Insgesamt kann also von einer sehr guten compliance ausgegangen werden, die sich auch zwischen den Bedingungen nicht unterschied.

Eine Grundidee der Untersuchung bestand darin, durch die verschiedenen Arten der Verarbeitung auch den Ursprung der Information (internal, external) bzw. den Status der Person (aktiv, passiv) und somit die Tiefe der Verarbeitung zu berücksichtigen. Dabei wurde davon ausgegangen, dass Vorstellen im Vergleich zu Beobachten und Generieren im Vergleich zu Lesen aufgrund des internalen Informationsursprungs bzw. aktiven Status' der Person und der Kontrolliertheit der Prozesse tiefere Formen der Verarbeitung darstellen und somit zu einer gesteigerten Verarbeitungsleichtigkeit bzw. Vertrautheit führen sollten. Bekanntermaßen äußert sich Vertrautheit bzw. Verarbeitungsleichtigkeit von Stimuli in einem Rekognitionstest darin, dass diese eher als alt eingeschätzt werden – korrekter- oder fälschlicherweise (siehe Abschnitt I.1.3). So werden z.B. Items, die bei der Enkodierung gar nicht, jedoch bei der Abfrage mehrmals präsentiert werden, bei wiederholter Darbietung manchmal fälschlicherweise als aus der Enkodierphase stammend bzw. alt eingestuft (siehe *repetition-lag-Paradigma*; Underwood & Freund, 1970; Jennings & Jacoby, 1997). Daher sollte zunächst grundlegend gezeigt werden können, dass – nach jedweder Art der Verarbeitung – Handlungsanweisungen, die nicht in Phase 1 dargeboten, jedoch fünfmalig in Phase 2 verarbeitet wurden, vermehrt fälschlich der Phase 1 zugeschrieben werden, verglichen mit solchen Handlungsanweisungen, die weder in Phase 1 noch in Phase 2 dargeboten wurden. In der Tat zeigte sich dieser Effekt für jede der vier Verarbeitungsformen; Vorstellen: $t(14)=8.87$, $p<.05$, $\eta^2=.849$; Beobachten: $t(14)=4.68$, $p<.05$, $\eta^2=.610$; Generieren: $t(14)=7.55$, $p<.05$, $\eta^2=.803$; Lesen: $t(14)=3.52$, $p<.05$, $\eta^2=.470$. Gesteigerte Vertrautheit durch tiefere Formen der Verarbeitung in Phase 2 sollte sich entsprechend in einer höheren Quote falscher Zuordnungen äußern. Mit anderen Worten war also zu erwarten, dass eine mehrfache tiefe Verarbeitung in Phase 2 mehr Fehlattributionen zu Phase 1 hervorrufen würde als eine mehrfache oberflächliche Verarbeitung. Aus den deskriptiven Statistiken ging in der Tat hervor, dass der Zuwachs an Fehlattributionen nach fünf- im Vergleich zum nullmaligem Vorstellen größer war als der entsprechende Zuwachs beim Beobachten. Beim Generieren war dieser Zuwachs im Vergleich zum Lesen sogar noch deutlicher ausgeprägt. Inferenzstatistisch ließ sich die Interaktion allerdings nur für den Vergleich Generieren – Lesen absichern (Generieren-Lesen: $t(28)=2.35$, $p<.05$, $\eta^2=.165$; Vorstellen- Beobachten: $t(28)=1.38$, $p>.05$). Die Interaktion wurde allerdings ebenfalls signifikant, wenn beide Formen tiefer Verarbeitung mit beiden Formen oberflächlicher Verarbeitung verglichen wurden: Vorstellen/ Generieren vs. Beobachten/Lesen; $t(58)=2.46$, $p<.05$, $\eta^2=.094$. Dieses Ergebnismuster lässt sich aus meiner Sicht als Bestätigung einer gelungenen Manipulation der Verarbeitungstiefe auffassen. Zwar konnte einer der möglichen Vergleiche inferenzstatistisch nicht abgesichert werden – allerdings basierte die experimentelle Planung, insb. die Bestim-

mung der optimalen Stichprobengröße, auch nicht auf diesen Vergleichen. Deskriptivstatistisch weist das Ergebnismuster in die erwartete Richtung und wäre vermutlich bei größerer Stichprobe auch inferenzstatistisch abzusichern gewesen.

Eine weitere – nur kurz ausgeführte – Annahme bestand darin, dass Vorstellen und Beobachten aufgrund der Anreicherung mit sensorischen/ motorischen Details unabhängig von der Tiefe der Verarbeitung zu einer stärkeren Gedächtnisspur führen sollten als die Verarbeitungsformen ohne sensorische/ motorische Anreicherung (Generieren und Lesen). Dies sollte bereits nach einer einzigen Exposition festgestellt werden können. Daher wurde – der bereits erläuterten Logik folgend – verglichen, ob eine Interaktion zwischen der Frequenz (1x vs. 0x) und der Art der Verarbeitung (sensorisch/ motorisch angereichert (Vorstellen, Beobachten) vs. sensorisch/ motorisch nicht angereichert (Generieren, Lesen)) hinsichtlich solcher Handlungsanweisungen existierte, die fälschlich als aus Phase 1 stammend erinnert wurden. Angenommen wurde, dass eine solche Interaktion existierte, dass also die Differenz zwischen ein- und nullmaliger Darbietung nach Vorstellen und Beobachten größer sein sollte als nach Generieren und Lesen. Diese Annahme wurde durch die Ergebnisse bestätigt ($t(58)=2.54$, $p<.05$, $\eta^2=.100$) und zeigt die Bedeutung sensorischer und/ oder motorischer Informationen für die Erinnerung auf.

I.5.3.3 Hypothesenprüfung

Zunächst sollte überprüft werden, ob sich innerhalb der einzelnen Formen der Verarbeitung in der zweiten Phase inflation-Effekte nachweisen ließen. Hierzu wurde für jede Art der Verarbeitung einzeln verglichen, ob falsche *ausgeführt*-Antworten häufiger nach fünfmaliger als nach nullmaliger Präsentation in Phase 2 auftraten.

Ein solcher Anstieg an falschen *ausgeführt*-Antworten konnte nach wiederholtem Vorstellen ($t(14)=3.80$, $p<.05$, $\eta^2=.508$) und Beobachten ($t(14)=3.61$, $p<.05$, $\eta^2=.483$), jedoch nicht nach wiederholtem Generieren ($t(14)=.26$, $p>.05$) oder Lesen ($t(14)=1.38$, $p>.05$) nachgewiesen werden.

Des Weiteren wurde geprüft, ob signifikante Interaktionen zwischen der Frequenz (5x vs. 0x) und bestimmten Formen der Verarbeitung (Vorstellen vs. Generieren; Beobachten vs. Lesen) vorlagen. Um diese Interaktionen gerichtet überprüfen zu können, wurde die Differenz falscher *ausgeführt*-Antworten nach fünf- vs. nullmaliger Präsentation für die verschiedenen Formen der Verarbeitung anhand von t-Tests für unabhängige Stichproben miteinander verglichen. Hierbei ergaben sich signifikante Interaktionen zwischen der Frequenz der Präsentation

in Phase 2 und dem Vorstellen bzw. Generieren ($t(16.50)=3.72, p<.05, \eta^2=.331$) sowie dem Beobachten bzw. Lesen ($t(18.76)=2.80, p<.05, \eta^2=.219$). Demnach lässt sich der inflation-Effekt auch nach Kontrolle des Generierens im Falle des Vorstellens bzw. nach Kontrolle des wiederholten Lesens im Falle des Beobachtens nachweisen.

Für die Interpretation dieser Effekte könnten zwei Aspekte als problematisch angesehen werden. Erstens wurden 9 Kombinationen gebildet, aber nur 15 anstatt der angezielten 18 TeilnehmerInnen pro Gruppe untersucht. Demnach konnten die Materialkombinationen innerhalb der Gruppen nicht gleich häufig eingesetzt werden. Dies könnte zur Konsequenz haben, dass mögliche Materialeffekte nicht ausgeglichen werden konnten und somit den intraindividuellen Vergleich bzgl. der Frequenzen beeinflussten. Aufgrund des unsystematischen Ausfalls von TeilnehmerInnen und den begrenzten Ressourcen konnten zudem nicht in allen Formen der Verarbeitung alle Kombinationen gleich häufig eingesetzt werden, was wiederum die interindividuellen Vergleiche beeinflusst haben könnte. Diesem Problem ist jedoch entgegenzusetzen, dass die Materialpools mit größter Sorgfalt parallelisiert wurden, um mögliche Materialeffekte von vornherein auszuschließen (vgl. Abschnitt I.5.2.2). Auch fallen die Abweichungen bei der Nutzung der Materialpools deutlich geringer aus als die Übereinstimmungen. So entsprechen bspw. die Materialkombinationen von 15 Personen, die in der 2. Phase die Handlungsanweisungen lasen, in 14 Fällen denjenigen von Personen, die in Phase 2 die Handlungsausführung beobachteten. Insgesamt gab es also nur bei 2 von 30 TeilnehmerInnen beider Gruppen Abweichungen. Die gefundenen Effekte sind jedoch sehr stark, weshalb es unplausibel erscheint, dass diese (allein) durch die geringen Abweichungen – dazu auch noch differentiell und systematisch in den unterschiedlichen Arten der Verarbeitung – hervorgerufen worden sein könnten. Aber nicht nur argumentativ, sondern auch auf empirischer Ebene lässt sich diese Problematik bewerten. Hierfür wurde überprüft, ob unter den Bedingungen Generieren und Lesen, in denen kein inflation-Effekt nachzuweisen war, und unter den Bedingungen Vorstellen und Beobachten, in denen inflation-Effekte gefunden wurden, ein signifikanter Einfluss der Materialkombination auf das Ausmaß falscher Erinnerungen existierte. Hierzu wurde eine Varianzanalyse mit dem interindividuellen Faktor Materialkombination und der abhängigen Variable *falsche ausgeführt*-Antworten für die Frequenzen 5x und 0x durchgeführt. Unter den Bedingungen Vorstellen und Beobachten zeigte sich für keine der beiden Frequenzen ein signifikanter Einfluss der Materialkombination (5x: $F(8,21)=2.13, p>.05$; 0x: $F(8,21)=2.28, p>.05$); das gleiche Bild zeichnete sich für die Bedingungen Generieren und Lesen ab (5x: $F(8,21)=1.94, p>.05$, 0x: $F(8,21)=1.19, p>.05$).

Diese Überlegungen bzw. Ergebnisse räumen demnach die Möglichkeit aus, dass etwaige materialspezifische Effekte für die gefundenen Ergebnismuster verantwortlich sein könnten.

I.5.3.4 Weitere Befunde

Weitere Analysen sollten dazu dienen, die falschen Erinnerungen näher zu charakterisieren und zusätzliche Hinweise auf die ihnen zugrundeliegenden Mechanismen zu generieren. Zunächst sollen jedoch die vier sehr unterschiedlichen Formen der Verarbeitung hinsichtlich ihrer Auswirkungen auf korrekte Erinnerungsleistungen analysiert werden. Einerseits könnte dies Anhaltspunkte bzgl. der Vergleichbarkeit der Prozesse und damit bzgl. der Interpretierbarkeit der Ergebnisse liefern, andererseits könnte das Ergebnismuster ebenfalls aufschlussreich hinsichtlich des Zustandekommens von falschen *ausgeführt*-Urteilen sein.

Univariate Varianzanalysen ergaben bei einer von den Frequenzstufen unabhängigen Betrachtung keine signifikanten Effekte des Faktors „Art der Verarbeitung“ auf die Rekognitionsquoten von in Phase 1 ausgeführten ($F(3,56)=.13, p>.05$) bzw. gelesenen ($F(3,56)=.28, p>.05$) Items. Alle vier Prozesse führten also zu einer vergleichbaren Anzahl an korrekt rekognizierten ausgeführten bzw. gelesenen Handlungsanweisungen. Dasselbe traf auch für die korrekten *ausgeführt*-Antworten im Quellengedächtnistest zu ($F(3,56)=1.33, p>.05$). Auf deskriptivstatistischer Ebene zeigt sich bei letzteren Erinnerungsurteilen aber interessanterweise ein Einfluss der Frequenz: Nach fünfmaliger Präsentation in Phase 2 gaben die TeilnehmerInnen, die sich die Handlungsausführung vorgestellt bzw. diese beobachtet hatten, mehr korrekte *ausgeführt*-Antworten ab als nach nullmaliger Präsentation – eine Analogie also zum Ergebnismuster der falschen *ausgeführt*-Antworten. Ein solcher Frequenzeffekt fand sich nach Generieren und Lesen nicht (signifikante Interaktion Frequenz (5x vs. 0x) x Art der Verarbeitung (Vorstellen vs. Beobachten vs. Generieren vs. Lesen) hinsichtlich korrekter *ausgeführt*-Antworten: $F(3,56)=4.68, p<.05, \eta^2=.200$)! Dieselbe Tendenz ließ sich – allerdings etwas weniger stark ausgeprägt – auch für die Rekognition der ausgeführten Handlungsanweisungen nachweisen ($F(3,56)=3.27, p<.05, \eta^2=.149$).

Hinsichtlich der falschen *ausgeführt*-Antworten wurde zunächst überprüft, wie viele TeilnehmerInnen in den einzelnen Bedingungen einen inflation-Effekt (mehr falsche *ausgeführt*-Antworten nach fünfmaliger verglichen mit nullmaliger Präsentation in Phase 2) aufwiesen. Es zeigte sich dasselbe Muster wie auch beim Mittelwertsvergleich: Nach Generieren bzw. Lesen wiesen lediglich 3 bzw. 2 der 15 TeilnehmerInnen einen inflation-Effekt auf, während dies für 11 bzw. 10 der 15 TeilnehmerInnen nach Vorstellen bzw. Beobachten der Fall war. Dabei weicht die Verteilung der TeilnehmerInnen in der Bedingung „Vorstellen“ signifikant von derjenigen in der Bedingung „Generieren“ ab ($\chi^2(1,30)=8.57, p<.05; \Phi^2=.285$); das Gleiche gilt für die Bedingung „Beobachten“ im Vergleich zur Bedingung „Lesen“ ($\chi^2(1,30)=8.89, p<.05, \Phi^2=.296$).

Die Ergebnismuster bisheriger Untersuchungen zeigten – auch wenn dieser Effekt nicht inferenzstatistisch analysiert bzw. berichtet wurde –, dass die Quote falscher *ausgeführt*-Antworten für in Phase 1 präsentierte, aber nicht ausgeführte, Handlungsanweisungen durchgängig höher lag als für die nicht in Phase 1 präsentierten. Es konnte davon ausgegangen werden, dass sich dasselbe Ergebnismuster gerade in der vorliegenden Untersuchung herauskristallisieren würde, da ein entscheidender Unterschied zwischen den in Phase 1 gelesenen und den in Phase 1 nicht enkodierten Items darin bestand, dass die TeilnehmerInnen zu ersteren (wie auch bei der Ausführung) in jedem Falle ein Objekt in der Hand gehalten hatten. Aufgrund der verminderten Distinktheit sollte die Entscheidung über die Handhabung einer Handlungsanweisung in ersterem Falle schwerer fallen und somit fehleranfälliger sein. Für die Überprüfung dieser Annahme wurden die falschen *ausgeführt*-Antworten für in Phase 1 gelesene und für in Phase 1 nicht enkodierte Items über die drei Präsentationsfrequenzen gemittelt und miteinander verglichen. Abgesehen vom Lesen, das ohnehin fast keine falschen Erinnerungen hervorrief, zeigten sich dabei für alle weiteren Arten der Verarbeitung deskriptivstatistisch höhere Quoten falscher *ausgeführt*-Antworten für die in Phase 1 gelesenen im Vergleich zu den nicht enkodierten Handlungsanweisungen. Inferenzstatistisch ließ sich dies allerdings nur für das Vorstellen und das Generieren bestätigen (Vorstellen: $t(14)=3.46$, $p<.05$, $\eta^2=.461$; Beobachten: $t(14)=1.46$, $p>.05$; Generieren: $t(14)=3.00$, $p<.05$, $\eta^2=.392$; Lesen: $t(14)=1.38$, $p>.05$).

Die Sicherheit, mit der TeilnehmerInnen fälschlicherweise erinnerten, Handlungen ausgeführt zu haben, war insgesamt relativ hoch. Die folgende Tabelle stellt eine Übersicht über die Mittelwerte der Einschätzungen auf der siebenstufigen Ratingskala dar. Zusätzlich ist der Anteil an Handlungsanweisungen, die mit absoluter Sicherheit, also der Ratingstufe 7, versehen wurden, für fünfmal präsentierte Handlungsanweisungen angegeben.

Tab. 15: Sicherheitsratings (siebenstufige Skala) bzgl. falscher *ausgeführt*-Antworten im ersten Experiment.

Frequ.	Vorstellen			Beobachten			Generieren			Lesen		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
5x	41	5.68	1.42	26	5.58	1.42	7	5.29	0.76	3	3.00	0.00
1x	20	5.30	1.81	6	5.50	2.07	8	5.88	0.83	4	4.25	2.22
0x	5	5.20	2.05	7	5.57	1.13	8	4.00	1.51	0	-	-
∅	66	5.53	1.58	39	5.56	1.45	23	5.04	1.33	7	3.71	1.70
	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>
5x	41	39%	16	26	35%	9	7	0%	0	3	0%	0

Anmerkungen: *N*=Anzahl an Fehlern; *f_%(7)*=Prozentuale Häufigkeit der höchsten Ratingstufe (7).

Auf eine inferenzstatistische Auswertung soll aufgrund der ungleichen Zellenbesetzungen und der teilweisen Abhängigkeit der Werte (eine Person kann mehrere falsche Erinnerungen und daher Sicherheitsratings produziert haben) verzichtet werden. Die Werte lassen aber erkennen, dass falsche *ausgeführt*-Antworten nach wiederholtem Lesen mit der geringsten Sicherheit behaftet waren; nach Vorstellen, Beobachten und Generieren lag die Sicherheit deutlich höher, wobei Vorstellen und Beobachten näher beieinander lagen und eine etwas höhere Sicherheit erzeugten als das Generieren. Auch wurde die höchste Stufe der Sicherheit etwa gleich häufig nach Vorstellen und Beobachten, jedoch gar nicht nach Generieren oder Lesen vergeben. Ein Frequenzeffekt im Sinne ansteigender Sicherheit mit zunehmender Verarbeitungshäufigkeit zeigt sich deutlich für das Generieren und weniger deutlich für das Vorstellen; in der Bedingung Beobachten existiert kein solcher Effekt.

Die falschen Erinnerungen an die Handlungsausführung nach Beobachtung wurden hinsichtlich eines Einflusses der beobachteten AkteurInnen analysiert. Hierzu wurde ausgewertet, wie häufig falsche Erinnerungen in Abhängigkeit vom Geschlecht des Akteurs bzw. der Akteurin und zusätzlich in Abhängigkeit vom Geschlecht des Beobachters bzw. der Beobachterin auftraten.

Tab. 16: Falsche *ausgeführt*-Antworten in Abhängigkeit vom Geschlecht der Akteurin/ des Akteurs bzw. der Beobachterin/ des Beobachters im ersten Experiment.

		Geschlecht Beobachter(in)		
		männlich	weiblich	
Geschlecht Akteur(in)	männlich	3	13	16
	weiblich	4	12	16
		7	25	32

Diese Verteilung wurde aufgrund der Tatsache, dass eine Person auch mehrere Fehlattributionen erbracht haben konnte, so dass die Werte nicht als unabhängig betrachtet werden können, ebenfalls nicht inferenzstatistisch ausgewertet. Allerdings erscheint bei dieser fast vollkommen identischen Verteilung falscher Erinnerungen hinsichtlich des Geschlechts des Akteurs bzw. der Akteurin eine inferenzstatistische Absicherung auch redundant. Die Tatsache, dass Frauen mehr Fehlattributionen vornahmen als Männer, ist auf die deutlich höhere Quote an Frauen in dieser Stichprobe zurückzuführen (13 Frauen, 2 Männer). Da sich das Verhältnis jedoch nicht perfekt entsprach (den Teilnehmern unterliefen verhältnismäßig mehr Fehler als den Teilnehmerinnen), wurden weitere Maße herangezogen, um einen signifikanten Einfluss des Faktors Geschlecht ausschließen zu können. Daher wurde ausgezählt, wie viele Personen nur dann falschen Erinnerungen unterlagen, wenn sie eine gleichgeschlechtliche Person beob-

achtet hatten. Von den 13 TeilnehmerInnen der Bedingung Beobachten, die überhaupt Erinnerungsfehler aufwiesen, wiesen lediglich zwei Frauen falsche Erinnerungen ausschließlich an durch die Akteurin ausgeführte Handlungen auf. Die restlichen 11 TeilnehmerInnen wiesen entweder falsche Erinnerungen für Handlungen, die von beiden AkteurInnen ausgeführt wurden (7), bzw. für Handlungen, die nur vom gegengeschlechtlichen Akteur bzw. von der gegengeschlechtlichen Akteurin ausgeführt wurden (4), auf.

Ein Blick auf die fälschlich als ausgeführt erinnerten Handlungsanweisungen ergab zwar, dass die Verteilung über die verschiedenen Pools relativ ungleichmäßig aussah, allerdings ist hierbei zu berücksichtigen, dass diese auch ungleich häufig eingesetzt wurden. Die beiden am häufigsten fälschlicherweise als ausgeführt erinnerten Handlungsanweisungen waren mit jeweils acht Mal „Legen Sie die Seife in die Dose!“ und „Werfen Sie den Ball von der einen in die andere Hand!“. Insgesamt acht Handlungsanweisungen hingegen wurden niemals falsch erinnert (u.a. „Mischen Sie die Karten!“, „Spitzen Sie den Bleistift an!“, „Biegen Sie eine Büroklammer auseinander!“, „Betätigen Sie die Klingel!“). Es ließ sich allerdings aus meiner Sicht keine Systematik erkennen – weder existierte ein verbindendes Element unter den häufig falsch erinnerten, noch unter den nie falsch erinnerten Handlungen. Auch ergab sich keine Systematik in Abhängigkeit von den verschiedenen Arten der Verarbeitung.

Bisher wurde die Frequenz 1x bei der Betrachtung ausgeklammert, da die experimentelle Planung sich auf den Vergleich 5x-0x konzentriert hatte. Wird jedoch die Häufigkeit falscher *ausgeführt*-Urteile nach einmaliger Vorstellung bzw. Beobachtung einbezogen, zeigen sich Unterschiede zwischen beiden Bedingungen. Anhand der deskriptiven Statistiken wird bereits deutlich, dass die einmalige Verarbeitung nur in der Bedingung „Vorstellen“ zu einem Zuwachs an falschen Erinnerungen bzw. zu einem inflation-Effekt führte. Dieser lässt sich auch inferenzstatistisch absichern, wenn anstatt des Vergleichs 5x-0x, nun der Vergleich 1x-0x gezogen wird; $t(14)=3.87$, $p<.05$, $\eta^2=.517$. Auch der Vergleich 5x-1x wird nach Vorstellen signifikant; $t(14)=2.30$, $p<.05$, $\eta^2=.275$. Nach Beobachtung ergab sich hingegen kein signifikanter Effekt für den Vergleich 1x-0x ($t(14)=.91$, $p>.05$), jedoch für den Vergleich 5x-1x ($t(14)=3.73$, $p>.05$, $\eta^2=.498$). Dies bedeutet, dass bereits das einmalige Vorstellen die Erinnerung bedeutsam verfälschte, das fünfmalige Vorstellen jedoch eine noch stärkere Verfälschung hervorrief. Beim Beobachten unterscheiden sich einmalige und nullmalige Verarbeitung dagegen nicht voneinander; erst nach fünfmaligem Beobachten trat eine deutliche Gedächtnistäuschung auf. Somit ergibt sich auch insgesamt für das Vorstellen ein höheres Ausmaß falscher *ausgeführt*-Antworten im Vergleich zum Beobachten; $t(28)=2.25$, $p<.05$, $\eta^2=.153$. Deskriptivstatistisch zeigt sich nach Vorstellung ebenfalls ein größerer inflation-Effekt als nach Beobachtung; dieser Unterschied war jedoch nicht signifikant; $t(21.72)=1.60$, $p>.05$.

Nach jeder Präsentation innerhalb der zweiten Phase der ersten Sitzung wurden die TeilnehmerInnen u.a. gebeten, die Häufigkeit der Verarbeitung anzugeben. Es ergaben sich folgende Kennwerte ($M(SD)$): Vorstellen: 4.31 (2.02), Beobachten: 2.24 (.21), Generieren: 2.54 (1.54), Lesen: 5.24 (1.17). Trotz aller Bemühungen (siehe Abschnitt I.5.2) variierten die Verarbeitungsfrequenzen also stark, wobei zu berücksichtigen ist, dass beim Generieren lediglich die Häufigkeit des Lesens nach erfolgter Eingabe angegeben wurde. Allerdings ergab sich für die Gesamtstichprobe kein signifikanter Zusammenhang zwischen der Verarbeitungshäufigkeit und dem Ausmaß des inflation-Effekts; $r(60) = .039$, $p > .05$. Auch innerhalb der beiden Bedingungen Vorstellen und Beobachten, innerhalb derer überhaupt nur ein inflation-Effekt gefunden wurde, ergaben sich keine bedeutsamen Zusammenhänge (Vorstellen: $r(15) = .168$, $p > .05$; Beobachten: $r(15) = .125$, $p > .05$), so dass dieser Aspekt bei der Diskussion der Ergebnisse nicht berücksichtigt werden muss.

Zwischen der erfassten Kognitions-/ Intuitionsneigung (NFC/ FII), den Items des HZI, VIVIQ und FDS und dem Ausmaß des imagination- respektive observation-inflation-Effekts wurden innerhalb der einzelnen Bedingungen wie auch über die Bedingungen hinweg Korrelationen errechnet. Es ergaben sich keine systematischen und bedeutsamen Zusammenhänge.

I.5.4 Diskussion

I.5.4.1 Diskussion unter methodischen Gesichtspunkten

Anhand von Studie I konnte der bekannte imagination-inflation-Effekt repliziert werden. Darüber hinaus konnte erstmals gezeigt werden, dass falsche Erinnerungen an die Handlungsausführung auch durch die bloße wiederholte Beobachtung einer anderen Person ausgelöst werden können – ein Phänomen, das analog zum imagination-inflation- als observation-inflation-Effekt bezeichnet werden soll. Allein nach fünfmaligem Generieren und Lesen der Handlungsanweisungen fand sich keine erhöhte Quote von falschen *ausgeführt*-Antworten.

Bevor auf die theoretischen und praktischen Implikationen dieses Befundes eingegangen werden soll, sollen verschiedene Aspekte der Validität der Hypothesenprüfung beleuchtet werden.

Die theoretische Hypothese wurde adäquat operationalisiert, in Testhypothesen überführt und statistisch valide überprüft. Die Testhypothesen können allesamt als angenommen gelten, da die vorhergesagten Unterschiede statistisch nachgewiesen werden konnten bzw. sich keine Unterschiede ergaben, wenn diese auch nicht erwartet worden waren. Dabei sind die erwartete

ten Effekte nicht nur als signifikant ($p < .05$), sondern auch als substantiell ($\eta^2 \geq .14$) zu bezeichnen.

Bei der Operationalisierung der Konstrukte und der Untersuchungsplanung wurde besonderes Augenmerk auf eine strenge, valide und präzise Hypothesenprüfung gelegt. Dennoch existieren Aspekte, die als problematisch erachtet werden könnten und daher im Folgenden beleuchtet werden sollen.

Eine Grundvoraussetzung für die Interpretierbarkeit der Ergebnisse ist, dass die TeilnehmerInnen, die autonom arbeiteten, in der ersten Phase auch tatsächlich wie instruiert vorgehen, d.h. Handlungen ausführten, wenn sie dazu aufgefordert wurden, bzw. nicht ausführten, wenn sie instruiert wurden, die Handlungsanweisungen nur zu lesen. Da die TeilnehmerInnen während der Untersuchungsdurchführung nicht beobachtet wurden, wurde die Validität der Daten anhand von vier Indikatoren festgemacht. Dabei bestätigten alle Indikatoren deutlich, dass die Handlungsanweisungen im Rahmen der ersten Phase auf die gewünschte Art und Weise gehandhabt wurden. Darüber hinaus zeigten weitere Parameter, dass auch in der zweiten Phase die Instruktionen wie gewünscht befolgt wurden. Insgesamt kann eine hohe compliance und daher Validität der Daten angenommen werden.

Aufgrund der beschriebenen Maßnahmen zur Kontrolle von allgemeinen und speziellen Störeffekten kann des Weiteren davon ausgegangen werden, dass die interne Validität gesichert war. Eine Einschränkung der internen Validität kann allerdings auch durch mögliche Materialeffekte hervorgerufen werden. Solche hätten in der vorliegenden Untersuchung auftreten können, da innerhalb sowie zwischen den jeweiligen Gruppen nicht jede Materialkombination gleich häufig eingesetzt werden konnte. Diese Möglichkeit konnte jedoch sowohl argumentativ als auch empirisch ausgeräumt werden.

Ein weiterer möglicher Kritikpunkt könnte in der Tatsache gesehen werden, dass die Verarbeitungshäufigkeiten der unterschiedlichen Prozesse voneinander divergierten. So wurden bereits in der ersten Phase Handlungsanweisungen häufiger gelesen als ausgeführt. In der zweiten Phase lag z.B. die durchschnittliche Häufigkeit des Vorstellens über derjenigen des Beobachtens und die des Lesens über derjenigen des Generierens. Diese Unterschiedlichkeiten sind darin begründet, dass im Rahmen des ersten Experiments – in Anlehnung an alle Vorgängerstudien – die Zeit und eben nicht die Häufigkeit der Verarbeitung für alle Prozesse konstant gehalten wurde. Das umgekehrte Vorgehen – nämlich die Häufigkeit und nicht die Zeit der Verarbeitung gleich zu halten (also z.B. jede Handlung genau einmal ausführen oder vorstellen zu lassen usw.), hätte umgekehrt zur Konsequenz gehabt, dass die Dauer weder innerhalb

einer Form der Verarbeitung noch zwischen verschiedenen Formen der Verarbeitung hätte konstant gehalten werden können, da die Ausführung einiger Handlungen schlicht länger dauert als diejenige anderer bzw. das Ausführen länger dauert als das Lesen usw. Daher wird es – egal, welchen Weg man wählt – immer einen „Pferdefuß“ bei der Umsetzung geben. Wichtig ist aber, dass kein Zusammenhang zwischen der Häufigkeit der Verarbeitung und dem Ausmaß falscher Erinnerungen zu finden war, so dass davon auszugehen ist, dass keine Konfundierungen vorliegen. Selbstverständlich unterscheiden sich die vier Prozesse nicht nur in der erwähnten Verarbeitungshäufigkeit stark voneinander, sondern sie sind genuin äußerst unterschiedlich – was natürlich einerseits intendiert und der Fragestellung inhärent ist, andererseits den Vergleich der Prozesse problematisch erscheinen lassen könnte. Explorative Analysen ergaben jedoch, dass sich alle vier Verarbeitungsformen hinsichtlich korrekter Erinnerungsleistungen im Rekognitions- und Quellengedächtnistest nicht grundlegend unterscheiden (keine signifikanten Haupteffekte bzgl. der verschiedenen Itemarten). Demnach kann der Vergleich der Prozesse als legitim angesehen werden.

Da auf bewährte Operationalisierungen zurückgegriffen wurde bzw. diese noch optimiert wurden, ist die Konstruktvalidität als gegeben zu erachten. Es ist allerdings zu berücksichtigen, dass nur eine Form des Generierens – nämlich das Lösen umgestellter Sätze – sowie eine Form des Beobachtens – nämlich die Beobachtung per Videodarbietung – realisiert werden konnten. Ob gleiche Ergebnisse bei anderen Operationalisierungen erzielt worden wären, bleibt spekulativ. Es ist allerdings davon auszugehen, dass die gewählten Operationalisierungen gleichzeitig die für die vorliegende Untersuchung geeignetsten darstellten sowie die strengste und präziseste Hypothesenprüfung erlaubten.

Die zentrale abhängige Variable stellt im vorliegenden Paradigma der Anteil an falschen *ausgeführt*-Antworten dar. Da ein solches Quellengedächtnisurteil nur nach erfolgter Rekognition abgegeben wird, bestünde eine Möglichkeit der Auswertung darin, die falschen *ausgeführt*-Antworten an der Anzahl rekognizierter Items zu relativieren. Bei diesem Vorgehen würden jedoch die Zellbesetzungen stark variieren. So erinnerten z.B. nur wenige TeilnehmerInnen fälschlicherweise, dass null- oder einmalig in Phase 2 präsentierte Items aus Phase 1 stammten. Damit würden *ausgeführt*-Antworten auf Items, die nur in sehr geringem Ausmaß der Phase 1 zugeordnet wurden, ein überproportionales Gewicht erhalten. Würde bspw. eine Teilnehmerin lediglich zwei Items, die gar nicht in Phase 1, aber einmal in Phase 2 präsentiert wurden, als aus Phase 1 stammend rekognizieren, davon eins als *ausgeführt* und eins als nicht *ausgeführt* einstufen, ginge die *ausgeführt*-Antwort mit einer Quote von 50% in die Berechnung ein. Da dies wenig sinnvoll erschien, wurden alle Erinnerungsurteile an der gültigen Anzahl an Items pro Zelle relativiert, also im Normalfall an der Zahl 5; wenn die TeilnehmerInnen

jedoch ungültige Items produziert hatten, an einer entsprechend geringeren Anzahl¹⁰. Dies entspricht letztlich dem Modell dreier Antwortmöglichkeiten: neu, alt und ausgeführt bzw. alt und nicht ausgeführt, d.h. die entsprechenden relativen Häufigkeiten dieser drei Antwortoptionen addieren sich zu eins (siehe auch Thomas et al., 2002, 2003).

Die Untersuchung wurde an einer sehr selektiven Stichprobe, nämlich an Psychologie-Studierenden im Grundstudium überprüft. Allerdings ist im Hinblick auf die zugrundeliegenden Gedächtnisprozesse, die sich auf ganz alltägliches Material beziehen, nicht davon auszugehen, dass sich Psychologiestudierende und andere junge Erwachsene voneinander unterscheiden könnten. Unterschiede mögen hinsichtlich der untersuchten Prozesse eher mit dem Alter zusammenhängen. So ist bspw. anzunehmen, dass der Effekt bei älteren TeilnehmerInnen noch ausgeprägter auftreten würde (siehe auch Abschnitt II); Stichproben von Kindern wurden bis dato noch nicht anhand des imagination-inflation-Paradigmas untersucht. Auch ist zu bedenken, dass nur eine Art von Handlungen – einfache, alltägliche Handlungen – untersucht wurde. Die Ergebnisse sind nicht auf andere, komplexere Handlungen oder Ereignisse übertragbar, denn es ist davon auszugehen, dass sich das Spektrum und die Gewichtung von Erinnerungsspuren, die als diagnostisch für die Handlungsausführung angesehen werden, je nach Komplexität der Handlung unterscheiden. Allerdings wurde die Art des zu erinnernden Materials bereits in der Hypothese spezifiziert, so dass diesbzgl. keine zusätzlichen Einschränkungen getroffen werden müssen. Die Frage, inwiefern die Ergebnisse auf andere Situationen bzw. in den Alltag übertragen werden können, soll im nächsten Abschnitt im Zusammenhang mit der praktischen Bedeutsamkeit der Ergebnisse diskutiert werden. Insgesamt kann also davon ausgegangen werden, dass die interne sowie die Ableitungsvalidität auf den verschiedenen Ebenen der Untersuchung gegeben war. Da alle Testhypothesen statistisch nachgewiesen werden konnten, kann die Hypothese

Falsche Erinnerungen an die Ausführung einfacher Handlungen basieren zumeist auf einer Fehlattribution von sensorischen und/ oder motorischen Repräsentationen, weshalb analog zum imagination-inflation- ein observation-inflation-Effekt existiert. Sie werden zumeist nicht durch Vertrautheit oder den (internalen) Ursprung der Information hervorgerufen.

als vorläufig bewährt gelten. Dabei bestehen bzgl. der externen Validität die üblichen Einschränkungen eines Laborexperiments, das an einer angefallenen Stichprobe durchgeführt

¹⁰ Ungültige Items sind solche, die von der Analyse ausgeschlossen wurden, da die instruierte Handhabung in der ersten Phase (ausführen – lesen) nicht mit der von den TeilnehmerInnen angegebenen Handhabung übereinstimmte; vgl. Abschnitt 1.5.3.2.

wurde. Die Übertragbarkeit auf andere Situationen, insb. jedoch die theoretischen und praktischen Implikationen des Befundes, werden im Folgenden diskutiert.

I.5.4.2 Diskussion unter theoretischen und praktischen Gesichtspunkten

Das gefundene stimmt perfekt mit dem vorhergesagten Ergebnismuster überein. Auf theoretischer Ebene lässt sich daher festhalten, dass falsche Erinnerungen an die eigene Handlungsausführung nicht durch bloße Vertrautheit bzw. durch die Verarbeitungsleichtigkeit der Handlungsanweisungen und auch nicht durch den bloßen internalen Informationsursprung zustande kommen. Sie scheinen auf einer Fehlattribution von sensorischen und/ oder motorischen Merkmalen zu basieren; eine Unterscheidung zwischen diesen beiden Mechanismen erlaubt die vorliegende Untersuchung allerdings nicht. Dieses Ergebnismuster kommt nicht nur im Ausmaß falscher *ausgeführt*-Antworten zum Ausdruck, sondern spiegelt sich auch in der Anzahl an Personen wider, die einen solchen inflation-Effekt aufwiesen.

Der bemerkenswerteste Befund der Untersuchung besteht zweifelsohne darin, dass zum ersten Mal ein observation-inflation-Effekt gezeigt werden konnte: Falsche Erinnerungen daran, eine Handlung ausgeführt zu haben, können durch die bloße Beobachtung einer anderen Person hervorgerufen werden. Bis dato wurde dieser Effekt ausschließlich durch Vorstellung evoziert. Bemerkenswerterweise ähneln sich Beobachten und Vorstellen jedoch nicht nur hinsichtlich falscher *ausgeführt*-Antworten, auch hinsichtlich korrekter Antworten im Quellengedächtnistest sowie hinsichtlich der Rekognitionsleistung führen beide Prozesse zu auffällig ähnlichen Ergebnismustern. Zudem scheinen die durch beide Formen der Verarbeitung hervorgerufenen falschen Erinnerungen von ähnlicher Qualität zu sein – jedenfalls werden sie mit vergleichbar hoher Sicherheit beurteilt. Ein augenfälliger Unterschied besteht darin, dass bereits nach einmaliger Vorstellung, jedoch erst nach fünfmaliger Beobachtung ein Effekt auftritt und das Ausmaß falscher Erinnerungen für die Vorstellung insgesamt etwas höher liegt als für die Beobachtung (, wenn auch kein signifikant größerer inflation-Effekt zu verzeichnen war). Ein Unterschied zwischen beiden Arten der Verarbeitung zeigte sich auch hinsichtlich der verschiedenen Itemarten; so wurden nach Vorstellung signifikant mehr Items fälschlich als ausgeführt erinnert, die bereits in Phase 1 präsentiert worden waren, verglichen mit solchen Items, die erst in der zweiten Phase dargeboten wurden. Eine derartige Tendenz fand sich auch für das Beobachten, ließ sich jedoch nicht statistisch absichern. Dies mag allerdings – anstatt einen grundlegenden Unterschied beider Prozesse darzustellen – eher darauf zurückzuführen sein, dass der kleinere Effekt nach Beobachten bei dem gegebenen Stichprobenumfang nicht nachgewiesen werden konnte. Die ausgeprägte Ähnlichkeit beider Ergebnismuster könnte dafür sprechen, dass sowohl durch Vorstellung als auch Beobachtung ähnliche Prozesse angestoßen werden – im Falle des

werden – im Falle des Vorstellens eine bewusst initiierte, internale, kontrollierte Form der sensorischen und motorischen Simulation der eigenen Handlungsausführung, im Falle des Beobachtens eine unbewusst initiierte, automatisch ablaufende Form der motorischen Simulation bzw. eine Anreicherung der Handlungsanweisungen mit sensorischen Informationen durch die Wahrnehmung external dargebotener Stimuli. Dennoch – wie lassen sich die (leichten) Unterschiede zwischen beiden Ergebnismustern interpretieren? Eine Dimension des Unterschiedes zwischen beiden Prozessen stellt die Verarbeitungstiefe dar, die aufgrund des internalen Ursprungs und des kontrollierten Prozesses nach Vorstellung größer sein sollte als nach Beobachtung. Es wäre denkbar, dass Verarbeitungstiefe und die Anreicherung mit sensorischen/motorischen Informationen interagieren – wenn auch Verarbeitungstiefe bzw. die aus ihr resultierende Verarbeitungsleichtigkeit allein keine falschen Erinnerungen hervorrief, wie das Ergebnismuster nach Lesen bzw. Generieren nahe legt (siehe auch Thomas et al., 2003). Aber nicht nur die Tiefe der Verarbeitung, sondern auch die Tiefe der motorischen Simulation könnte sich unterscheiden, d.h. stärker nach Vorstellung und schwächer nach Beobachtung sein. Einen ersten empirischen Anhaltspunkt hierfür liefern Léonard und Tremblay (2007; siehe auch Clark, Tremblay & Ste-Marie, 2003); weitere Studien wären jedoch vonnöten, um diese Annahme zu untermauern. Andererseits unterscheiden sich Vorstellung und Beobachtung (möglicherweise) nicht nur in der Stärke oder Tiefe der motorischen Simulation, sondern auch in der Art bzw. Qualität der perzeptuellen Simulation – die Vorstellung könnte der Ausführung ähnlicher sein, da sie es erlaubt, alle Sinne (z.B. auch taktile Eindrücke) einzubeziehen. Allerdings kann wohl keine Vorstellung an Klarheit an die Beobachtung bzw. tatsächliche Wahrnehmung heranreichen, die diesbzgl. der Ausführung näher kommt – wobei einschränkend zu beachten ist, dass die Perspektive bei der Beobachtung der fremden vs. Ausführung oder Vorstellung der eigenen Handlungsausführung eine andere ist. Vielleicht könnte also die größere Schwierigkeit, Vorstellung und Ausführung zu trennen, auch durch die Tatsache hervorgerufen worden sein, dass Vorstellung in verschiedener Hinsicht (Stärke und Spektrum) eine wahrnehmungsnähere Simulation darstellt als Beobachtung. Welcher der beiden Prozesse – Vorstellen oder Beobachten – der Ausführung ähnlicher ist, ist aber vor dem gegebenen Hintergrund letztlich nicht zu beurteilen; diese Frage könnte jedoch in zukünftigen Untersuchungen verfolgt werden.

Generieren und Lesen evozierten insgesamt fast gar keine falschen *ausgeführt*-Antworten; demnach scheinen beide Prozesse sehr distinkt zum Prozess der Handlungsausführung zu sein (siehe auch Thomas et al., 2003), obwohl ebenfalls Vertrautheit mit den Handlungsanweisungen sowie (im Falle des Generierens) ein internaler Informationsursprung erzeugt wurden. Bemerkenswert ist jedoch, dass beim Generieren eine – von der Frequenz der Präsentation nicht beeinflusste – höhere Rate an falschen Erinnerungen für in der ersten Phase präsentier-

te, jedoch nicht ausgeführte, im Vergleich zu in der ersten Phase nicht enkodierten Handlungsanweisungen auftrat. Demnach wurde in dieser Bedingung schon die bloße Tatsache, dass eine Handlungsanweisung in der ersten Phase überhaupt präsentiert worden war, als Hinweis auf die Handlungsausführung gewertet. Die TeilnehmerInnen schienen also qualitative Erinnerungsspuren an den Gegenstand/ die Gegenstände, die ja für die in Phase 1 präsentierten im Vergleich zu den nur in Phase 2 präsentierten Handlungsanweisungen vorlagen, als partiell diagnostisch für die Ausführung anzusehen. Generieren und Lesen unterscheiden sich am deutlichsten hinsichtlich der Rekognition von Handlungsanweisungen, die in Phase 1 gar nicht enkodiert wurden. Nach mehrfachem Generieren wurden deutlich mehr Handlungsanweisungen fälschlich als aus Phase 1 stammend rekogniziert, die tatsächlich erst in Phase 2 präsentiert worden waren – vermutlich ein Effekt der Verarbeitungstiefe bzw. Verarbeitungseichtigkeit der mehrfach generierten Handlungsanweisungen.

Abschließend soll das Ergebnismuster in seinem Gesamt noch einmal kritisch betrachtet und gefragt werden, ob zu der gewählten Interpretationsmöglichkeit Alternativen existieren, die sich nicht auf bereits diskutierte Validitätsaspekte beziehen. Dabei besteht die gewählte Interpretationsmöglichkeit in der Behauptung, die zusätzlichen sensorischen und/ oder motorischen Informationen, die die Vorstellung und das Beobachten vom Generieren und Lesen unterscheiden, seien kritisch für das Zustandekommen von falschen Erinnerungen an die eigene Handlungsausführung. Zwei Befunde könnten gegen diese Interpretation sprechen. Einerseits ließe sich argumentieren, dass der Anstieg falscher *ausgeführt*-Antworten lediglich dem Anstieg der *Phase 1*-Antworten geschuldet sei. So steigen die Quoten rekognizierter Handlungsanweisungen, die in der ersten Phase nicht ausgeführt wurden, nach Beobachtung und Vorstellung mit zunehmender Frequenz der Darbietung an – wie auch die Quoten falscher *ausgeführt*-Antworten auf diese Handlungsanweisungen. Dieser Interpretation lässt sich allerdings auf verschiedenen Ebenen widersprechen. Auf theoretischer Ebene haben bereits Glisky und KollegInnen (1995) nachgewiesen, dass Rekognitions- und Quellengedächtnis dissoziiert sein können, also keine Abhängigkeit bestehen muss (siehe auch Schacter, Harbluk & McLachlan, 1984; aber siehe Murnane & Bayen, 1996, 1998; Näheres hierzu in Abschnitt II.1.2). Auf Ebene der Daten ist der Interpretation entgegenzuhalten, dass dieses Ergebnismuster nicht konsistent auftritt. Auch in den Bedingungen Generieren und Lesen gibt es einen der Frequenz folgenden Anstieg in der Rekognition nicht ausgeführter Handlungsanweisungen, der in der Stärke sogar teilweise demjenigen in den Bedingungen Vorstellen und Beobachten entspricht – dennoch ergibt sich nach Generieren und Lesen keine erhöhte Quote von falschen *ausgeführt*-Antworten. Zudem entspricht das Verhältnis des Anstiegs von Rekognitions- und Quellengedächtnisurteilen einander nicht; während z.B. die Quote falscher *ausgeführt*-Antworten nach mehrfacher Vorstellung von in der ersten Phase gelesenen Handlungsanweisungen um das

fast siebenfache ansteigt (von .05 nach nullmaliger auf .33 nach fünfmaliger Präsentation), ergibt sich eine deutlich geringere Steigerungsrate bei der Rekognition, obwohl das Maximum nicht erreicht wurde (von .60 nach nullmaliger auf .84 nach fünfmaliger Präsentation). Demnach scheidet die Möglichkeit aus, dass der Zuwachs der falschen Quellengedächtnisurteile allein eine Konsequenz des Zuwachses der Rekognitionsurteile darstellt. Zusätzlich spricht gegen die Interpretationsmöglichkeit, dass der Anteil an *ausgeführt*-Antworten nicht an der Anzahl rekognizierter Items, sondern an der Anzahl an (gültigen) Items pro Zelle relativiert wurde; also auch auf dieser Ebene keine Abhängigkeit besteht. Insgesamt bleibt allerdings festzustellen, dass es sich bei der eingesetzten Rekognitionsaufgabe, nämlich zu bestimmen, ob ein Item aus Phase 1 stammt oder nicht, nicht um eine klassische Rekognitionsaufgabe im Sinne einer alt/neu-Unterscheidung handelt, sondern – wie auch Goff und Roediger (1998) kritisch bemerken – bereits um eine Unterscheidung von zwei (zeitlichen) Quellen, nämlich Phase 1 und 2. Dies erschwert insgesamt die Interpretation, v.a. da nicht sichergestellt werden kann, ob die TeilnehmerInnen diese Unterscheidung vollständig verstanden und im gewünschten Sinne umgesetzt haben.

Weitaus kritischerer hinsichtlich der Ergebnisinterpretation stellt sich jedoch die Tatsache dar, dass sowohl nach Vorstellung als auch Beobachtung, nicht jedoch nach Generieren und Lesen, ein der Frequenz folgender Anstieg auch der korrekten *ausgeführt*-Antworten existiert. Folglich zeigt sich ein inflation-Effekt nach Vorstellen und Beobachten hinsichtlich aller Itemarten. Während also die Quote von *ausgeführt*-Antworten im Vergleich der Frequenzen 5x und 0x nach Generieren und Lesen stets konstant bleibt, ist nach Vorstellen und Beobachten stets ein Anstieg nach fünfmaliger im Vergleich zu nullmaliger Präsentation zu verzeichnen. Demnach könnte es sich bei dem beschriebenen Phänomen gar nicht um eine Erinnerungstäuschung sondern vielmehr um eine Tendenz handeln, bei wiederholter Präsentation eher ausgeführt zu antworten („Quellen-Bias“; siehe auch Murnane & Bayen 1996). Dies würde allerdings die gesamte Interpretation des inflation-Effekts als Gedächtnisphänomen in Frage stellen. Alternativ könnte das Ergebnismuster aber auch dadurch zustande kommen, dass (wiederholte) Vorstellung und Beobachtung bestimmte Erinnerungsqualitäten stärken, die als diagnostisch für die Handlungsausführung angesehen werden; mutmaßlich sensorische und/ oder motorische Repräsentationen.

Bzgl. der praktischen Bedeutsamkeit der Ergebnisse lässt sich festhalten, dass es sich zwar bei der hergestellten Laborsituation um eine sehr künstliche Gegebenheit handelt, jedoch bei den verwendeten Materialien um Anweisungen zu ganz alltäglichen Handlungen. Und auch die erfassten Prozesse des Ausführens, Vorstellens und Beobachtens dieser Handlungen kommen im Alltag vor; wenn auch sicher nicht in der extremen Geballtheit aufgrund der mehrfachen

Wiederholungen und natürlich nicht unter den kontrollierten Situationsbedingungen im Labor. Das Lesen und Generieren derartiger Handlungsanweisungen hingegen ist weniger alltäglich, könnte aber möglicherweise der bloßen Aktivierung oder der Bewusstwerdung einer Handlung(sabsicht) (Lesen) bzw. dem Nachdenken über eine Handlung (Generieren) ähnlich sein. Daher ist denkbar, dass mit den verwendeten Operationalisierungen Prozesse approximiert werden, die uns allen aus unserem Alltag sehr bekannt sind: Jede(r) hat schon einmal die Situation erlebt, sich nicht sicher zu sein, ob er oder sie eine einfache Handlung (z.B. das Abschalten des Herdes) tatsächlich ausgeführt hat oder nicht. Wenn man sich die Handlungsausführung nun vorstellt bzw. vorgestellt hat oder eine andere Person bei der Handlung beobachtet bzw. beobachtet hat, könnte man den Ergebnissen der Studie zufolge (richtiger- oder falscherweise) eher geneigt sein, anzunehmen, dass die Handlung tatsächlich umgesetzt wurde. Eine Verwechslung ist in ähnlicher Form sicherlich auch bei einer früheren eigenen Ausführung möglich. Wenn man sich die Handlung bloß „ins Bewusstsein ruft“ oder über die Handlung nachdenkt bzw. dies getan hat, sollte dies weniger der Fall sein. Im Alltag werden allerdings noch viele andere Aspekte eine Rolle spielen, die in der Laborsituation keine Rolle spielten bzw. konstant gehalten wurden und somit die Situationsvalidität einschränken. Hierzu zählen insb. die Aufmerksamkeit, die man der Handlung schenkte, das Intervall zwischen Enkodierung und Abruf, Aspekte der Plausibilität und Gewohnheit und – wie bereits ausgeführt – die Komplexität der Handlung. Insgesamt ist jedoch festzuhalten, dass es sich bei den erfassten Gedächtnisprozessen durchaus um solche handelt, die für unseren Alltag von zentraler Relevanz sind. Fälschlicherweise zu erinnern, den Herd ausgestellt, die Tabletten genommen oder die Tür zugezogen zu haben, kann gravierende Konsequenzen nach sich ziehen. Daher ist die Erforschung des Phänomens nicht nur grundlagenwissenschaftlich relevant – sollte der Mechanismus eingegrenzt sein, könnte hierauf aufbauend überlegt werden, wie sich die Leistung steigern bzw. wie sich falsche Erinnerungen reduzieren lassen. Dies könnte insbesondere für ausgewählte Populationen, z.B. ältere gesunde, aber auch kognitiv beeinträchtigte Menschen von Relevanz sein, die bekanntermaßen Einschränkungen insbesondere des episodischen Gedächtnisses und daher auch der Quellendiskrimination aufweisen (vgl. Abschnitt II). Wollte man diese jedoch anhand des eingesetzten Paradigmas untersuchen, müsste dieses zunächst hinsichtlich einiger Aspekte (z.B. hinsichtlich des autonomen Arbeitens) modifiziert und insgesamt verschlankt werden, da es in der jetzigen Form schon für jüngere, gesunde TeilnehmerInnen eine starke Beanspruchung darstellt.

I.5.4.3 Résumé und Ausblick

Hinsichtlich des Zustandekommens von falschen Erinnerungen an die Handlungsausführung scheiden die bloße Vertrautheit bzw. der Informationsursprung als Ursachen aus; ein Ergebnis, das bestehende Befunde bestätigt und erweitert (z.B. Thomas et al., 2003). Die Anreicherung der Handlungsanweisungen mit sensorischen und/ oder motorischen Informationen schien kritisch für das Zustandekommen der falschen Erinnerungen. Daher ließ sich nicht nur der bekannte imagination-inflation-Effekt replizieren, sondern erstmals auch ein observation-inflation-Effekt nachweisen. Eine Unterscheidung der beiden postulierten Mechanismen (Fehlattri-bution sensorischer vs. motorischer Informationen) konnte durch die Untersuchung allerdings nicht geleistet werden und steht daher aus. Da jedoch aufgrund des Datenmusters nicht ausgeschlossen werden kann, dass es sich bei den vermeintlichen falschen Erinnerungen nicht bloß um Antworttendenzen bzw. einen Quellen-Bias handelt, soll Gegenstand der zweiten Studie zunächst die Manipulation des Antwortkriteriums sein. Da des Weiteren die Interpretierbarkeit der Datenstruktur aufgrund der Tatsache eingeschränkt ist, dass im Grunde zwei Quellengedächtnistests eingefordert werden, zudem die Prozedur insgesamt, insb. jedoch die zweite Phase, als anstrengend empfunden wurde und als mit weniger belastbaren Populationen nicht einsetzbar angesehen wird, soll ebenfalls die Verschlinkung der Prozedur Gegenstand der zweiten Untersuchung sein.

Der sicherlich zentralste Befund der ersten Studie ist insgesamt darin zu sehen, dass Vorstellen und Beobachten zu erstaunlich ähnlichen (korrekten wie falschen) Erinnerungsleistungen führten. Dies eröffnet die Möglichkeit, Vorstellungen – also für einen Außenstehenden nicht zugängliche Prozesse – durch Beobachtungen zu ersetzen, bei denen die zu verarbeitenden Stimuli vollständig kontrollierbar sind. Dies stellt besonders dann einen nicht zu unterschätzenden Vorteil dar, wenn mit Populationen gearbeitet werden soll, die kognitive Beeinträchtigungen aufweisen und bei denen von daher nicht sicherzustellen ist, ob die Instruktion zum Vorstellen überhaupt verstanden und umgesetzt werden kann. Daneben führt der Einsatz von Beobachtung zu einer besseren Vergleichbarkeit zwischen den ProbandInnen, da interindividuelle Unterschiede in der Vorstellungsfähigkeit sich nicht auswirken können.

Insgesamt soll also die zweite Studie auf die Beobachtung fokussieren und zwei Ziele verfolgen: Einerseits soll die Prozedur verschlinkt werden, andererseits soll der observation-inflation-Effekt unter Bedingungen strikterer Antwortkriterien repliziert werden.

Die Induktion strikterer Antwortkriterien bei der Abfrage ist in der Gedächtnispsychologie auf verschiedenen Wegen erfolgt, bspw. durch Instruktionen, die spezifizieren, auf welche Erinne-

rungsspuren die Gedächtnisurteile basiert werden sollen, Warnungen vor möglichen Gedächtnistäuschungen oder auch durch eine Änderung des Abfrageformates (z.B. Thomas & Bulevich, 2006; Lindsay & Johnson, 1989; McDermott & Roediger, 1998; Zaragoza & Lane, 1994). Im Rahmen der zweiten Studie sollen zwei Varianten an Instruktionen zum Einsatz kommen, eine Instruktion zur Quellenprüfung und eine Warnung, deren genaue Ausgestaltung und Herleitung Gegenstand des Abschnitts I.6.2.1 ist. An dieser Stelle soll jedoch kurz darauf eingegangen werden, wie die Wirkung solcher Manipulationen theoretisch aussehen kann. Dabei wird auf Warnungen fokussiert, die nach erfolgter Enkodierung gegeben werden. Der alternative Fall, die Warnung vor der Enkodierung auszusprechen, ist bereits hinsichtlich mehrerer Gedächtnistäuschungen eingesetzt worden (z.B. Gallo, Roediger & McDermott, 2001); da aber eben nicht eine Manipulation der Enkodier-, sondern der Abrufprozesse intendiert war, wurde von dieser Möglichkeit abgesehen (siehe z.B. auch Echterhoff, Hirst & Hussy, 2005).

Dem source memory framework zufolge stellt ein Erinnerungsurteil das Resultat zweier Komponenten dar – der verfügbaren Erinnerungsspuren einerseits und der Bewertung dieser Erinnerungsspuren andererseits. Angestrebt ist im Rahmen der Studie II eine Manipulation des letzten Aspekts – nämlich der Evaluation der Gedächtnisspuren. Sowohl eine Instruktion zur Quellenprüfung als auch eine Warnung sollten Personen, die den Gedächtnistest durchlaufen, dafür sensibilisieren, verstärkt auf bestimmte für die Handlungsausführung diagnostische Kriterien zu achten und nicht leichtfertig bzw. rein intuitiv zu urteilen (z.B. Chambers & Zaragoza, 2001; Echterhoff, Hirst & Hussy, 2005; Thomas & Bulevich, 2006). Demnach sollte bspw. ein bloßes Vertrautheitsgefühl nicht als ausreichende Grundlage erachtet werden, um zu urteilen, dass eine Handlung tatsächlich ausgeführt wurde. Qualitative Gedächtnisspuren, die auf eine Ausführung hinweisen, müssen notwendig vorhanden sein, wenn die Person ein *ausgeführt*-Urteil fällt. Verschwände der Effekt unter der Bedingung einer genaueren Überprüfung und Evaluation der Erinnerungsspuren, würde es sich somit nicht um eine echte Gedächtnistäuschung, sondern um eine in der Art der Bewertung der Erinnerungsspuren begründete Täuschung bzw. um einen Effekt des Antwortkriteriums handeln. Blicke der Effekt jedoch bestehen, wenn die Person ein striktes Antwortkriterium anwendet, so würde dies bedeuten, dass die Erinnerungsspuren beider Quellen so ähnlich sind, dass Gedächtnistäuschungen nahezu unausweichlich sind. Beide Varianten sind in der gedächtnispsychologischen Forschung bereits aufgetreten. So konnte bspw. der Falschinformationseffekt durch Warnungen reduziert werden, falsche Erinnerungen im DRM-Paradigma allerdings ließen sich so nicht (immer) vermindern (z.B. Blank, 1998; Chambers & Zaragoza, 2001; Greene, Flynn & Loftus, 1982; McDermott & Roediger, 1998; Gallo, Roediger & McDermott, 2001).

Es existiert nur eine einzige Studie zum imagination-inflation-Paradigma für einfache Handlungen, innerhalb derer der Abruf manipuliert wurde (Thomas & Bulevich, 2006). In dieser Untersuchung wurde eine Instruktion zur Quellenprüfung eingesetzt, die die TeilnehmerInnen dazu aufforderte, auf Hinweise für die tatsächliche Ausführung (vs. bloße Vorstellung) der Handlungen zu fokussieren. Thomas und Bulevich (2006) baten ihre TeilnehmerInnen, ihre Erinnerungsurteile auf für die Handlungsausführung bzw. Vorstellung diagnostische Hinweise zu basieren:

Research has determined that accurate memory judgments are made when one tries to remember not only the event or action but also the context in which the action was presented. Memory for perceptual and contextual information is a useful indicator of whether an action was performed, imagined, or never presented during the first session of the experiment. For example, if you remember performing an action, you may also remember how an object felt in your hand, or how something looked or smelled. You may remember the color of an object, or even how you felt performing that action. If an action was imagined in Session 1 you may have different contextual information associated with that memory of imagination. For example, you may remember closing your eyes or the sound of the experimenter's voice when the action was read. You may even remember the action that preceded or followed. If an action was never presented during Session 1, then the cues associated with memory for those actions would have been generated solely in Session 2 (Thomas & Bulevich, 2006, S. 385).

Diese Fokussierung auf Hinweisreize für die Handlungsausführung respektive Vorstellung hatte jedoch bei den jungen TeilnehmerInnen keinerlei Einfluss auf das Ausmaß falscher Erinnerungen an bloß vorgestellte Handlungen (siehe Abschnitt II.1.4). Aufgrund der gefundenen, ausgeprägten Ähnlichkeit von Vorstellung und Beobachtung und mangels weiterer empirischer Evidenz wird dieser Befund auf den observation-inflation-Effekt übertragen. Daher wird folgende Hypothese aufgestellt, die im Rahmen der zweiten Studie überprüft werden soll:

Beim observation-inflation-Effekt handelt es sich zumeist nicht um das Resultat einer Antworttendenz bzw. eines laxen Antwortkriteriums, sondern um eine a posteriori kaum vermeidliche Gedächtnistäuschung.

I.6 Experiment II

I.6.1 Voruntersuchung

I.6.1.1 Zielsetzung der Voruntersuchung

Neben der zentralen inhaltlich-theoretischen Zielsetzung – den Einfluss von Abrufmanipulationen auf den observation-inflation-Effekt zu erfassen – wurde im Rahmen des zweiten Experiments auch die Zielsetzung verfolgt, das experimentelle Design eindeutiger interpretierbar zu gestalten und so zu verkürzen, dass es auch zur Untersuchung älterer gesunder (evtl. auch dementer) TeilnehmerInnen eingesetzt werden könnte.

Um Letzteres zu gewährleisten kamen mehrere Variationen in Betracht – z.B. eine Kürzung der Anzahl an Items pro Zelle, die Realisierung anderer Frequenzstufen usw. Nach Abwägung aller Alternativen fiel die Entscheidung letztlich auf die Kürzung um eine Itemart – und zwar um diejenigen Items, die nicht in der ersten, sondern erst in der zweiten Phase präsentiert wurden (in der Terminologie des Versuchsplans also Items, die in Phase 1 nicht enkodiert, aber in Phase 2 fünf- bzw. einmal präsentiert wurden). Der entscheidende Grund hierfür war, dass das Ergebnismuster dieser Items demjenigen der in der ersten Phase präsentierten, aber nur gelesenen Items, entsprach, jedoch weniger stark ausgeprägt war. Diese schwächere Duplikation des Ergebnismusters schien verzichtbar, zumal es nicht nur den Umfang der zweiten Phase um ein Drittel reduzierte, sondern gleichzeitig die Eindeutigkeit und damit Interpretierbarkeit des Designs erhöhte. Dadurch, dass die Einführung von neuen Items in Phase 2 entfiel, konnte nämlich beim Gedächtnistest ebenfalls auf diese Unterscheidung bzw. das Rekurrieren auf eine bestimmte Phase verzichtet werden. D.h. an die Stelle der Frage, ob ein Item innerhalb der *ersten Phase* der ersten Sitzung präsentiert worden sei, konnte die Frage treten, ob ein Item innerhalb der ersten Sitzung *überhaupt* präsentiert worden bzw. ob es alt oder neu sei. Diese einfache alt-neu-Diskrimination sollte für die TeilnehmerInnen leichter verständlich sein und zu einem eindeutiger interpretierbaren Ergebnismuster führen als die Unterscheidung der Phasen. Außerdem konnte durch die Reduzierung um eine Itemart an der bewährten Anzahl von Items pro Zelle (5) festgehalten werden.

Ein weiteres Ziel lag in der Neugestaltung der ersten Phase, in der die TeilnehmerInnen zum Zwecke der besseren Kontrollierbarkeit nicht mehr – wie bisher – autonom arbeiten sollten. Im Rahmen der Voruntersuchung sollte jedoch zunächst die aufgezeigte Reduktion des Umfangs vorgenommen und auf ihre Brauchbarkeit hin getestet werden; die Abänderung des Prozederes der ersten Phase ist Gegenstand der Hauptuntersuchung.

I.6.1.2 Veränderungen des Designs gegenüber Experiment I

Wie bereits erläutert, sollte anhand der Voruntersuchung insbesondere überprüft werden, ob der observation-inflation-Effekt auch bei Reduktion des Untersuchungsumfangs noch nachzuweisen war, da nicht ausgeschlossen werden konnte, dass eine gewisse Art von „Überflutung“ durch die Masse an Material in der zweiten Phase dem Effekt zugute kam. Hiefür war also die Untersuchung nur einer Gruppe – einer Gruppe, die in der zweiten Phase die Aufgabe hatte, die Handlungsausführung zu beobachten – notwendig.

Sowohl die berichteten Daten der Vor- als auch der Hauptuntersuchung der zweiten Studie wurden an der University of Alberta, Edmonton, Kanada erhoben. Daher wurden alle Handlungsanweisungen sowie Instruktionen ins Englische übersetzt. Die Voruntersuchung sollte ebenfalls zeigen, ob diese Übersetzungen verständlich waren.

Ansonsten sollte so wenig wie möglich geändert werden, um möglichst selektiv den Einfluss der Umfangsreduktion erfassen zu können. Einige weitere Veränderungen waren jedoch unumgänglich und werden im Folgenden erläutert. Abgesehen von diesen war die Untersuchung jedoch identisch mit dem ersten Experiment.

In Studie I hatten sich einige Objekte als nicht ganz einfach in der Handhabung erwiesen. Weitere Objekte schienen für den geplanten Einsatz bei älteren TeilnehmerInnen und evtl. PatientInnen zu gefährlich – wie z.B. das Bedienen eines Feuerzeuges. Beide Arten von Objekten wurden (im ersten Falle durch gleichartige, aber besser handhabbare, im zweiten Falle durch neue Objekte und damit auch Handlungsanweisungen) ersetzt. Dies hatte zur Konsequenz, dass auch die Handlungsausführung erneut gefilmt werden musste. Für die Voruntersuchung nun wurden zum Großteil alte Videos präsentiert; ein kleiner Teil jedoch musste durch neue Videos ersetzt werden, um den neuen Objekten bzw. Handlungsanweisungen Rechnung zu tragen. Die neuen Videos waren hinsichtlich Szenerie und AkteurInnen analog zu den bereits bestehenden, weshalb die Ersetzung einiger Videos keinen großen Einfluss ausgeübt haben sollte. Um dies sicherzustellen bzw. eine mögliche Konfundierung auszuschließen, wurde jedoch die Anzahl an ersetzten Videos pro Pool, Subpool und pro AkteurIn ausbalanciert, so dass auch einige Videos ersetzt wurden, die die Handlungsausführung von Handlungsanweisungen bzw. mit Objekten zeigten, die nicht neu eingeführt worden waren. Insgesamt wurden so 18 Videos ausgetauscht; eine genaue Auflistung ist der Tabelle im Anhang B.3 zu entnehmen.

In Phase 1 wurden – wie im ersten Experiment – 30 Handlungsanweisungen präsentiert, die zur Hälfte zu lesen bzw. auszuführen waren. Es wurde jedoch eine kürzere Präsentationsdauer von 10 (anstatt – wie bisher – 15) Sekunden gewählt, da Studie I einerseits gezeigt hatte, dass innerhalb von 15 Sekunden sehr häufig gelesen bzw. ausgeführt werden konnte, andererseits die englischen Handlungsanweisungen kürzer – und damit schneller zu lesen – waren als die deutschen.

In Phase 2 kam die vorgenommene Umfangsreduktion zum Tragen: Die Phase bestand nun nur noch aus 60 anstatt 90 Präsentationen, da keine neuen Handlungsanweisungen gezeigt wurden. Nach jeder Darbietung wurde (wie auch in Studie I) sowohl nach der Häufigkeit der Beobachtung wie nach dem Ausmaß der Konzentration gefragt.

Um den Einsatz der Handlungsanweisungen auszubalancieren, wurde an den bestehenden neun Materialkombinationen festgehalten. Der in der ersten Sitzung nicht präsentierte Pool wurde beim Abruf gezeigt, so dass durch dieses Vorgehen auch eine Ausbalancierung bzgl. solcher Handlungsanweisungen stattfand, die nur im Gedächtnistest präsentiert wurden. Um ein ausgeglichenes Verhältnis von alten und neuen Handlungsanweisungen bei der Abfrage zu erzielen und so möglichen Antworttendenzen entgegenzuwirken, wurden zusätzlich 15 weitere Distraktortests präsentiert. Diese unterschieden sich von den vorher genannten lediglich darin, dass ihr Einsatz nicht ausbalanciert wurde (siehe Experiment I). In der Terminologie des Versuchsplans handelt es sich hierbei um Items, die in Phase 1 nicht enkodiert und in Phase 2 nicht präsentiert (Frequenz 0x) wurden.

In Phase 3 wurden also die 30 in Phase 1 bzw. 2 präsentierten, sowie die 15 nicht präsentierten (enkodierten) Items des Gesamtpools dargeboten. 15 Distraktoren, die stets nur in Phase 3 zum Einsatz kamen, wurden zusätzlich eingesetzt, so dass in dieser Phase insg. 60 Items (30 alte, 30 neue) dargeboten wurden. Im Unterschied zu Studie I wurden die TeilnehmerInnen bei der Abfrage zunächst gebeten, zu unterscheiden, ob eine Handlungsanweisung alt oder neu war – anstatt zu unterscheiden, ob sie in Phase 1 präsentiert worden war oder nicht (diese Überlegungen sind in der Tabelle 17, folgende Seite, zusammengefasst). Ansonsten verlief diese Phase identisch zu derjenigen in Experiment I (, d.h. im Falle einer *alt*-Antwort folgten Quellen- sowie Sicherheitsurteile). Allerdings konnte aus praktischen Gründen in der Voruntersuchung das Intervall zwischen erster und zweiter Sitzung nicht konstant gehalten werden und lag zwischen 7 und 11 Tagen ($M=8.00$, $SD=1.36$). Damit war der zeitliche Abstand deutlich kürzer als noch in der ersten Untersuchung.

Die Instruktionen zur Voruntersuchung des zweiten Experiments sind dem Anhang B.4 zu entnehmen.

Tab. 17: Darbietung der intraindividuell variierten unabhängigen und der abhängigen Variablen der Voruntersuchung zum zweiten Experiment (am Beispiel der Materialkombination 1).

Phase 1: Enkodierung	UV-A	Art der Enkodierung			
	Stufe d. UV-A	ausgeführt	gelesen	nicht enkodiert	
	Anzahl	15	15	(15)	Σ 30
	Pools	1A,1B,1C	2A,2B,2C	3A,3B,3C	
	Dauer	10 Sekunden/Handlungsanweisung			
	Reihenfolge	vollständig randomisiert			
Phase 2: Verarbeitung*	UV-B	Frequenz der Verarbeitung			
	Stufe d. UV-B	5x	1x	0x	
	Anzahl	10	10	(10)	Σ 60
	Pools	1A,2A	1B,2B	1C,2C	
	Dauer	15 Sekunden/Handlungsanweisung			
	Reihenfolge	fixiert randomisiert			
Phase 3: Abruf**	AVn	Rekognitions-/ Quellengedächtnis-/ Sicherheitsurteile			
	Anzahl	45 + 15 Distraktoren			Σ 60
	Pools	1A,1B,1C,2A,2B,2C,3A,3B,3C,4A,4B,4C			
	Dauer	nicht vorgegeben			
	Reihenfolge	vollständig randomisiert			

Anmerkungen: * 5 Minuten nach Phase 1; ** durchschnittlich 8 Tage nach Phase 1.

I.6.1.3 Stichprobe

An der Voruntersuchung nahmen 16 Studierende der University of Alberta, die im Bachelorstudium einen Einführungskurs Psychologie studierten und mit ihrer Teilnahme einen Teil der curricular vorgeschriebenen Versuchspersonenstunden ableisteten, teil. Die Ergebnisse einer Person, die sich ganz offenkundig instruktionswidrig verhielt, wurden aus der Analyse ausgeschlossen, so dass letztlich Daten von 15 TeilnehmerInnen in die Analyse eingingen. 10 Teilnehmerinnen waren weiblich, 5 Teilnehmer männlich. Das Alter reichte von 18 bis 23 Jahren ($M=19.20$, $SD=1.42$).

I.6.1.4 Ergebnisse

Die folgende Tabelle beinhaltet die relativen Häufigkeiten von *alt*- bzw. *ausgeführt*-Antworten, d.h. sowohl Rekognitions- als auch Quellengedächtnisurteile.

Tab. 18: Relative Häufigkeiten von *alt*- bzw. *ausgeführt*-Antworten im Rekognitions- bzw. Quellengedächtnistest der Voruntersuchung zum zweiten Experiment.

		Art der Abfrage/ Antwort:			
		Rekognitions- test/ <i>alt</i>		Quellen- gedächtnistest/ <i>ausgeführt</i>	
Art der Enkodierung in Phase 1: Ausgeführt		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.99	.05	.91	.13
	1x	.97	.07	.95	.09
	0x	.95	.12	.77	.21
Art der Enkodierung in Phase 1: Gelesen		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.96	.11	.32	.30
	1x	.89	.17	.19	.21
	0x	.67	.27	.11	.15
Art der Enkodierung in Phase 1: Nicht enkodiert		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	0x	.03	.04	.01	.02

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an „alt“- bzw. „ausgeführt“- Antworten relativiert an der Anzahl aller (gültigen) Handlungsanweisungen in der entsprechenden Zelle. $N=15$.

Die Voruntersuchung fokussierte v.a. auf die Frage, ob der observation-inflation-Effekt, d.h. mehr falsche *ausgeführt*-Antworten nach fünfmaliger im Vergleich zu nullmaliger Beobachtung, auch bei reduziertem Materialumfang nachweisbar ist – und darauf soll auch die Ergebnisdarstellung fokussieren. Weitere Befunde sollen nur dargestellt werden, sofern sie entweder zur Sicherung der Datenvalidität beitragen oder wichtige Hinweise auf die Konsequenzen der Abwandlung des Designs oder die zugrundeliegenden Mechanismen liefern. Ansonsten gelten die gleichen Überlegungen wie bzgl. Studie I.

Da die TeilnehmerInnen während der ersten Sitzung – wie schon in Studie I – autonom gearbeitet hatten, schien eine Absicherung der Datenvalidität (analog dem Vorgehen bei Studie I), wichtig.

Um sicherzustellen, dass die TeilnehmerInnen in der ersten Phase – beim Ausführen bzw. Lesen der Handlungsanweisungen – instruktionsgemäß gehandelt hatten, wurden drei Vergleiche herangezogen. Zunächst sollten Items häufiger gelesen als ausgeführt worden sein. Dies war der Fall: Im Durchschnitt wurden die Handlungsanweisungen 5.92 Mal gelesen ($SD=3.55$), aber nur 4.40 Mal ausgeführt ($SD=2.04$); $t(14)=2.28$, $p<.05$, $\eta^2=.270$. Die Variabilität der Häufigkeit lag jedoch – wie angenommen – mit durchschnittlich 3.59 ($SD=2.01$) beim Ausführen höher als beim Lesen mit ($M=1.50$, $SD=.97$); $t(14)=4.35$, $p<.05$, $\eta^2=.575$. Des Weiteren wurden die ausgeführten Handlungsanweisungen deutlich besser rekogniziert als die bloß gelesenen (enactment-Effekt); $t(14)=4.09$, $p<.05$, $\eta^2=.544$. Dieses Bild weist darauf hin, dass die Handlungsanweisungen in Phase 1 wie gewünscht gehandhabt wurden und wird abgerundet durch zwei Ratings, die nach dem Gedächtnistest in der zweiten Sitzung erhoben wurden. Wie schon in Studie I wurde explizit nach dem Befolgen der Instruktionen in Phase 1 wie auch in Phase 2 gefragt. Wiederum wurde eine siebenstufige Skala verwendet (höhere Werte = höhere compliance). Für Phase 1 lag die Einschätzung im Schnitt bei 6.80 ($SD=.41$), für Phase 2 bei 6.20 ($SD=.86$). Insgesamt kann also davon ausgegangen werden, dass die Instruktionen im intendierten Sinne umgesetzt wurden und die Daten diesbzgl. valide sind.

Bereits das deskriptivstatistische Bild zeigte eine deutliche Zunahme an falschen *ausgeführt*-Antworten nach fünfmaligem Videobeobachten in Phase 2 verglichen mit nullmaliger Video-präsentation; der observation-inflation-Effekt ließ sich auch inferenzstatistisch absichern: $t(14)=3.55$, $p<.05$, $\eta^2=.474$. 10 der 15 TeilnehmerInnen wiesen einen entsprechenden Effekt auf. Ebenfalls zeigte sich eine Zunahme an falschen *ausgeführt*-Antworten nach einmaligem Videobeobachten in Phase 2 verglichen mit nullmaligem Videobeobachten; dieser Effekt war allerdings nicht signifikant: $t(14)=1.39$, $p>.05$. Ebenso unterschieden sich auch einmalige und

fünfmalige Präsentation nicht im Ausmaß falscher *ausgeführt*-Antworten: $t(14)=1.44, p>.05$. Es ergab sich eine schwache, positive Korrelation zwischen dem (variierenden) Intervall zwischen erster und zweiter Sitzung und dem Ausmaß des observation-inflation-Effekts, die jedoch nicht signifikant wurde ($r(15)=.270, p>.05$). Tendenziell also waren falsche *ausgeführt*-Antworten nach wiederholter Beobachtung mit zunehmendem Zeitintervall wahrscheinlicher. Umgekehrt ergaben sich durchweg negative, teilweise signifikante Zusammenhänge zwischen dem Intervall und der Rekognitionsleistung. Tendenziell war also die Rekognition bei längerem Intervall schlechter.

Obwohl – wie schon in Untersuchung I – nicht alle Materialkombinationen gleich häufig eingesetzt werden konnten, soll auf eine Überprüfung hinsichtlich eines möglichen Materialeffekts an dieser Stelle verzichtet werden, da bei 9 Materialkombinationen und 15 TeilnehmerInnen die Zellbesetzungen viel zu klein wären. Dieselben Überlegungen, die schon im Rahmen der ersten Studie getroffen wurden, legen aber nahe, dass es sich bei dem gefundenen observation-inflation-Effekt nicht um einen reinen Materialeffekt handeln kann (siehe Abschnitte I.5.3.3 und I.5.4.1).

Im Folgenden werden die Ergebnisse einiger Ratings berichtet, die zum Abschluss der zweiten Sitzung erbeten wurden. Diese Einschätzungen könnten im Vergleich mit denjenigen der folgenden Hauptuntersuchung interessant sein.

Tab. 19: Einschätzungen verschiedener Items (siebenstufige Skala) in der Voruntersuchung zum zweiten Experiment.

Item	<i>M</i> (<i>SD</i>)
Item 1: I expected that my memory for the action statements would be tested when I was here for the first session.	3.80 (2.11)
Item 2: I expected exactly this kind of memory test when I was here for the first session.	2.93 (2.09)
Item 3: The memory task I just fulfilled was very easy.	5.00 (1.89)
Item 4: I was very sure about which actions I actually performed.	4.93 (1.91)
Item 5: I was very sure about which actions I merely read.	4.80 (1.61)
Item 6: My memory for the last session is very clear and vivid.	4.87 (1.41)
Item 7: I didn't think of the video-observation task while I was doing the memory task.	3.93 (2.12)
Item 8: While I was doing the memory task, I feared that my memory for the first phase could be distorted by the videos I observed within the third phase of the first session.	4.67 (1.88)

Anmerkungen: $N=15$.

I.6.1.5 Résumé und Ausblick

Der Voruntersuchung lag keine theoretisch-inhaltliche Fragestellung zugrunde. Vielmehr sollte überprüft werden, ob eine Abwandlung im Sinne einer Verschlankung des Prozederes sich auf das Zustandekommen des observation-inflation-Effekts auswirkt. Daher soll an dieser Stelle auch auf ausführliche Überlegungen zu verschiedenen Validitätsaspekten bzw. alternativen Interpretationsmöglichkeiten, die im Übrigen denjenigen der Studie I ähneln würden, verzichtet und mit einem Satz resümiert werden: Bei einer offensichtlich hohen compliance der Versuchspersonen konnte ein starker observation-inflation-Effekt auch nach Modifikation des Prozederes nachgewiesen werden. Dies ist umso erstaunlicher, als keine perfekte Realisation gewährleistet werden konnte – so wurden alte und neue Videos eingesetzt und das Intervall zwischen Enkodierung und Abruf variierte. Insgesamt sind die Ergebnisse daher ermutigend und sprechen dafür, die gewählte Modifikation beizubehalten.

I.6 2 Versuchsplanung

I.6.2.1 Operationalisierung und Variablenvalidität

Da weiterhin dasselbe Paradigma mit seinen zwei intraindividuell variierten unabhängigen Variablen „Art der Enkodierung in Phase 1“ und „Frequenz der Verarbeitung in Phase 2“ sowie den abhängigen Variablen „Rekognitions“- , „Quellengedächtnis“- und „Sicherheitsurteile“ zum Einsatz kam, kann auf die diesbzgl. Überlegungen zur Operationalisierung und Variablenvalidität an anderer Stelle (siehe I.5.1.1) zurückgegriffen werden.

Die zweite Studie fokussierte nun auf falsche Erinnerungen an die Handlungsausführung, die durch Beobachtung hervorgerufen werden. Da vor dem Hintergrund der bisherigen Ergebnisse nicht ausgeschlossen werden konnte, dass es sich beim observation-inflation-Effekt gar nicht um eine Gedächtnistäuschung, sondern um eine Antworttendenz handelt (vgl. Abschnitt I.5.4), bestand die zentrale theoretisch-inhaltliche Zielsetzung der zweiten Studie darin, zu prüfen, ob sich der Effekt durch die Induktion strengerer Antwortkriterien eliminieren oder zumindest reduzieren ließ. Dabei sollte die Manipulation des Antwortkriteriums anhand von zusätzlichen Instruktionen bei der Abfrage erfolgen (vgl. Abschnitt I.5.4). Zum Vergleich wurde daher eine Kontrollgruppe benötigt, die bei der Abfrage keine zusätzliche Instruktion erhielt. Daneben sollten – in Analogie zur ersten Studie – Daten einer weiteren Kontrollgruppe erhoben werden, die die Handlungsanweisungen in Phase 2 nur lesen sollte. Hierdurch sollte wiederum sichergestellt werden, dass der Effekt (auch nach Veränderungen am Prozedere;

siehe weiter unten) nicht allein auf die bloße Vertrautheit der Handlungsanweisungen zurückgeführt werden könnte.

In Anlehnung an Thomas und Bulevich (2006; vgl. Abschnitt I.5.4.3) wurde eine Instruktion zur Quellenprüfung entwickelt, die die TeilnehmerInnen dazu veranlassen sollte, beim Gedächtnistest verstärkt auf Hinweise für die Handlungsausführung zu achten, also ihre Erinnerungen genau zu überprüfen und ihre Urteile auf qualitative Erinnerungsspuren zu basieren, d.h. bspw. nicht auf Basis eines bloßen Vertrautheitsgefühls zu urteilen (siehe Abschnitt I.5.4.3). Die TeilnehmerInnen wurden z.B. instruiert, sich beim Abruf zu fragen, ob sie sich daran erinnern könnten, wie sich das Objekt in ihrer Hand anfühlte, welche Geräusche entstanden oder wie genau sie die Handlung ausführten. Um die TeilnehmerInnen nicht fehlzuleiten, wurden sie ausdrücklich darauf aufmerksam gemacht, dass sensorische Eindrücke – z.B. Geräusche – auch dem Video entstammen könnten und nicht zwangsläufig mit der eigenen Handlungsausführung verknüpft sein müssten. Da allerdings eine ähnlich geartete Instruktion in der Studie von Thomas und Bulevich (2006) keinen Einfluss auf das Ausmaß an falschen Erinnerungen hatte, wurde neben diese eine zweite Variante gestellt.

Viele Studien haben gezeigt, dass falsche Erinnerungen durch Warnungen vor derartigen Effekten beim Abruf reduziert werden können; dies scheint besonders der Fall zu sein, wenn eine soziale Komponente involviert ist (z.B. Chambers & Zaragoza, 2001; Echterhoff, Hirst & Hussy, 2005). Daher sollte eine weitere Gruppe eine Warnung vor dem observation-inflation-Effekt, d.h. vor falschen Erinnerungen aufgrund der Beobachtung einer anderen Person, erhalten (siehe Abschnitt I.5.4.3). Es wurde vermutet, dass die Warnung einen gleichartigen, aber möglicherweise stärkeren Effekt bewirken könnte als die Instruktion zur Quellenprüfung. Die TeilnehmerInnen waren sich nämlich – so zeigten Experiment I und die Voruntersuchung zu Experiment II – ihrer falschen Erinnerungen generell sehr sicher, was darauf hindeuten könnte, dass sie ihre Erinnerungsurteile ohnehin (auch) auf qualitative Merkmale basierten (siehe auch Thomas & Bulevich, 2006). Die Instruktion zur Quellenprüfung verbalisierte daher vermutlich nur dieselbe Strategie, die sie auch ohne eine derartige Instruktion verfolgten. Ein expliziter Verweis auf eine mögliche (soziale) Gedächtnisbeeinflussung könnte jedoch zu stärkeren Reaktionen führen als die „bloße“ Anleitung zu einem vermutlich sowieso umgesetzten Vorgehen, die eher als überflüssig wahrgenommen werden könnte (z.B. Chambers & Zaragoza, 2001; Echterhoff, Hirst & Hussy, 2005).

Beide Abrufinstruktionen sollten interindividuell variiert werden, um Übertragungseffekt zu vermeiden. Der genaue Wortlaut ist dem Ablauf (Abschnitt I.6.3.4) zu entnehmen.

Ein Ablaufschema ist in der folgenden Tabelle dargestellt.

Tab. 20: Ablauf des zweiten Experiments.

Phase	Variable	Ausprägungen	Art der Variation
Phase 1: Enkodierung	UV-A: Art der Enkodierung	ausgeführt gelesen nicht enkodiert	intraindividuell
Phase 2: Verarbeitung*	UV-B: Frequenz der Verarbeitung	5x 1x 0x	intraindividuell
	UV-C: Art der Verarbeitung	beobachten lesen	interindividuell
Phase 3: Abruf**	UV-D: Art der Abruf- manipulation***	keine Quellenprüfung Warnung	interindividuell
	AV-1: Rekognition	alt neu	--
	AV-2: Quellengedächtnis	ausgeführt nicht ausgeführt	--
	AV-3: Sicherheit	siebenstufige Skala	--

Anmerkungen: * 5 Minuten nach Phase 1; ** 2 Wochen nach Phase 2; *** Quellenprüfung und Warnung wurden nur nach Beobachtung in Phase 2 realisiert.

Die Hypothesenprüfung bezieht sich – wie schon in Studie I – auf den Vergleich der Quote falscher *ausgeführt*-Antworten nach fünf- vs. nullmaliger Präsentation. Die Hypothese

Beim observation-inflation-Effekt handelt es sich zumeist nicht um das Resultat einer Antworttendenz bzw. eines laxen Antwortkriteriums, sondern um eine a posteriori kaum vermeidliche Gedächtnistäuschung.

soll als bewährt angesehen werden, wenn alle folgenden Bedingungen erfüllt sind (die Notation bezieht sich auf den im folgenden Abschnitt präsentierten Versuchsplan):

- (1) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation nicht signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsanweisungen lasen und beim Abruf keine zusätzliche Instruktion erhielten.

$$H_0: \mu_{A2B1C1D1} \leq \mu_{A2B3C1D1}; H_1: \mu_{A2B1C1D1} > \mu_{A2B3C1D1}$$

- (2) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und keine Abrufmanipulation erhielten.

$$H_1: \mu_{A2B1C2D1} > \mu_{A2B3C2D1}; H_0: \mu_{A2B1C2D1} \leq \mu_{A2B3C2D1}$$

- (3) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und beim Abruf eine Instruktion zur Quellenprüfung erhielten.

$$H_1: \mu_{A2B1C2D2} > \mu_{A2B3C2D2}; H_0: \mu_{A2B1C2D2} \leq \mu_{A2B3C2D2}$$

- (4) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und beim Abruf eine Warnung bzgl. des observation-inflation-Effekts erhielten.

$$H_1: \mu_{A2B1C2D3} > \mu_{A2B3C2D3}; H_0: \mu_{A2B1C2D3} \leq \mu_{A2B3C2D3}$$

- (5) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und keine Abrufmanipulation erhielten, als wenn sie in der zweiten Phase die Handlungsanweisungen lasen und ebenfalls keine Abrufmanipulation erhielten.

$$H_1: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} > \mu_{A2B1C1D1} - \mu_{A2B3C1D1};$$

$$H_0: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} \leq \mu_{A2B1C1D1} - \mu_{A2B3C1D1}$$

- (6) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist nicht signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und keine Abrufmanipulation erhielten, als wenn sie in der zweiten Phase die Handlungsausführung beobachteten und beim Abruf eine Instruktion zur Quellenprüfung erhielten.

$$H_0: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} \leq \mu_{A2B1C2D2} - \mu_{A2B3C2D2};$$

$$H_1: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} > \mu_{A2B1C2D2} - \mu_{A2B3C2D2}$$

- (7) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist nicht signifikant höher, wenn die TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und keine Abrufmanipulation erhielten, als wenn sie in der zweiten Phase die Handlungsausführung beobachteten und beim Abruf eine Warnung erhielten.

$$H_0: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} \leq \mu_{A2B1C2D3} - \mu_{A2B3C2D3};$$

$$H_1: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} > \mu_{A2B1C2D3} - \mu_{A2B3C2D3}$$

Neben der bereits im Rahmen der Voruntersuchung überprüften Modifikation der zweiten Phase erschien eine Modifikation der ersten Phase aus drei Gründen wichtig: Erstens hatten zwar alle bisherigen Studien erbracht, dass die TeilnehmerInnen mit dem autonomen Arbeiten gut

zurecht kamen. Dennoch war so – trotz Angabe der Versuchspersonen und Absicherung über verschiedene Parameter – eine vollständige Kontrolle der instruktionsgemäßen Handhabung nicht möglich. Zweitens war es aufwändig, für jeden Teilnehmer einzeln zu bestimmen, ob und welche Handlungsanweisungen entgegen der Instruktion gehandhabt worden waren und diese bei der Berechnung der Ergebnisse dann ggf. unberücksichtigt zu lassen. Drittens war diese Vorgehensweise im Hinblick auf den Einsatz mit älteren gesunden (und evtl. später auch dementen) TeilnehmerInnen kaum denkbar. Die Situation, einen Gegenstand in der Hand zu haben, ihn dann aber nicht zu gebrauchen, war aufgrund des Aufforderungscharakters der Gegenstände und Handlungsanweisungen schon für junge gesunde TeilnehmerInnen teilweise schwierig und sollte daher im Hinblick auf weitere in die Untersuchung einzubeziehende Populationen vermieden werden. Zusätzlich wurde davon ausgegangen, dass das Wiederholen des Lesens bzw. Ausführens insb. von den älteren TeilnehmerInnen, die nicht an derartige experimentpsychologische Maßnahmen gewöhnt sind, als seltsam empfunden und daher verweigert werden könnte. Deshalb sollte ein Prozedere entwickelt werden, das sowohl das autonome Arbeiten der TeilnehmerInnen als auch die Wiederholungen sowie das Berühren von nicht gebrauchten Objekten ersetzte. Die genaue Ausgestaltung der ersten Phase wird im Abschnitt I.6.3.4 erläutert.

Ob sich diese Änderungen auf das Ausmaß des Effekts auswirken würden, war jedoch ungewiss. Einerseits sollte die Diskrimination ausgeführt – nicht ausgeführt hierdurch erleichtert werden, da die Tatsache, dass ein Objekt überhaupt nicht in der Hand gehalten wurde, als zusätzlicher Hinweis gegen die Ausführung fungieren könnte (siehe die niedrigere Quote falscher *ausgeführt*-Antworten auf in Phase 1 nicht enkodierte im Vergleich zu gelesenen Items innerhalb des ersten Experiments). Andererseits sollte der Wegfall der Wiederholungen des Ausführens bzw. Lesens zu einer weniger tiefen Erinnerungsspur führen und daher die Diskrimination ausgeführt – nicht ausgeführt erschweren. Insgesamt wurde daher von einem Ausgleich beider Tendenzen und somit von einer analogen Effektgröße wie in der Voruntersuchung ausgegangen (s. Abschnitt I.6.1).

I.6.2.2 Versuchsplananlage und Versuchsplan

Der Untersuchung lagen also die intraindividuell variierte, dreistufige Variable „Art der Enkodierung in der 1. Phase“ (UV-A: ausgeführt, gelesen, nicht enkodiert) und die ebenfalls intraindividuell variierte, dreistufige Variable „Frequenz der Präsentation in der 2. Phase“ (UV-B: 5x, 1x, 0x) zugrunde, wobei für die in Phase 1 nicht enkodierten Handlungsanweisungen nur noch eine der drei Frequenzstufen realisiert wurde (0x). Es wurden vier Gruppen benötigt, um die anvisierten Untersuchungsziele zu erreichen: Eine Gruppe sollte beim Abruf eine In-

struktion zur Quellenprüfung, eine andere eine Warnung erhalten, nachdem in Phase 2 die Handlungsausführung beobachtet worden war (UV-C: Beobachten, UV-D: Quellenprüfung, Warnung). Eine Kontrollgruppe sollte keine zusätzliche Abrufinstruktion erhalten, jedoch in Phase 2 ebenfalls die Videos beobachten (UV-C: Beobachten, UV-D: keine Abrufmanipulation). Eine weitere Kontrollgruppe sollte keine zusätzliche Abrufinstruktion erhalten und in Phase 2 auch nicht die Videos beobachten, sondern nur die entsprechenden Handlungsanweisungen lesen (UV-C: Lesen, UV-D: keine Abrufmanipulation). Es sollten keine Gruppen realisiert werden, die die Handlungsanweisungen in Phase 2 lesen sollten und eine Abrufmanipulationen erhielten, da davon ausgegangen wurde, dass – wie in Studie I – der Effekt unter dieser Verarbeitungsart gar nicht auftreten würde. Als abhängige Variablen wurden wiederum Rekognitions- und Quellengedächtnisurteile sowie Sicherheitsratings bzgl. der Erinnerungsurteile erhoben. Es resultierte also eine vierfaktorielle Versuchsplananlage (VPL-A4, Hussy & Jain, 2002; siehe Tabelle 21).

Die Umsetzung dieser Versuchsplananlage in den konkreten Versuchsplan (VPL4-RRRR) basierte auf denselben Überlegungen wie in Studie I. Um Übertragungseffekte zu vermeiden, wurde nicht nur die Art der Verarbeitung der Handlungsanweisungen in der zweiten Phase, sondern auch die Art der Instruktion in Phase 3 (UV-D) interindividuell variiert.

Tab. 21: Versuchsplan des zweiten Experiments.

				UV-C: Art der Verarbeitung in Phase 2			
				Lesen	Beobachten		
				UV-D: Art der Abrufmanipulation in Phase 3			
				keine	keine	Quellenprüfung	Warnung
UV-A: Art der Enkodierung in Phase 1	ausgeführt	UV-B: Frequenz der Prä- sentation in Phase 2	5x	A ₁ B ₁ C ₁ D ₁	A ₁ B ₁ C ₂ D ₁	A ₁ B ₁ C ₂ D ₂	A ₁ B ₁ C ₂ D ₃
			1x	A ₁ B ₂ C ₁ D ₁	A ₁ B ₂ C ₂ D ₁	A ₁ B ₂ C ₂ D ₂	A ₁ B ₂ C ₂ D ₃
			0x	A ₁ B ₃ C ₁ D ₁	A ₁ B ₃ C ₂ D ₁	A ₁ B ₃ C ₂ D ₂	A ₁ B ₃ C ₂ D ₃
	gelesen		5x	A ₂ B ₁ C ₁ D ₁	A ₂ B ₁ C ₂ D ₁	A ₂ B ₁ C ₂ D ₂	A ₂ B ₁ C ₂ D ₃
			1x	A ₂ B ₂ C ₁ D ₁	A ₂ B ₂ C ₂ D ₁	A ₂ B ₂ C ₂ D ₂	A ₂ B ₂ C ₂ D ₃
			0x	A ₂ B ₃ C ₁ D ₁	A ₂ B ₃ C ₂ D ₁	A ₂ B ₃ C ₂ D ₂	A ₂ B ₃ C ₂ D ₃
	nicht enkodiert		0x	A ₃ B ₃ C ₁ D ₁	A ₃ B ₃ C ₂ D ₁	A ₃ B ₃ C ₂ D ₂	A ₃ B ₃ C ₂ D ₃

I.6.2.3 Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang

Die Überlegungen entsprachen denjenigen, die bereits bzgl. des ersten Experiments erörtert wurden. Wiederum basierte die Versuchsplanung auf dem Vergleich der Frequenzen 5x und 0x, so dass auch diesmal die Bedingung 1x nicht in die (hypothesenprüfende) Auswertung einfluss.

Die Studie von Thomas und Bulevich (2006) hat gezeigt, dass jüngere TeilnehmerInnen nicht von einer Instruktion zur Quellenprüfung profitierten bzw. der Effekt genauso stark ausgeprägt war, wie ohne die Abrufhilfe. Dieser Befund lässt sich mutmaßlich auf den observation-inflation-Effekt übertragen; schließlich führten beide Prozesse schon in Studie I zu sehr ähnlichen Ergebnismustern. In Ermangelung empirischer Evidenz bzgl. der Wirkung von Warnungen auf den imagination- bzw. observation-inflation-Effekt, wurde unter beiden Abrufmanipulationen von einem gleich bleibend großen Effekt und somit von derselben optimalen Stichprobengröße wie in Experiment I ausgegangen: 18 Personen pro Gruppe, d.h. 72 TeilnehmerInnen insgesamt ($\eta^2=.14$, $\alpha=.05$, $\beta=.20$).

I.6.2.4 Kontrolle von allgemeinen Störeffekten

Siehe Experiment I

I.6.2.5 Kontrolle von speziellen Stör- und von Materialeffekten

Siehe Experiment I.

I.6.2.6 Ethische Gesichtspunkte

Die Überlegungen entsprechen denjenigen zur ersten Untersuchung (vgl. Abschnitt I.5.1.6). Die Durchführung an der University of Alberta erforderte zusätzlich eine Genehmigung des zuständigen Reviewboardes (Faculties of Arts, Science, and Law) unter ethischen Gesichtspunkten. Diese wurde ohne Auflagen erteilt (#1632 DK-11-16-07-070). Das Prozedere sah (wie schon in Studie I umgesetzt) den Abschluss eines Vertrages zwischen Versuchsleiterin und TeilnehmerInnen vor (Einverständniserklärung, siehe Anhang B.1) sowie die Anfertigung einer Beschreibung der Untersuchung, die den TeilnehmerInnen am Ende ausgehändigt werden sollte (Aufklärungsformular, siehe Anhang B.2).

I.6.3 Versuchsdurchführung

I.6.3.1 Stichprobe

Die TeilnehmerInnen wurden über das Online-System des Departments Psychologie der University of Alberta auf die Untersuchung aufmerksam gemacht. 82 Bachelor-Studierende, die einen Einführungskurs Psychologie belegten, nahmen an der ersten Untersuchungssitzung teil. Acht von ihnen erschienen nicht zur zweiten Sitzung. Von den verbliebenen 74 vollständigen Datensätzen gingen jedoch nur 56 in die Auswertung ein. Sechs der insg. 18 nicht in die Analyse eingeschlossenen Datensätze stammten von Personen, die vorläufige Varianten (insb. der Instruktion zur Quellenprüfung) erhielten, anhand derer die endgültigen Versionen erst entwickelt wurden. Andere ausgeschlossene Datensätze stammten von Personen, die sich offensichtlich nicht instruktionsgemäß verhalten hatten, nicht über ausreichende Sprachkenntnisse verfügten oder sich altersmäßig deutlich von der Gesamtgruppe abhoben (sechs TeilnehmerInnen). Durch eine beschädigte Festplatte gingen bereits erhobene Daten von sieben Personen aus der zweiten Sitzung verloren; von sechs der sieben verlorenen Datensätze existierten jedoch glücklicherweise Aufzeichnungen, die teilweise in die Auswertung einbezogen wurden. Des Weiteren mussten vier zum Zeitpunkt des Festplatten-Ausfalls noch ausstehende Datensätze der zweiten Sitzung im paper-pencil Verfahren erhoben werden. Um eine Vergleichbarkeit der Datengewinnung sicherzustellen, wurden diese vier Datensätze bis auf einen von der Analyse ausgeschlossen. Insgesamt entfielen durch den Ausfall der Festplatte somit sechs Datensätze.

Die Stichprobe von 56 TeilnehmerInnen, 14 pro Gruppe, lässt sich folgendermaßen charakterisieren: Das Alter umfasste eine Spannweite von 18 bis 33 Jahren mit einem Mittelwert von 19.67 Jahren ($SD=2.64$), wobei 33 Teilnehmerinnen weiblich und 23 Teilnehmer männlich waren. Die Verteilung des Alters und des Geschlechts war zwischen den Gruppen vergleichbar.

I.6.3.2 Material

Die bereits in Experiment I erfolgreich eingesetzten Pools und Subpools von Handlungsanweisungen blieben bis auf minimale Änderungen, die durch den Austausch von Handlungsanweisungen entstanden (siehe Voruntersuchung), bestehen. Es wurden ebenfalls dieselben Kombinationen der Pools und Subpools eingesetzt wie in Experiment I.

Die Videos der ersten Untersuchung wurden aufgrund des Austauschs einiger Objekte bzw. Handlungsanweisungen durch neue Videos ersetzt (siehe Abschnitt I.6.1.2). Da sich im Rah-

men der ersten Studie keinerlei Einfluss des Geschlechts der AkteurInnen auf die falschen Erinnerungen gezeigt hatte und anvisiert wurde, möglichst homogene Videos zu nutzen, wurde nur noch ein (männlicher) Akteur gewählt. Die Videos glichen ansonsten denjenigen aus Studie I. Eine Auflistung der Handlungsanweisungen ist dem Anhang B.3 zu entnehmen.

Die Instruktion zur Quellenprüfung wurde in Anlehnung an Thomas und Bulevich (2006) formuliert, die Warnung wiederum in Anlehnung an diese bzw. bereits erfolgreich eingesetzte Warnungen (siehe z.B. Landau & von Glahn, 2004; Neuschatz, Benoit & Payne, 2003).

I.6.3.3 Hilfsmittel und Geräte

Hilfsmittel und Geräte waren identisch mit denjenigen der ersten Studie. Allerdings wurde diesmal ein MacBook eingesetzt (2,16 GHz mit Intel Core 2 Duo Prozessor), das an einen 19"-Monitor (Bildschirmauflösung 1024x768) angeschlossen war und über eine externe Tastatur und Maus verfügte.

I.6.3.4 Ablauf

Die Untersuchung lief – bis auf die Fillertask, die im paper-pencil-Verfahren durchgeführt wurde, – vollständig computergestützt ab.

Bisher war das Augenmerk darauf gerichtet, in der ersten Phase eine gleich lange Enkodierung der Handlungsanweisungen zu gewährleisten – d.h. das Lesen oder Ausführen sollte für die Dauer von 15 Sekunden immer wieder wiederholt werden. Im Zuge der Modifikation der ersten Phase wurde beschlossen, nicht die Zeit, sondern die Häufigkeit des Lesens bzw. Ausführens sowie das Prozedere zu vereinheitlichen (siehe Abschnitt I.6.2). Dennoch sollte die Präsentationsdauer pro Handlungsanweisung gleich gehalten werden, um eine vergleichbare Exposition zu gewährleisten. So kam folgender Ablauf zustande:

Nach der Begrüßung und dem Abschluss des Vertrages wurden den TeilnehmerInnen 30 Handlungsanweisungen in zufälliger Reihenfolge für jeweils 10 Sekunden präsentiert. Vorgeschaltet war – wie auch in Studie I – die Nennung des in die jeweilige Handlungsanweisung eingeschlossenen Objektes bzw. der eingeschlossenen Objekte. Um die Tatsache auszugleichen, dass die TeilnehmerInnen die Objekte nun nicht mehr selber auswählten und sich hierdurch automatisch einen (v.a. konkreten visuellen) Eindruck von den Gegenständen verschafften, und um die Tiefe der Objektverarbeitung für ausgeführte und nicht ausgeführte Handlungsanweisungen vergleichbar zu gestalten, wurde ihnen innerhalb dieser Studie nicht nur die Objektbezeichnung, sondern auch ein Photo des Objektes für jeweils 5 Sekunden

die Objektbezeichnung, sondern auch ein Photo des Objektes für jeweils 5 Sekunden präsentiert und sie wurden gebeten, anzugeben, an wie vielen Tagen einer gewöhnlichen Woche sie den Gegenstand/ die Gegenstände benutzen würden (die Photos der Objekte sind in Anhang B.3 abgedruckt). Die Objekte befanden sich hinter einem Sichtschutz und wurden von der Versuchsleiterin für jede Handlungsanweisung (auch für die nur zu lesenden) einzeln hervorgeholt und in Sichtweite der ProbandInnen gelegt¹¹. Nach der Eingabe der Benutzungshäufigkeit erschien die entsprechende Handlungsanweisung zusammen mit der Aufforderung, wie mit dieser zu verfahren sei. 15 Handlungsanweisungen sollten ausgeführt werden. Hierzu wurde den TeilnehmerInnen das entsprechende Objekt gereicht. Nach einmaliger Handlungsausführung wurde dieses von der Versuchsleiterin wieder hinter dem Sichtschutz verborgen. 15 Handlungsanweisungen sollten gelesen werden. Um das Prozedere zu vereinheitlichen, wurden die TeilnehmerInnen gebeten, auch in diesem Falle nach dem Lesen der Anweisung eine Handlung auszuführen. Dabei handelte es sich jedoch um eine bedeutungslose Handlung, die aus vier Grundbewegungen bestand, die mit einer Hand auszuführen waren (Faust, Handkante, Handfläche, Zeigefinger). Zusammen mit jeder Handlungsanweisung, die nur zu lesen war, wurde eine zufällige Kombination von drei bis fünf dieser Handbewegungen durch die Versuchsleiterin vorgemacht, die zeitgleich von den TeilnehmerInnen imitiert werden sollte. Die Kombinationen waren von unterschiedlicher Länge, da auch die Handlungsausführungen von unterschiedlicher Dauer waren und somit eine Äquivalenz hergestellt werden konnte. Neben einer Vereinheitlichung des Prozederes wurde so sichergestellt, dass die TeilnehmerInnen dem Objekt auch bei Nicht-Ausführung ebenso lange ausgesetzt waren, wie bei Ausführung der Handlung – die Handbewegung wurde nämlich genau neben dem Objekt ausgeübt. Anschließend wurde das Objekt durch die Versuchsleiterin – wie im Falle des Ausführens – außer Sichtweite gebracht.

Der Ablauf der ersten Phase ist in Abbildung 10 dargestellt.

¹¹ Photos der Objekte wurden zusätzlich auf dem Bildschirm präsentiert, da Voruntersuchungen gezeigt hatten, dass die TeilnehmerInnen sehr auf den Monitor fixiert waren und daher die in Sichtweite gebrachten Objekte im Falle des Nicht-Ausführens kaum beachteten.

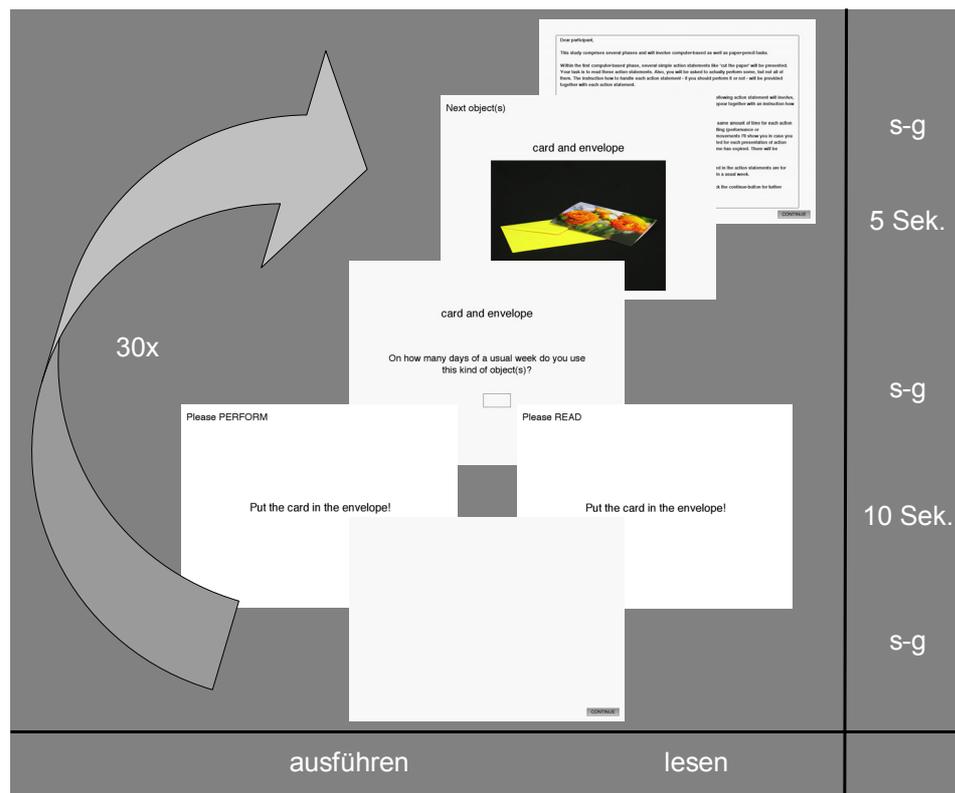


Abb. 10: Ablauf der ersten Phase des zweiten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.

Als Fillertask kam eine Aufgabe zum Einsatz, die das Generieren von Wörtern zu einem vorgegebenen Anfangsbuchstaben bzw. zu einer vorgegebenen Kategorie erforderte. Diese Aufgabe war Teil der neuropsychologischen Testbatterie, die die ebenfalls eingeschlossenen Gruppen älterer gesunder TeilnehmerInnen durchliefen und wurde bei der Untersuchung jener Population aus Gründen der Zeitersparnis bei vergleichbarer Dauer an die Stelle des NFC/FII gesetzt (siehe Abschnitt II). Um die Ergebnisse beider Altersgruppen vergleichen zu können, wurden auch die jüngeren TeilnehmerInnen dieser Prozedur unterworfen.

Der Ablauf der zweiten Phase war identisch mit demjenigen innerhalb von Studie I, allerdings wurden keine neuen Items dargeboten (siehe Voruntersuchung). Nach jeder Präsentation wurde jedoch nur noch eine Eingabe getätigt. Im Falle des *Beobachtens* wurde – wie in Experiment I – jeweils eine Handlungsanweisung zusammen mit dem Video, das die entsprechende Handlungsausführung zeigte, für 15 Sekunden präsentiert. Um zu gewährleisten, dass die TeilnehmerInnen die Handlungsausführung auch nach wiederholter Präsentation aufmerksam verfolgten, sollte nach jeder Darbietung angegeben werden, wie häufig die Handlung von der beobachteten Person vollständig ausgeführt worden war. Im Gegensatz zur ersten Studie wurden auch bzgl. der *Lesen*-Bedingung nun nicht mehr die Dauer, sondern die Häufigkeit der Verarbeitung sowie das Prozedere parallelisiert. Hierzu wurde jeweils eine Handlungsanwei-

sung zusammen mit einer Reihe bedeutungsloser visueller Stimuli – verschiedene geometrische Formen (Kreis, Quadrat, Dreieck), die in zufälliger Reihenfolge und für eine zufällige Dauer erschienen – für 15 Sekunden präsentiert. Die TeilnehmerInnen wurden instruiert, jede Handlungsanweisung einmal zu lesen und dann die geometrischen Figuren zu beobachten. Nach jeder Präsentation sollte angegeben werden, wie viele Kreise erschienen waren (siehe Abbildung 11).



Abb. 11: Ablauf der zweiten Phase des zweiten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.

Phase 3 lief so ab, wie in der Voruntersuchung, d.h. es wurde ein einfacher Rekognitionstest für die erste Sitzung (alt-neu-Diskrimination) mit einem Quellengedächtnistest (ausgeführt - nicht ausgeführt) kombiniert. Auch die Sicherheit bzgl. der Erinnerungsurteile wurde erhoben. Insg. wurden 60 Items in zufälliger Reihenfolge präsentiert.

In der Bedingung „*Instruktion zur Quellenprüfung*“ wurde zusätzlich zur allgemeinen Instruktion die folgende Anleitung zum Vorgehen gegeben:

If an action statement is OLD, you might have a vivid memory for the objects. You might remember, for example, that I brought out black scissors and a white sheet of paper. If an action statement is NEW, you probably will NOT have a vivid memory for the objects, because they were never shown to you within this study.

BE CAREFUL! You should make sure that – if you have a vivid memory – it comes from this study, and not from somewhere else: For example, you might have used scissors at home yesterday. We only want you to say 'old' to action statements that were presented within this study.

If an action statement was PERFORMED by you, you might in addition have a vivid memory for performing the action. You might clearly remember the way you performed the action, and/or how it felt, sounded, or looked when you were performing it. For example, you might remember how it felt when you were cutting the paper and that you cut it in half. Also, you might remember what you felt or thought while performing the action. For example, you might remember that you were not sure whether you should cut the paper the whole way through or not. Or you might, for example, remember that you thought that the scissors were similar to the ones you use at home.

You should make sure that – if you have a vivid memory for performing an action – it actually comes from your OWN performance. Remember, you also saw videos showing another person performing some of these actions, and so you have to make sure not to mix up performance of the person in the video with your own performance. For example, try to remember if you actually held the scissors in your hand, and whether you viewed cutting the paper from your own point of view.

If you DIDN'T PERFORM an action statement, you might also NOT remember the way you performed the action, and/ or how it felt, sounded, or looked when you were performing it, because you never touched the objects. You might nevertheless remember what you felt or thought in conjunction with that action statement; for example you might have been disappointed or glad that you were not supposed to perform a certain action

So, try to ask yourself the following questions for each action statement and base your memory judgments on your answers to these questions:

- Do I have a vivid memory for the object?
- Do I have a memory for the way I performed the action?
- Do I have a memory for how action felt, looked, or sounded?
- Do I remember what I was feeling or thinking when I saw this action statement?

MAKE SURE you refer to what you experienced within THIS study and not other contexts. And make sure you refer to your OWN action performance.

In der Bedingung mit *Warnung* wurde neben der allgemeinen Instruktion folgende Warnung ausgesprochen:

In the first session, you watched some videos showing another person performing some actions that you DID actually perform yourself. But also, you watched some videos showing another person performing some actions that you DID NOT actually perform yourself.

BE CAREFUL concerning these videos about actions that you DID NOT perform yourself!

Research has determined that observing another person performing actions, especially repeatedly, can lead to false memories for action performance. That is, people might sometimes think that they actually performed an action when in fact they only observed another person performing it.

So, you should be careful when performing this memory test because you watched some videos showing another person performing some actions that you didn't actually perform yourself.

Try not to let these videos influence your judgments about whether you performed an action or not.

Die Instruktionen zur Hauptuntersuchung des zweiten Experiments sind im Anhang B.5 vollständig abgedruckt.

Dem Erinnerungstest folgten 13 Aussagen zur Untersuchungsteilnahme, die bzgl. ihres Zutreffens auf einer siebenstufigen Skala einzuschätzen waren.

I.6.4 Ergebnisse

I.6.4.1 Deskriptive Statistiken

Die folgenden Tabellen beinhalten die relativen Häufigkeiten von *alt-* bzw. *ausgeführt-* Antworten, d.h. sowohl Rekognitions- als auch Quellengedächtnisurteile. Die statistische Auswertung entspricht derjenigen des ersten Experiments. Anmerkungen zu den eingesetzten Verfahren sind daher dem entsprechenden Abschnitt (I.5.3.1) zu entnehmen.

Tab. 22: Relative Häufigkeiten von *alt*-Antworten im Rekognitionstest des zweiten Experiments.

		Art der Verarbeitung in Phase 2:							
		Lesen				Beobachten			
		Art der Abrufmanipulation in Phase 3:							
		keine		keine		Quellenprüfung		Warnung	
		Art der Enkodierung in Phase 1: Ausgeführt							
		<i>M SD</i>		<i>M SD</i>		<i>M SD</i>		<i>M SD</i>	
Frequenz der Präsentation in Phase 2:	5x	.94	.12	.97	.07	1.00	.00	.96	.09
	1x	.94	.09	.97	.07	.97	.07	.99	.05
	0x	.91	.15	.91	.15	.93	.13	.91	.15
		Art der Enkodierung in Phase 1: Gelesen							
		<i>M SD</i>		<i>M SD</i>		<i>M SD</i>		<i>M SD</i>	
Frequenz der Präsentation in Phase 2:	5x	.66	.20	.94	.09	.90	.17	1.00	.00
	1x	.59	.33	.84	.26	.90	.17	.89	.20
	0x	.40	.25	.50	.22	.40	.25	.46	.29
		Art der Enkodierung in Phase 1: Nicht enkodiert							
		<i>M SD</i>		<i>M SD</i>		<i>M SD</i>		<i>M SD</i>	
Frequenz der Präsentation in Phase 2:	0x	.09	.06	.06	.06	.04	.03	.04	.05

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *alt*-Antworten relativiert an der Anzahl aller Handlungsanweisungen in der entsprechenden Zelle. $N=56$, d.h. 14 pro Gruppe.

Tab. 23: Relative Häufigkeiten von *ausgeführt*-Antworten im Quellengedächtnistest des zweiten Experiments.

		Art der Verarbeitung in Phase 2:							
		Lesen				Beobachten			
		Art der Abrufmanipulation in Phase 3:							
		keine		keine		Quellenprüfung		Warnung	
		Art der Enkodierung in Phase 1: Ausgeführt							
		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>	
Frequenz der Präsentation in Phase 2:	5x	.84	.21	.87	.17	.91	.13	.74	.28
	1x	.86	.12	.87	.17	.93	.13	.89	.15
	0x	.83	.21	.74	.21	.77	.23	.86	.20
		Art der Enkodierung in Phase 1: Gelesen							
		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>	
Frequenz der Präsentation in Phase 2:	5x	.09	.15	.29	.28	.29	.20	.23	.17
	1x	.04	.09	.23	.17	.21	.20	.27	.24
	0x	.04	.09	.06	.09	.03	.07	.07	.19
		Art der Enkodierung in Phase 1: Nicht enkodiert							
		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>		<i>M</i> <i>SD</i>	
Frequenz der Präsentation in Phase 2:	0x	.01	.02	.02	.02	.00	.01	.01	.03

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *ausgeführt*-Antworten relativiert an der Anzahl aller Handlungsanweisungen in der entsprechenden Zelle. *N*=56, d.h. 14 pro Gruppe.

I.6.4.2 Manipulation-Checks

Da die erste Phase der ersten Sitzung nun unter vollständiger Kontrolle der Versuchsleiterin war, war es nicht mehr notwendig, die Handhabung der Items anhand verschiedener Parameter abzusichern. Um jedoch mögliche Auswirkungen der Modifikation des Prozederes der ersten Phase erfassen zu können, wurden die Daten dennoch hinsichtlich eines enactment-Effekts analysiert. Es zeigte sich – wie auch in Experiment I bzw. der Voruntersuchung zu Experiment II – sehr deutlich, dass ausgeführte Handlungsanweisungen besser rekogniziert wurden als nicht ausgeführte, d.h. gelesene Handlungsanweisungen, obwohl die Häufigkeit des Ausführens nun auf einmal beschränkt war ($M(SD)$: .95 (.07) vs. .71 (.18), $t(55)=11.41$, $p<.05$, $\eta^2=.703$).

Die Einschätzung des instruktionsgemäßen Verhaltens in der zweiten Phase lag im Schnitt – über alle Gruppen hinweg – bei 5.50 ($SD=1.39$) und war für alle Gruppen vergleichbar; $F(3,48)=.57$, $p>.05$.

Da die Variation des Antwortkriteriums auf bewährten Operationalisierungen beruhte, die in der Gedächtnispsychologie standardmäßig eingesetzt werden, wurden im Vorhinein keine Indikatoren festgelegt, um eine gelungene Manipulation zu überprüfen. Im Rahmen der Diskussion wird jedoch darauf eingegangen, inwiefern eine erfolgreiche Manipulation an den Daten festgemacht werden kann.

I.6.4.3 Hypothesenprüfung

Zunächst wurde überprüft, ob ein inflation-Effekt – wie vorhergesagt – zwar in allen Gruppen, die die Handlungsausführung in der zweiten Phase beobachtet hatten, nicht jedoch in der Gruppe, die in der zweiten Phase nur bedeutungslose visuelle Stimuli beobachtet hatte („Lesengruppe“), auftrat.

In der Tat ließ sich nach der Beobachtung eines bedeutungslosen visuellen Stimulus kein inflation-Effekt nachweisen ($t(13)=1.15$, $p>.05$), nach der Beobachtung der Handlungsausführung ohne anschließende Abrufmanipulation, mit Instruktion zur Quellenprüfung und mit Warnung bei Abruf hingegen schon (ohne Abrufmanipulation: $t(13)=3.89$, $p<.05$, $\eta^2=.538$; mit Instruktion zur Quellenprüfung: $t(13)=5.83$, $p<.05$, $\eta^2=.723$; mit Warnung: $t(13)=2.35$, $p<.05$, $\eta^2=.298$).

In einem nächsten Schritt wurde das Ausmaß der inflation-Effekte zwischen den Bedingungen verglichen. Als Kontrollgruppe für die Beobachtungsgruppe ohne Abrufmanipulation diente die Lesengruppe. Es zeigte sich eine signifikante Interaktion zwischen der Präsentationsfrequenz in Phase 2 (5x vs. 0x) und der Art der Verarbeitung in dieser Phase (Beobachten der Handlungsausführung vs. Beobachten eines bedeutungslosen visuellen Stimulus): $t(26)=2.67$, $p<.05$, $\eta^2=.215$. Demnach ließ sich der observation inflation-Effekt auch nach Kontrolle des Lesens nachweisen.

Keine signifikante Interaktion ergab sich zwischen der Beobachtungsgruppe mit Instruktion zur Quellenprüfung vs. ohne Abrufmanipulation und der Präsentationsfrequenz (5x vs. 0x): $t(26)=.39$, $p>.05$. Das Ausmaß des Effekts war also in beiden Gruppen vergleichbar.

Dasselbe zeigte sich beim Vergleich der Gruppe, die die Warnung erhalten hatte, mit der Kontrollgruppe ohne Abrufmanipulation: $t(26)=.80$, $p>.05$. Auch die Warnung übte also keinen signifikanten Einfluss auf das Ausmaß des observation-inflation-Effekts aus.

Da erneut – aufgrund unsystematischer Ausfälle – nicht alle Kombinationen innerhalb sowie zwischen den Gruppen gleich häufig eingesetzt werden konnten, sollte der Einfluss des Merkmals „Materialkombination“ auf das Ausmaß falscher *ausgeführt*-Antworten untersucht werden. Die Analyse basiert auf den drei Beobachtungsgruppen, da innerhalb dieser überhaupt nur inflation-Effekte gefunden wurden. Es zeigten sich – wie bereits in Experiment I – keine signifikanten Effekte der Materialkombination auf falsche *ausgeführt*-Antworten nach fünfmaliger ($F(8,33)=.80$, $p>.05$) bzw. nullmaliger ($F(8,33)=.95$, $p>.05$) Präsentation.

I.6.4.4 Weitere Befunde

Deskriptivstatistisch unterscheiden sich die verglichenen Bedingungen hinsichtlich der korrekten Erinnerungsleistungen sowohl im Rekognitions- als auch im Quellengedächtnistest nicht wesentlich voneinander; so existieren (bis auf eine Ausnahme) auch keine Haupteffekte der verschiedenen Verarbeitungsformen (Rekognition, ausgeführte Items als alt erinnert: $F(3,52)=.60$, $p>.05$; Rekognition, gelesene Items als alt erinnert: $F(3,52)=6.68$, $p<.05$, $\eta^2=.278$; Quellengedächtnis, ausgeführte Items als ausgeführt erinnert: $F(3,52)=.39$, $p>.05$). Der signifikante Haupteffekt bzgl. der Rekognition von in Phase 1 gelesenen Items geht bei näherer Betrachtung darauf zurück, dass das Ausmaß rekognizierter Items in der Lesengruppe geringer war als in den Beobachtungsgruppen (es ergab sich kein signifikanter Unterschied, wenn nur die drei Beobachtungsgruppen in die Analyse einbezogen wurden: $F(2,39)=.33$, $p>.05$). Eine Gegenüberstellung der Lesen- und der Beobachtungsgruppe ohne

Abrufmanipulation zeigte, dass signifikante Unterschiede zwischen den Bedingungen nur nach fünf- und einmaliger, aber nicht nach nullmaliger Präsentation in Phase 2 auftraten (5x: $t(26)=4.86$, $p<.05$, $\eta^2=.476$; 1x: $t(26)=2.29$, $p<.05$, $\eta^2=.168$; 0x: $t(26)=1.13$, $p>.05$).

Auch wenn die Bedingung mit Warnung sich bzgl. falscher Erinnerungen an die Handlungsausführung nicht von der Bedingung ohne Warnung (bzw. Abrufmanipulation) unterschied, zeigt das Ergebnismuster doch zwei Auffälligkeiten. In allen bisherigen Untersuchungen gab es nach Beobachtung auch in den korrekten *ausgeführt*-Antworten einen der Präsentationsfrequenz folgenden Anstieg: Nach fünfmaligem Beobachten der Videos wurden mehr korrekte *ausgeführt*-Antworten gegeben als nach nullmaligem Beobachten, ein Befund der u.a. zu der Manipulation der Abrufinstruktion geführt hatte. Auffälligerweise kehrte sich dieses Muster in der Bedingung mit Warnung jedoch um: Nach nullmaliger Videopräsentation in Phase 2 wurden in dieser Bedingung mehr Handlungsanweisungen korrekt als ausgeführt erinnert als nach fünfmaliger Präsentation! Diese Interaktion (Beobachten mit Warnung beim Abruf vs. ohne Abrufmanipulation; Frequenz 5x vs. 0x) war statistisch signifikant; $t(26)=2.07$, $p<.05$, $\eta^2=.141$. Zudem lag im Rahmen der zweiten Untersuchung zwar die Rate an falschen *ausgeführt*-Antworten nach einmaliger durchgängig höher als nach nullmaliger Präsentation (siehe auch Voruntersuchung zur Studie II; in Studie I hatte es einen solchen Unterschied nicht gegeben; vgl. weiter unten), war aber stets niedriger als diejenige nach fünfmaliger Präsentation – außer in der Bedingung mit Warnung. In dieser Bedingung wurden mehr falsche *ausgeführt*-Antworten nach einmaliger im Vergleich zu fünfmaliger Beobachtung gegeben. Die Interaktion (Beobachten mit Warnung beim Abruf vs. ohne Abrufmanipulation; Frequenz 5x vs. 1x) ließ sich jedoch nicht statistisch absichern; $t(26)=1.05$, $p>.05$.

Um die Möglichkeit vollständig auszusräumen, dass der Anstieg falscher *ausgeführt*-Antworten lediglich auf den Anstieg korrekter *alt*-Antworten zurückgehen könnte, wurde eine Interaktion zwischen der Frequenz (5x vs. 0x) und der Art des Gedächtnistests (Rekognition vs. Quellengedächtnis) für jede (Beobachtens-)Bedingung einzeln berechnet. Diese Interaktionen wurden durchweg signifikant und zeigen, dass der Anstieg der falschen *ausgeführt*-Antworten disproportional zum Anstieg der *alt*-Antworten ist (Beobachten ohne Abrufmanipulation: $t(13)=2.79$, $p<.05$, $\eta^2=.374$; Beobachten mit Instruktion zur Quellenprüfung: $t(13)=2.97$, $p<.05$, $\eta^2=.405$; Beobachten mit Warnung: $t(13)=5.01$, $p<.05$, $\eta^2=.659$).

Nach Betrachtung auf Personenebene zeigte sich folgendes Muster: Von 14 TeilnehmerInnen in jeder Gruppe wiesen 3 in der Lesenbedingung, 9 in der Beobachtungsbedingung ohne Abrufmanipulation, 12 in der Beobachtungsbedingung mit Instruktion zur Quellenprüfung und 11 in der Beobachtungsbedingung mit Warnung einen inflation-Effekt auf. Dabei weicht die Ver-

teilung der TeilnehmerInnen in der Bedingung „Beobachten ohne Abrufmanipulation“ signifikant von derjenigen in der Bedingung „Lesen ohne Abrufmanipulation“ ab ($\chi^2(1,28)=5.25, p<.05, \Phi^2=.187$). Keine signifikanten Abweichungen ergaben sich beim Vergleich der Bedingungen „Beobachtung mit Instruktion zur Quellenprüfung“ und „Beobachtung ohne Abrufmanipulation“ ($\chi^2(1,28)=1.71, p>.05$) bzw. „Beobachtung mit Warnung“ und „Beobachtung ohne Abrufmanipulation“ ($\chi^2(1,28)=.70, p>.05$).

Im Folgenden sind die Sicherheitsratings bzgl. falscher *ausgeführt*-Antworten in den einzelnen Bedingungen dargestellt.

Tab. 24: Sicherheitsratings (siebenstufige Skala) bzgl. falscher *ausgeführt*-Antworten im zweiten Experiment.

Art der Verarbeitung in Phase 2												
Lesen						Beobachten						
Art der Abrufmanipulation in Phase 3												
keine			keine			Quellenprüfung			Warnung			
Frequ.	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
5x	6	5.67	1.51	20	5.25	1.59	20	5.40	1.43	16	5.44	1.15
1x	3	5.67	1.15	16	5.50	1.37	15	5.07	1.62	19	5.21	1.55
0x	3	5.67	2.31	4	4.00	1.54	2	5.00	2.83	5	3.60	1.82
∅	12	5.67	1.50	40	5.23	1.49	37	5.24	1.53	40	5.10	1.52
	<i>N</i>	<i>f</i> _{%(7)}	<i>f</i> (7)	<i>N</i>	<i>f</i> _{%(7)}	<i>f</i> (7)	<i>N</i>	<i>f</i> _{%(7)}	<i>f</i> (7)	<i>N</i>	<i>f</i> _{%(7)}	<i>f</i> (7)
5x	6	30%	2	20	20%	4	20	30%	6	16	19%	3

Anmerkungen: *N*=Anzahl an Fehlern; *f*_{%(7)}=Prozentuale Häufigkeit der höchsten Ratingstufe (7).

Diese Ratings zeigen insb. zwei Auffälligkeiten: In den Beobachtungsbedingungen steigt die Sicherheit bzgl. falscher *ausgeführt*-Antworten mit der Frequenz an, während sie in der Lesenbedingung konstant bleibt. Im Vergleich der verschiedenen Beobachtungsbedingungen ist v.a. zu konstatieren, dass die Warnung den größten Unterschied in der Sicherheit der Erinnerung fünf- vs. nullmalig präsentierter Handlungsanweisungen produziert. Der geringste Unterschied zwischen den Frequenzstufen und gleichzeitig die größte Sicherheit im Sinne des Anteils an mit absoluter Sicherheit beurteilten Erinnerungen findet sich in der Beobachtungsbedingung mit Instruktion zur Quellenprüfung. Die Gruppe, die nach Beobachtung keine Abrufmanipulation erhielt, ist bzgl. der Sicherheit zwischen diesen beiden anzusiedeln.

Die fälschlich als *ausgeführt* erinnerten Handlungsanweisungen verteilten sich relativ ungleichmäßig über die Pools und Subpools, allerdings bei ungleich häufig eingesetzten Materialkombinationen. Die beiden am häufigsten fälschlicherweise als *ausgeführt* erinnerten Handlungsanweisungen lauteten in Studie II „Tun Sie das Ei in den Eierbecher!“ mit neun Fehlattributionen, „Nehmen Sie einen Becher vom Stapel!“ mit acht Fehlattributionen und „Klappen Sie

die Sonnenbrille zusammen!“ mit sieben Fehlattributionen. Sechs der insgesamt 45 Handlungsanweisungen hingegen wurden gar nicht falsch erinnert (z.B. „Mischen Sie die Karten!“, „Spitzen Sie den Bleistift an!“, „Stellen Sie die Uhr eine Stunde vor!“, „Ziehen Sie den Handschuh an!“). Erneut ließ sich aus meiner Sicht keine Systematik erkennen – weder existierte ein verbindendes Element unter den häufig falsch erinnerten, noch unter den nie falsch erinnerten Handlungen. Auch erschienen die nach den verschiedenen Arten der Verarbeitung bzw. der Abrufmanipulation falsch erinnerten Handlungen nicht systematisch unterschiedlich. Allein auffällig war, dass zwei Handlungsanweisungen sowohl in Studie I als auch in Studie II niemals fälschlich als ausgeführt erinnert wurden – nämlich „Mischen Sie die Karten!“ und „Spitzen Sie den Bleistift an!“.

Wie bereits angesprochen, lag die Häufigkeit falscher *ausgeführt*-Antworten im Rahmen der 2. Studie bereits nach einmaliger Präsentation deutlich höher als nach nullmaliger. Eine explorative Analyse zeigte sogar, dass sich in allen drei Beobachtungsbedingungen das Ausmaß an falschen Erinnerungen nach einmaliger Präsentation signifikant von der nullmaligen, jedoch nicht signifikant von der fünfmaligen Präsentation unterschied. Mit anderen Worten wurde die Erinnerung an die Handlungsausführung schon durch die einmalige Videobeobachtung bedeutend verzerrt, und erfuhr durch die fünfmalige Präsentation keine bedeutsame Steigerung (1x vs. 0x: Beobachtung ohne Abrufmanipulation: $t(13)=3.71$, $p<.05$, $\eta^2=.514$; Beobachtung mit Instruktion zur Quellenprüfung: $t(13)=3.79$, $p<.05$, $\eta^2=.525$; Beobachtung mit Warnung: $t(13)=3.37$, $p<.05$, $\eta^2=.467$; 5x vs. 1x: Beobachtung ohne Abrufmanipulation: $t(13)=.84$, $p>.05$; Beobachtung mit Instruktion zur Quellenprüfung: $t(13)=1.05$, $p>.05$; Beobachtung mit Warnung: $t(13)=.64$, $p>.05$).

Die Quoten falscher *ausgeführt*-Antworten auf neue, d.h. in der ersten Sitzung überhaupt nicht präsentierte, Handlungsanweisungen, fielen in allen Gruppen vergleichbar niedrig aus. Auf eine inferenzstatistische Auswertung soll aufgrund der Bodeneffekte verzichtet werden.

Nach dem Gedächtnistest in der zweiten Sitzung wurde die Einschätzung einiger Items (u.a. bzgl. des Befolgens der Instruktionen) auf einer siebenstufigen Skala erbeten (siehe Tabelle 25, folgende Seite). Item 4 und 5 bezogen sich dabei auf den vorausgegangenen Gedächtnistest und erforderten eine generelle Einschätzung der Sicherheit, dass eine Handlung tatsächlich ausgeführt worden (Item 4) bzw. nur gelesen worden (Item 5) sei. Es stellte sich die Frage, ob anhand dieser Items auch auf Ebene des subjektiven Empfindens ein enactment-Effekt im Sinne höherer Sicherheit nach Ausführen als nach Lesen nachgewiesen werden könnte. Dies war allerdings nur für die Bedingungen Lesen ($t(13)=3.63$, $p<.05$, $\eta^2=.504$) und Beobachten mit Instruktion zur Quellenprüfung ($t(13)=3.74$, $p<.05$, $\eta^2=.518$) der Fall. Nach Warnung verschwand dieser Effekt auch auf deskriptiver Ebene vollständig. Zwischen den Bedin-

gungen Beobachten und Lesen ohne Abrufmanipulation fanden sich deskriptivstatistisch große Unterschiede hinsichtlich folgender Einschätzungen, die z.T. auch inferenzstatistisch abgesichert werden konnten: Nach Beobachten im Vergleich zum Lesen wurde eher ein Gedächtnistest für die Handlungsanweisungen erwartet. Nach Lesen waren die TeilnehmerInnen sich jedoch sicherer, eine Handlung tatsächlich ausgeführt zu haben. Beim Gedächtnistest dachten sie auch seltener an die bedeutungslosen visuellen Stimuli aus Phase 2 und befürchteten auch weniger eine Gedächtnisverzerrung durch diese Stimuli. Zwischen den Bedingungen Beobachten mit Instruktion zur Quellenprüfung und Beobachten ohne Abrufmanipulation ergab sich nur ein deutlicher Unterschied: Nach Quellenprüfung waren die TeilnehmerInnen sich sicherer, welche Handlungen sie tatsächlich ausgeführt hatten. Zwischen der Bedingung mit Warnung und der Vergleichsbedingung ohne Abrufmanipulation ergaben sich gar keine bedeutsamen Abweichungen – auch das Ausmaß an Gedanken über die Videos sowie die Befürchtung einer Gedächtnisverzerrung durch die Videos wurde vergleichbar eingeschätzt.

Tab. 25: Einschätzung verschiedener Items (siebenstufige Skala) im zweiten Experiment.

	Art der Verarbeitung in Phase 2			
	Lesen		Beobachten	
	Art der Abrufmanipulation in Phase 3			
	keine	keine	Quellenprüfung	Warnung
	<i>M(SD)</i>	<i>M(SD)</i>	<i>M(SD)</i>	<i>M(SD)</i>
Item 1	3.00 (1.71) ¹	4.93 (1.98) ¹	3.79 (2.46)	4.40 (1.96)
I expected that my memory for the action statements would be tested.				
Item 2	2.57 (1.45)	3.21 (1.58)	2.79 (1.81)	3.50 (1.27)
I expected exactly this kind of memory test.				
Item 3	4.43 (1.16)	4.57 (1.34)	4.71 (1.33)	4.40 (0.84)
The memory task I just fulfilled was very easy.				
Item 4	5.14 (1.46)	4.14 (1.41)	5.21 (1.42)	4.20 (1.32)
While I was doing the memory task, I was very sure about which actions I actually performed.				
Item 5	3.50 (1.45)	3.86 (1.56)	4.14 (1.35)	4.20 (1.69)
While I was doing the memory task, I was very sure about which actions I merely read.				
Item 6	4.64 (1.08)	4.21 (1.31)	4.29 (1.14)	4.90 (1.20)
My memory for the first phase of the study (i.e. when I performed or merely read the action statements) is very clear and vivid.				
Item 7	3.29 (1.94) ¹	6.00 (0.78) ¹	5.79 (1.81)	5.50 (1.51)
I thought about the video-observation task while I was doing the memory task.				
Item 8	3.21 (1.89) ¹	4.79 (1.67) ¹	4.36 (2.13)	4.71 (1.33)
When I was doing the memory task, I feared that my memory for action performance could be distorted by the videos I observed later.				

Anmerkungen: $N=56$, d.h. 14 pro Gruppe, außer in der Bedingung „Beobachten mit Warnung“. In dieser ist $n=10$ für alle Items, außer Items 7 und 8 ($n=14$). 1: $p < .05$ (zweiseitig) beim Vergleich der beiden Gruppen „Beobachten vs. Lesen ohne Abrufmanipulation“.

Zwischen der Einschätzung der berichteten Items und dem Ausmaß des observation-inflation-Effekts ergaben sich keine systematischen und konsistenten Zusammenhänge. Eine sehr hohe Korrelation soll jedoch nicht verschwiegen werden: Die Einschätzung des Items 6, "My memory for the first phase of the study (i.e. when I performed or merely read the action statements) is very clear and vivid.", korrelierte signifikant und sehr hoch mit dem Ausmaß des observation-inflation-Effekts in der Gruppe, die zur Quellenprüfung instruiert wurde ($r(14)=.807$, $p<.05$, $r^2=.651$). In den anderen beiden Gruppen war die Korrelation weniger stark bzw. auch von umgekehrter Richtung (Beobachten ohne Abrufmanipulation: $r(14)=-.450$, $p>.05$; Beobachten mit Warnung: $r(14)=.234$, $p>.05$).

I.6.5 Diskussion

I.6.5.1 Diskussion unter methodischen Gesichtspunkten

Anhand von Studie II konnte der observation-inflation-Effekt unter Einsatz eines im Vergleich zu Studie I verschlankten Designs repliziert werden. Theoretisch bedeutsam war dabei jedoch die Variation des Antwortkriteriums beim Abruf innerhalb dieses verschlankten Designs. Diese zeigte, dass der Effekt nicht vermeidbar war – weder durch eine Instruktion zur Quellenprüfung noch durch eine beim Abruf gegebene explizite Warnung vor dem Effekt, also durch keine Variation des Antwortkriteriums, ließ sich das Ausmaß falscher Erinnerungen beeinflussen.

Bevor auf die theoretischen und praktischen Implikationen dieses Befundes eingegangen werden soll, sollen verschiedene Aspekte der Validität der Hypothesenprüfung beleuchtet werden.

Die theoretische Hypothese wurde adäquat operationalisiert und in Testhypothesen überführt, die statistisch valide überprüft wurden. Die Testhypothesen können allesamt als angenommen gelten, da die vorhergesagten Unterschiede statistisch nachgewiesen werden konnten bzw. sich keine Unterschiede ergaben, wenn diese auch nicht erwartet worden waren. Dabei sind die erwarteten Effekte nicht nur signifikant ($p<.05$), sondern auch substantiell ($\eta^2 \geq .14$).

Wiederum wurde eine möglichst strenge, valide und präzise Hypothesenprüfung angestrebt. Die vorgenommenen Modifikationen – z.B. die vollständige Kontrolle der ersten Phase durch die Versuchsleiterin – dienten ebenfalls diesem Ziel. Potentiell problematische Aspekte sollen im Folgenden beleuchtet werden.

Zunächst zeigten Ratings, dass sich die TeilnehmerInnen auch in der zweiten Phase, in der sie im Gegensatz zu Phase 1 nicht durchgehend kontrolliert bzw. beobachtet wurden, stark be-

mühten, die Instruktionen wie gewünscht zu befolgen. Es ergab sich allerdings wiederum das Problem möglicher Materialeffekte aufgrund eines ungleich häufigen Einsatzes der verschiedenen Materialkombinationen innerhalb sowie zwischen den jeweiligen Gruppen. Diese Möglichkeit konnte jedoch sowohl argumentativ (siehe Experiment I, Abschnitt I.5.4.1) als auch empirisch ausgeräumt werden. Aufgrund der auch schon für Studie I beschriebenen Maßnahmen zur Kontrolle von allgemeinen und speziellen Störeffekten kann insgesamt davon ausgegangen werden, dass die interne Validität gesichert war.

Durch die Modifikation der Lesenbedingung wurde diese deutlich vergleichbarer mit der Beobachtungsbedingung – z.B. existierte nun nicht mehr das Problem unterschiedlicher Häufigkeiten der verschiedenen Formen der Verarbeitung (Lesen, Beobachten) der Handlungsanweisungen in Phase 2. Auch die drei Beobachtungsbedingungen in Studie II unterschieden sich nur durch die zusätzlichen Instruktionen bei der Abfrage und waren ansonsten vollkommen identisch, weshalb keine weiteren Begründungen der Vergleichbarkeit der vier interindividuell variierten Bedingungen vorgenommen werden müssen, während dies in Studie I noch problematischer erschien, da sehr unterschiedliche Verarbeitungsformen, die dann auch unterschiedliche Verarbeitungshäufigkeiten aufwiesen, miteinander verglichen wurden.

Die zusätzlichen Abrufmanipulationen wurden an bereits eingesetzte Instruktionen angelehnt und im Rahmen von Vortests auf ihre Verständlichkeit hin überprüft und modifiziert. Auf empirischer Ebene zeigte sich jedoch hinsichtlich derjenigen Items, die augenscheinlich Maße für eine gelungene Manipulation darstellen könnten, kein Effekt. In allen drei Beobachtungsbedingungen wurden – unabhängig von der jeweiligen Abrufmanipulation – gleich hohe Ratings bzgl. des Ausmaßes, in dem beim Abruf an die Videos gedacht wurde, sowie bzgl. des Ausmaßes, in dem eine Gedächtnisbeeinflussung befürchtet wurde, abgegeben. Es wäre indes erwartbar gewesen, dass insb. die Warnung, aber auch schon die Instruktion zur Quellenprüfung, die ja für mögliche Einflüsse durch die Videobedingung sensibilisierte, die Einschätzung in eine positive Richtung hätte beeinflussen können. Ein Grund für das Ausbleiben dieses Effekts mag darin liegen, dass es sich bei Erinnerungen an Handlungen um eine besondere Form von Erinnerungen handelt, die generell mit einem hohen Sicherheitsgefühl einhergehen. Daher könnten TeilnehmerInnen, auch wenn sie eine Warnung bzw. Instruktion zur Quellenprüfung erhielten, die Gefahr einer möglichen Beeinflussung unterschätzen. Gleichzeitig zeigten sich bei der Einschätzung des Ausmaßes, in dem die TeilnehmerInnen an die Videos dachten, in allen Gruppen Deckeneffekte, die möglicherweise (mit)verantwortlich für den fehlenden Nachweis eines Effekts der Abrufmanipulation sein könnten (allerdings existiert noch nicht einmal eine Tendenz in diese Richtung). Andere Indikatoren legen jedoch stark nahe, dass die Instruktion zur Quellenprüfung und die Warnung in der gewünschten Weise verstanden und

umgesetzt wurden. In der Bedingung mit Quellenprüfung fand sich der stärkste observation-inflation-Effekt, was v.a. darauf zurückzuführen ist, dass die Rate falscher *ausgeführt*-Antworten nach nullmaliger Präsentation im Vergleich aller Bedingungen am niedrigsten war. Dies spiegelte sich ferner darin wider, dass in dieser Gruppe die Quote an TeilnehmerInnen, die einen observation-inflation-Effekt aufwiesen, am höchsten war. Auch waren die wenigen *ausgeführt*-Antworten nach null- und fünfmaliger Präsentation von vergleichbarer Sicherheit begleitet. Diese Befunde weisen darauf hin, dass die TeilnehmerInnen – wie instruiert – intensiv nach qualitativen Erinnerungsspuren suchten und diese auch streng evaluierten; bloße Vertrautheit bspw. reichte nicht als Indikator für eine Handlungsausführung aus. Somit attribuierten sie nur sehr wenige Handlungsanweisungen, die nicht mit qualitativen Merkmalen angereichert worden waren, fehl. Wenn sie dies taten, taten sie es mit einer hohen Sicherheit, so dass hier Verwechslungen mit außerhalb des Experiments liegenden Quellen angenommen werden können. Dieses Vorgehen könnte auch der Grund für die hohe positive Korrelation zwischen der eingeschätzten Klarheit und Lebendigkeit der Erinnerung und dem Ausmaß des observation-inflation-Effekts in dieser Bedingung sein. In dieselbe Richtung lässt sich die Tatsache interpretieren, dass diese TeilnehmerInnen – wie a posteriori erhobene Ratings zeigten – generell sicherer waren, eine Handlungsanweisung ausgeführt als gelesen zu haben und auch im Vergleich zur Bedingung ohne Abrufmanipulation eine höhere Sicherheit bzgl. der Handlungsausführung angaben. In der Bedingung mit Warnung existiert ein ganz eindeutiger Hinweis auf die Wirkung derselben: Es handelt sich um die einzige Beobachtungsbedingung überhaupt, in der nach fünfmaliger im Vergleich zu nullmaliger Präsentation weniger *korrekte* *ausgeführt*-Antworten abgegeben wurden. Dies lässt sich dahingehend interpretieren, dass die TeilnehmerInnen ihre Erinnerungsurteile bzgl. der Handlungsausführung auf andere Kriterien basierten bzw. andere Hinweisreize als diagnostisch für die Handlungsausführung heranzogen als bspw. die TeilnehmerInnen, die keine Abrufmanipulation erhielten. Gleichzeitig ist diese Bedingung die einzige, innerhalb derer nach einmaliger im Vergleich zu fünfmaliger Präsentation mehr falsche *ausgeführt*-Antworten gegeben wurden; auch dies kann als Zeichen eines veränderten Antwortkriteriums angesehen werden. Zudem zeigen die Sicherheitsratings bzgl. falscher *ausgeführt*-Urteile eine große Differenz zwischen fünf- und nullmaliger Präsentation; die globale Beurteilung der Sicherheit bzgl. der Erinnerung an ausgeführte vs. nicht ausgeführte Handlungen (Items 4 und 5) ergab keinen Unterschied. Diese Befunde legen nahe, dass die TeilnehmerInnen sehr vorsichtig vorgehen und ihre Erinnerungen sehr kritisch evaluierten. Die angesprochenen Aspekte werden im Rahmen der theoretischen Diskussion noch näher zu diskutieren sein; insgesamt ist aber anzunehmen, dass die Manipulationen beim Abruf in der gewünschten Art und Weise aufgenommen wurden und ihre Wirkung entfalteten.

Des Weiteren konnte eindeutig gezeigt werden, dass der Anstieg der falschen *ausgeführt*-Antworten kein bloßes Abbild des Anstieges an *alt*-Antworten ist, der Effekt also mit anderen Worten kein methodologisches Artefakt darstellt.

Die Originaldaten von vier TeilnehmerInnen gingen verloren; glücklicherweise existierten jedoch Aufzeichnungen der Ergebnisse des Gedächtnistests, so dass hier – abgesehen von einem Großteil der Ratings der zusätzlich erhobenen Items – kein Verlust zu beklagen war. Da die Daten auf dieselbe Art und Weise gewonnen wurden, wie die übrigen Daten, ist hiermit auch kein Problem der Vergleichbarkeit der Datengewinnung verbunden. Anders sieht dies für die TeilnehmerInnen aus, denen der Gedächtnistest aufgrund der versagenden Technik in paper-pencil-Form vorgelegt werden musste. Allerdings gibt es keine Anhaltspunkte dafür, dass die analoge im Vergleich zur digitalen Abfrage einen Einfluss auf die Erinnerungsleistung ausgeübt hätte – die Ergebnisse beider Verfahren waren absolut vergleichbar. Zusätzlich wurden die analog gewonnenen Daten nur einer einzigen Teilnehmerin in die Auswertung einbezogen, weshalb das Problem als vernachlässigbar angesehen wird.

Bzgl. der Überlegungen zur Konstrukt-, Populations- und Situationsvalidität kann auf das entsprechende Kapitel der ersten Studie verwiesen werden. Gegenüber der ersten Untersuchung ist allerdings positiv anzumerken, dass die hier untersuchte Stichprobe weniger selektiv war, da Studierende verschiedener Fachrichtungen an der Untersuchung teilnahmen. Da dasselbe Ergebnismuster erzielt wurde, stellt dies einen Hinweis auf die Ausweitbarkeit der Ergebnisse auf die Population junger Erwachsener dar, wie sie bereits in Abschnitt I.5.4.2 vorgeschlagen wurde.

Insgesamt kann also davon ausgegangen werden, dass die interne sowie die Ableitungsvalidität auf den verschiedenen Ebenen der Untersuchung gegeben war; bzgl. der externen Validität bestehen auch für Studie II die üblichen Einschränkungen eines Laborexperiments, das an einer angefallenen Stichprobe durchgeführt wurde. Demnach kann die Hypothese

Beim observation-inflation-Effekt handelt es sich zumeist nicht um das Resultat einer Antworttendenz bzw. eines laxen Antwortkriteriums, sondern um eine a posteriori kaum vermeidliche Gedächtnistäuschung.

vor dem Hintergrund der beschriebenen Einschränkungen als vorläufig bewährt gelten.

I.6.5.2 Diskussion unter theoretischen und praktischen Gesichtspunkten

Die Ergebnisse der zweiten Studie zeigen zweierlei. Erstens konnte der innerhalb des ersten Experiments erstmalig gefundene observation-inflation-Effekt anhand eines leicht abgewandelten Designs repliziert werden – falsche Erinnerungen fanden sich nach fünfmaliger im Vergleich zu nullmaliger Beobachtung der Handlungsausführung, während die wiederholte Beobachtung bedeutungsloser visueller Stimuli keinen Anstieg falscher Erinnerungen hervorrief. Zweitens zeigte sich nach Beobachtung der Handlungsausführung kein Einfluss einer Quellenprüfung oder einer nachträglichen Warnung auf den Effekt – das Ausmaß falscher Erinnerungen war nach Quellenprüfung/ Warnung mit dem Ausmaß falscher Erinnerungen ohne Abrufmanipulation vergleichbar! Ein Ergebnismuster, das sich auch in der Anzahl an Personen, die einen inflation-Effekt aufwiesen, widerspiegelt, und das perfekt mit der Vorhersage korrespondiert. Die Datenstruktur belegte zudem zweifelsfrei, dass es sich bei dem gefundenen Effekt nicht um ein methodologisches Artefakt handelt.

Auf theoretischer Ebene bedeutet dies einerseits, dass falsche Erinnerungen daran, eine Handlung ausgeführt zu haben, tatsächlich nicht durch bloße Vertrautheit mit den Handlungsanweisungen zustande kommen. Eine Anreicherung mit sensorischen und/ oder motorischen Informationen (z.B. durch Beobachtung), die beim Abruf als diagnostisch für die Handlungsausführung angesehen werden, scheint kritisch zu sein; eine Replikation der Ergebnisse der ersten Studie. Dieser qualitative Unterschied, der zwischen dem Beobachten bedeutungsloser visueller Stimuli vs. der Handlungsausführung besteht, spiegelt sich auch deutlich in den Sicherheitsurteilen wider – so zeichnet sich nach Beobachtung der bedeutungslosen visuellen Stimuli im Vergleich zu der Beobachtung der Handlungsausführung kein Einfluss der Frequenz der Präsentation auf die Sicherheit der Erinnerungsurteile ab. Auffällig ist allerdings, dass in dieser Bedingung überhaupt ein gewisses Ausmaß an falschen Erinnerungen auftrat, die zudem von hoher Sicherheit begleitet waren, während dies in der entsprechenden Bedingung des ersten Experiments (Bedingung Lesen) fast gar nicht der Fall war. Dies könnte sowohl auf die Parallelisierung der Verarbeitungsfrequenz anstatt der Verarbeitungsdauer in Phase 1 bzw. 2 als auch auf die geänderte Operationalisierung innerhalb der zweiten Phase zurückzuführen sein. Die nur noch einmalige Verarbeitung der Handlungsanweisungen innerhalb beider Phasen könnte zu einer weniger tiefen Gedächtnisspur führen und daher ein erhöhtes Konfusionspotential mit außerhalb der Untersuchung liegenden Quellen bergen (die hohe Sicherheit der falschen Erinnerungen legt nahe, dass diese auf qualitativen Erinnerungsspuren basieren). Auch ist denkbar, dass die Aufgabe, bedeutungslose visuelle Stimuli zu beobachten, nicht so kapazitätsbeanspruchend war, wie die Lesenaufgabe in Experiment I, so dass die TeilnehmerInnen die Handlungsanweisungen elaborierten, z.B. mit Vorstellungen anreicherten. Der Ver-

gleich des Ergebnismusters mit der Lesengruppen der zweiten Studie liefert diesbzgl. Hinweise. Es ist zu konstatieren, dass in Experiment I bei gleicher Rekognitionsquote ausgeführter Handlungsanweisungen mehr nicht ausgeführte Handlungsanweisungen rekogniziert wurden. Gleichzeitig wurden korrekte *ausgeführt*-Antworten seltener gegeben als in Experiment II. Demnach scheint das nur einmalige im Vergleich zum mehrmaligen Lesen in Phase 1 bzw. 2 zu einer weniger tiefen Erinnerungsspur zu führen (schlechtere Rekognition, schlechteres Quellengedächtnis für nicht ausgeführte Handlungen in Experiment II), aber gleichzeitig die Erinnerungsspuren, die diagnostisch für die Handlungsausführung sind, weniger zu überlagern (besseres Quellengedächtnis für ausgeführte Handlungsanweisungen in Experiment II, obwohl diese nur einmal ausgeführt wurden). Allerdings sind die beiden Rekognitionstests in Experiment I vs. II (Phasenzuordnung vs. alt-/neu-Diskrimination) nur schwerlich vergleichbar. Eine Betrachtung der Beobachtungsbedingungen des ersten, der Vor- und der Hauptuntersuchung des zweiten Experiments liefert weitere Anhaltspunkte bzgl. der Auswirkungen des verkürzten Designs bzw. der abgewandelten Handhabung der Items in der ersten Phase. Experiment I und die Voruntersuchung zu Experiment II unterscheiden sich v.a. im Untersuchungsdesign, das in Experiment II um eine Itemart (nämlich die in Phase 1 nicht enkodierten Items) gekürzt wurde. Die Vor- und Hauptuntersuchung der 2. Studie unterscheiden sich dann v.a. in der Handhabung der 1. Phase, bei der letztlich nicht mehr die Dauer, sondern die Häufigkeit der Verarbeitung einer Handlungsanweisung konstant gehalten wurde. Dabei ähneln sich die Ergebnismuster der Vor- und Hauptuntersuchung II mehr als die Ergebnismuster der Untersuchung I und Voruntersuchung II. Sowohl die Differenz zwischen fünf- und nullmaliger Präsentation ist in der Vor- und Hauptuntersuchung II deskriptivstatistisch größer als diejenige der ersten Untersuchung, aber auch nach einmaliger Präsentation findet sich innerhalb der Vor- und Hauptuntersuchung II ein Anstieg falscher *ausgeführt*-Antworten, der in Untersuchung I noch nicht zu Tage trat. Allerdings ist dieser Anstieg in der Vor- im Gegensatz zur Hauptuntersuchung II nicht signifikant, während in beiden Untersuchungen der Unterschied zwischen einmaliger und fünfmaliger Präsentation gleichermaßen insignifikant ist. Diese Vergleiche sind jedoch mit Vorsicht zu interpretieren, da auch diverse weitere Bedingungen, insb. zwischen Experiment I und II, aber auch zwischen Vor- und Hauptuntersuchung II, variierten – z.B. die Selektivität der Stichprobe, die Untersuchungssprache, das Intervall zwischen Enkodierung und Abruf usw. Bei aller Vorsicht der Interpretation zeichnet sich jedoch folgender Trend ab: Die Verkürzung des Designs scheint zu einem erhöhten Ausmaß falscher Erinnerungen insb. nach fünfmaliger Präsentation zu führen, die Modifikation der Handhabung scheint eine erhöhte Quote von falschen *ausgeführt*-Antworten insb. nach einmaliger Präsentation nach sich zu ziehen. Warum allerdings der verkürzte Materialumfang zu mehr Konfusionen nach fünfmaliger Präsentation führen sollte, ist eine offene Frage. Eine mögliche Erklärung wäre, dass TeilnehmerInnen sich intensiver auf das Material konzentrieren können und weniger ermüden; so

ließe sich auch erklären, warum bereits nach einmaliger Präsentation ein etwas stärkerer Effekt auftritt: Die einzelnen Präsentationen könnten weniger in der Masse untergehen. Dies könnte sich insbesondere dann negativ auswirken, wenn die Handlungen nur einmal ausgeführt wurden. Diese Vergleiche führen also zu einer ähnlichen Schlussfolgerung, wie sie schon für die Lesenbedingung formuliert worden war. Die letzte Annahme wird durch die Tatsache untermauert, dass die TeilnehmerInnen der Voruntersuchung, die ja die Handlungsanweisungen noch mehrfach ausführten respektive lasen, den Gedächtnistest (deskriptivstatistisch) leichter fanden, sich sicherer waren, welche Handlungen ausgeführt bzw. gelesen worden waren und angaben, über eine klarere und lebendigere Erinnerung für die erste Phase zu verfügen als die TeilnehmerInnen in der Beobachtungsbedingung der Hauptuntersuchung. Hierbei ist allerdings auch zu beachten, dass das Intervall in der Voruntersuchung nicht konstant gehalten werden konnte und im Schnitt kürzer war, und dass auch die Items noch leicht anders formuliert waren als in der Hauptuntersuchung; aus beiden Gründen wurde kein inferenzstatistischer Vergleich der Ratings der Vor- und Hauptuntersuchung vorgenommen. Ein signifikanter Zusammenhang zwischen dem Intervall und der Einschätzung der Items innerhalb der Voruntersuchung konnte allerdings nicht nachgewiesen werden.

Theoretisch interessant ist v.a. der Vergleich der Beobachtungsbedingungen mit bzw. ohne Manipulation beim Abruf, da dieser keine bloße Replikation darstellt, sondern neue Einsichten birgt. Es zeigte sich, dass weder die Instruktion, beim Erinnerungstest auf Hinweise für bzw. gegen die eigene Handlungsausführung zu fokussieren, noch die nachträgliche Warnung, dass das Beobachten der Handlungsausführung zu falschen Erinnerungen führen könnte, den observation-inflation-Effekt reduzieren konnte. Dieser Befund wurde zwar erwartet und stimmt mit demjenigen der Studie von Thomas und Bulevich (2006) überein, erstaunt aber dennoch besonders aufgrund der Tatsache, dass in jener Untersuchung Vorstellung und in dieser Beobachtung eingesetzt wurde. Beide Formen der Verarbeitung scheinen also fast unausweichlich zu falschen Erinnerungen zu führen (einschränkend ist allerdings anzumerken, dass Thomas und Bulevich (2006) nur die Instruktion zur Quellenprüfung, aber keine Warnung einsetzen). Anders formuliert ist der observation-inflation-Effekt nicht auf die Evaluation, sondern auf die Qualität der Erinnerungsspuren zurückzuführen, stellt also eine Gedächtnistäuschung im engeren Sinne dar. Dies wird durch die innerhalb der Bedingungen fast identischen Quoten falscher *ausgeführt*-Antworten auf neue, d.h. nur innerhalb des Gedächtnistests dargebotenen, Items unterstrichen. Auf diese Überlegungen soll im Folgenden näher eingegangen werden.

Thomas und Bulevich (2006) setzten in ihrer Untersuchung, in der anstelle von Beobachtung mit Vorstellung gearbeitet wurde, eine Instruktion zur Quellenprüfung ein, die fast identisch

mit der hier verwendeten war. Diese Instruktion sollte TeilnehmerInnen animieren, nach qualitativen Hinweisen für eine tatsächliche Handlungsausführung zu suchen (sich z.B. an taktile oder akustische Eindrücke zu erinnern) und die Erinnerungsurteile auf diese zu stützen, bewirkte jedoch bei den jungen TeilnehmerInnen keine Reduktion falscher Erinnerungen. Die AutorInnen nahmen daher an, dass diese auch ohne Instruktion zur Quellenprüfung dieselbe Strategie anwandten, falsche Erinnerungen aber trotzdem nicht vermeiden konnten, da die Erinnerungsspuren von Vorstellung und Ausführung so ähnlich waren, dass sie einfach nicht voneinander unterschieden werden konnten. Dieses Ergebnis konnte anhand von Beobachtung perfekt repliziert werden. Man könnte sogar davon sprechen, dass die zusätzliche Instruktion den observation-inflation-Effekt noch klarer zum Vorschein brachte: Falsche Erinnerungen nach nullmaliger, jedoch nicht nach fünfmaliger Präsentation wurden reduziert, so dass die Effektgröße diejenige in den anderen Beobachtungsbedingungen überstieg. Zudem war die Quote an TeilnehmerInnen, die den Effekt aufwiesen, in dieser Bedingung am höchsten – 12 von 14 Personen unterlagen dem Effekt. Da Handlungen generell mit großer Sicherheit erinnert werden und auch die falschen Erinnerungen stets mit subjektiv hoher Sicherheit einhergingen, wurde neben der Instruktion zur Quellenprüfung, die mutmaßlich eine bloße Verbalisierung der Strategie darstellte, die die TeilnehmerInnen sowieso einsetzten, eine Warnung vor dem Effekt ausgesprochen. Es wurde angenommen, dass die Gruppe mit Warnung ebenfalls stärker auf cues für die eigene Handlungsausführung fokussieren würde, jedoch aufgrund des Wissens um die Möglichkeit der Täuschung motivierter hierzu sein sollte als die Gruppe mit Instruktion zur Quellenprüfung. Während die Ergebnismuster der Bedingungen mit und ohne Instruktion zur Quellenprüfung nahezu identisch waren, gab es deutliche Unterschiede zwischen den Bedingungen mit und ohne Warnung. Diese Unterschiede belegen, dass die Warnung einen etwas anderen Prozess anstieß, als die Instruktion zur Quellenprüfung. Die Tatsache, dass sich das Ergebnismuster von *ausgeführt*-Antworten hinsichtlich der fünf- und nullmalig präsentierten, tatsächlich ausgeführten Handlungsanweisungen umkehrte, belegt, dass die TeilnehmerInnen ihr Antwortkriterium änderten. Sie schienen weniger visuelle und akustische Erinnerungsspuren, die ja nach fünfmaliger Präsentation der Videos besonders stark ausgeprägt waren und daher als verdächtig (Echterhoff, Groll & Hirst, 2007) angesehen werden konnten, sondern vielmehr andere Erinnerungsqualitäten, die ausschließlich auf die eigene Ausführung zurückgehen konnten – wie z.B. taktile Eindrücke –, als diagnostisch für die Handlungsausführung heranzuziehen. Hätte lediglich die Vorsicht bzw. Tendenz dominiert, die zusammen mit einem Video präsentierten Handlungsanweisungen als nicht ausgeführt zurückzuweisen, hätte die Quote von *ausgeführt*-Antworten nach nullmaliger Präsentation sich in der Warnungsbedingung nicht von derjenigen in der Bedingung ohne Warnung unterschieden. Dies ist jedoch nicht der Fall. Auch insgesamt lässt sich bzgl. der korrekten Erinnerungen an die Handlungsausführung kein negativer Effekt der Warnung verzeichnen. Im Schnitt über

die verschiedenen Frequenzstufen hinweg ergibt sich exakt die gleiche Quote an korrekten Erinnerungsurteilen wie in der Bedingung ohne Warnung. Trotz dieses offensichtlichen Shifts im Antwortkriterium konnten falsche Erinnerungen allerdings nicht reduziert werden¹². Derartige Gedächtnistäuschungen könnten demnach nicht nur auf lebendigen visuellen und akustischen Erinnerungsspuren, die ja gerade nach Warnung nicht als hinreichend für ein *ausgeführt*-Urteil angesehen werden konnten, sondern auch auf qualitativ andersartigen Erinnerungsspuren basieren. So ist es bspw. möglich, dass die TeilnehmerInnen gerade motorische Repräsentationen zur Grundlage ihrer Erinnerungsurteile machten. Bezeichnenderweise lag die Quote falscher *ausgeführt*-Antworten in der Gruppe mit Warnung nach einmaliger Präsentation gar höher als nach fünfmaliger und auch höher als in allen anderen Bedingungen. Dies lässt sich mutmaßlich darauf zurückführen, dass in der Warnung explizit davon gesprochen wurde, dass besonders das wiederholte Videobeobachten zu Erinnerungstäuschungen führen könnte; gleichzeitig erklärt dies die Umkehrung des Musters auch in den korrekten Erinnerungsurteilen. Da die einmalige Videobeobachtung keine solche „Bedrohung“ darstellte wie die fünfmalige, könnte bei diesen Items, bzgl. derer den TeilnehmerInnen die Beobachtung möglicherweise gar nicht mehr präsent war, ein laxeres Antwortkriterium angelegt worden sein, so dass sie in diesen Fällen v.a. aufgrund von lebendigen visuellen oder akustischen Merkmalen urteilten.

Die Tatsache, dass keine der beiden Abrufmanipulationen, jedoch insb. die Warnung, zu einer Reduktion falscher Erinnerungen führte, ist insofern erstaunlich, als die Beobachtung der fremden Handlungsausführung sich charakteristisch von der tatsächlichen eigenen Handlungsausführung unterscheidet. Diese Unterschiede beziehen sich v.a. auf das Spektrum und die Art der sensorischen Repräsentationen. So werden auf sensorischer Ebene bei der Beobachtung nur visuelle und akustische, keine taktilen Merkmale der Handlungsausführung enkodiert – die visuellen Merkmale sind dabei zwar von der Qualität her der Ausführung ähnlich, gleichzeitig jedoch durch die 2.-Person-Perspektive und das (im vorliegenden Falle) verkleinerte, zweidimensionale Format unähnlich. Die Geräusche sind bei der Beobachtung jedoch mit der tatsächlichen Ausführung nahezu identisch. Dennoch wurden viele Handlungsanweisungen fälschlich erinnert, die kaum mit Geräuschen verbunden sind, wie „Wickeln Sie das Band auf!“,

¹² Es ist allerdings zu konstatieren, dass sich das Ausmaß des observation-inflation-Effekts (definiert als Differenz falscher Erinnerungen nach fünf- bzw. nullmaliger Darbietung) deskriptivstatistisch verringerte. Möglicherweise wäre dieser Unterschied daher bei einer größeren Stichprobe auch inferenzstatistisch abzusichern gewesen. Da allerdings – wie im Folgenden ausgeführt – im Vergleich zu den anderen Bedingungen mehr falsche Erinnerungen nach einmaliger Videopräsentation produziert wurden, entsprechen sich die Quoten falscher *ausgeführt*-Antworten in den Beobachtungsgruppen bei einer von der Frequenz unabhängigen Betrachtung exakt. Daher wird – trotz der beschriebenen Reduktion nach fünfmaliger Präsentation – davon ausgegangen, dass die Bedingung mit Warnung mit den anderen Bedingungen hinsichtlich des Ausmaßes falscher Erinnerungen tatsächlich vergleichbar ist, wenn auch das Muster abweicht.

„Tun Sie das Ei in den Eierbecher!“ oder „Dehnen Sie das Gummiband!“). Auch wurden einige Handlungsanweisungen fälschlich als ausgeführt erinnert, bei denen die Objektperspektive bei eigener Ausführung (1.-Person-Perspektive) und Beobachtung einer anderen Person (2.-Person-Perspektive) stark voneinander abweicht, z.B. „Reißen Sie einen Zettel vom Block ab!“, „Schlagen Sie das Buch auf!“ oder „Nehmen Sie die CD aus der Hülle!“. Zusammen mit der Tatsache, dass die falschen Erinnerungen trotz veränderten Antwortkriteriums unvermeidlich waren, legen die Ergebnisse nahe, dass nicht primär eine Fehlattribution sensorischer (visueller und akustischer) Erinnerungsspuren, sondern möglicherweise auch motorische Repräsentationen, die durch beobachtungsinduzierte Simulationsprozesse hervorgerufen werden, eine kritische Rolle spielen könnten. Da beide Aspekte – motorische und sensorische Informationen – jedoch nicht direkt manipuliert wurden, lässt sich aus den vorliegenden Ergebnissen kein eindeutiger Schluss ziehen.

Unter praktischer Perspektive und bei aller Vorsicht bei der Generalisierung der Ergebnisse auf andere Situationen lassen die Befunde kaum Spielraum für die Hoffnung, dass der als groß zu klassifizierende Effekt durch metakognitive Strategien im Nachhinein verhindert werden kann.

I.6.5.3 Résumé und Ausblick

Im Rahmen der zweiten Studie konnte der observation-inflation-Effekt auch anhand eines modifizierten Paradigmas repliziert und erneut die Möglichkeit zurückgewiesen werden, dass er allein auf die Vertrautheit mit den Handlungsanweisungen zurückzuführen ist. Zusätzlich konnte ausgeschlossen werden, dass der Anstieg der falschen *ausgeführt*-Antworten bloß eine Folge des Anstiegs rekonzipierter Handlungsanweisungen darstellt. Besonders implikationsreich jedoch ist die Tatsache, dass nach Darbietung einer Instruktion zur Quellenprüfung und nach Warnung vor dem Effekt beim Abruf selbiger nicht reduziert werden konnte.

Während die Ergebnismuster der Gruppe mit und ohne Instruktion zur Quellenprüfung fast identisch waren, wich das Ergebnismuster der Gruppe mit Warnung derart von demjenigen der entsprechenden Kontrollgruppe ab, dass dies als Änderung des Antwortkriteriums interpretiert werden konnte. Die TeilnehmerInnen schienen weniger auf Basis von lebendigen visuellen und akustischen Details zu urteilen, als vielmehr auf Basis anderer für die Handlungsausführung diagnostischer Kriterien. Dies legt einerseits nahe, dass es sich beim observation-inflation-Effekt in der Tat um eine Gedächtnistäuschung und nicht bloß um ein Resultat laxer Antwortkriterien oder bestimmter Antworttendenzen (Quellen-Bias) handelt. Andererseits scheint es unwahrscheinlich, dass die nahezu unausweichliche Täuschung primär auf eine Fehlattribution sensorischer Eindrücke zurückgeht. So sind bspw. viele der fehlattribuierten

Handlungsanweisungen äußerst geräuscharm, so dass die falschen Erinnerungen durch visuelle Eindrücke hätten zustande kommen müssen, die ja aber gerade in der Bedingung mit Warnung nicht die primäre und alleinige Basis der Erinnerungsurteile darstellen konnten. Zusätzlich weichen diese bei eigener Ausführung und Beobachtung fremder Ausführung per Video in Perspektive und Dimensionalität charakteristisch voneinander ab, so dass es unwahrscheinlich erscheint, dass keine der Abrufmanipulationen zu einer Reduktion des Erinnerungsfehlers geführt hätte, wenn dieser tatsächlich alleinig bzw. vorrangig auf einer Fehlattribution von sensorischen, insb. visuellen, Eindrücken beruhen würde. Ein Einbezug auch motorischer Repräsentationen ist deutlich besser mit dem Ergebnismuster und der Unvermeidbarkeit der Gedächtnistäuschung vereinbar. Da beide Dimensionen – sensorische und motorische Informationen – jedoch im Rahmen der zweiten Studie nicht direkt manipuliert wurden, sollte dies im Rahmen der dritten Studie geschehen. Dieser unterlag daher die folgende Hypothese:

Der observation-inflation-Effekt kann sowohl auf einer Fehlattribution sensorischer als auch auf einer Fehlattribution motorischer Repräsentationen beruhen; die auf Simulation beruhenden motorischen Repräsentationen stellen jedoch zu meist den kritischeren Faktor dar.

Eine Reihe von Studien hat gezeigt, dass die Quellendiskrimination umso schwieriger ist, je perzeptuell ähnlicher die zu unterscheidenden Quellen sind bzw. je größer die sensorische Überlappung zwischen ihnen ist. So konnten bspw. Henkel, Franklin und Johnson (2000) zeigen, dass vorgestellte Ereignisse eher fälschlich als gesehen erinnert wurden, wenn TeilnehmerInnen während des Vorstellens die entsprechenden Geräusche vorgespielt wurden im Vergleich zum alleinigen Hören der Geräusche bzw. Vorstellen der Ereignisse (siehe auch Johnson, Foley & Leach, 1988; Johnson, Raye, Wang & Taylor, 1979). Daher soll im Rahmen der dritten Studie die sensorische Überlappung von Beobachtung und Ausführung sowohl maximiert als auch minimiert werden. Ließe sich hierdurch auch der Effekt maximieren bzw. minimieren, würde dies für eine bedeutende Rolle sensorischer Informationen sprechen. Des Weiteren soll die Wahrscheinlichkeit motorischer Simulation manipuliert werden, um gleichzeitig den Einfluss motorischer Repräsentationen auf das Ausmaß der Gedächtnistäuschung einschätzen zu können. Würden die motorischen Repräsentationen (wie postuliert) den kritischeren Faktor für das Zustandekommen des observation-inflation-Effekts darstellen, sollte die Variation dieser entscheidender sein als die Variation sensorischer Merkmale. Die genaue Umsetzung beider experimenteller Manipulationen ist dem nächsten Abschnitt zu entnehmen.

I.7 Experiment III

I.7.1 Versuchsplanung

I.7.1.1 Operationalisierung und Variablenvalidität

Anhand von Studie II konnte gezeigt werden, dass es sich beim observation-inflation-Effekt weder um ein methodologisches Artefakt noch um einen Effekt des Antwortkriteriums oder der Antworttendenz, sondern um einen echten Gedächtnisfehler handelt. Das Ergebnismuster lieferte Hinweise, dass dem Effekt nicht nur eine Fehlattribution von sensorischen Informationen, sondern (v.a.) auch eine Fehlattribution von motorischen Repräsentationen, hervorgerufen durch Simulation, zugrunde liegen könnte. Im Rahmen einer dritten Untersuchung sollten daher beide Mechanismen direkt manipuliert und hinsichtlich ihres Einflusses auf den observation-inflation-Effekt gegenübergestellt werden.

Wiederum wurde dasselbe verschlankte Paradigma mit seinen zwei intraindividuell variierten unabhängigen Variablen „Art der Enkodierung in Phase 1“ und „Frequenz der Verarbeitung in Phase 2“ sowie den abhängigen Variablen „Rekognitions-“, „Quellengedächtnis-“ und „Sicherheitsurteile“ eingesetzt. Daher wird bzgl. der Überlegungen zur Operationalisierung und Variablenvalidität auf die entsprechenden Kapitel der vorausgegangenen Studien verwiesen (siehe Abschnitte I.5.1.1 bzw. I.6.2.1).

Um die beiden postulierten Mechanismen gegenüberstellen zu können, sollten zusätzlich die folgenden zwei Bedingungen realisiert und jeweils mit einer Kontrollbedingung verglichen werden: Maximale Überlappung von sensorischen Informationen bei minimaler Wahrscheinlichkeit von Simulation und umgekehrt minimale Überlappung von sensorischen Informationen bei maximaler Wahrscheinlichkeit von Simulation. Um Übertragungseffekt zu vermeiden, sollten diese Manipulationen interindividuell vorgenommen werden.

Auf Ebene der sensorischen Eindrücke lässt sich feststellen, dass eigene und fremde Handlungsausführung hinsichtlich akustischer und visueller Informationen überlappen. Aufgrund der Videopräsentation war die Überlappung hinsichtlich olfaktorischer Reize ausgeschlossen; allerdings waren Letztere im Rahmen der vorliegenden Untersuchung sowieso kaum vorhanden. Während dies hinsichtlich der Geräusche keine Rolle spielt, unterscheidet sich die visuelle Perspektive bei der Beobachtung fremder vs. eigener Handlungsausführung. Bei manchen Handlungen dürften sich die beiden *Objekt*perspektiven allerdings nur minimal unterscheiden – z.B. ist es von der Objektansicht her irrelevant, ob ein weißer Plastikbecher selbst oder von

einer anderen Person von einem Stapel genommen wird, da das Objekt keine unterscheidbare Vorder- oder Rückseite hat. Natürlich unterscheiden sich hierbei Selbst- und Fremdausführung in der Perspektive auf die beobachtete *Person* bzw. auf die Position ihrer Arme/ Hände. Bei anderen Handlungen jedoch – wie z.B. dem Durchblättern eines Buches – ist neben der Personen- auch die Objektperspektive distinkt und kann beim Abruf als Hinweis für oder gegen die eigene Ausführung genutzt werden. Um also sensorische Eindrücke – womit v.a. visuelle Objekt- und Personeneindrücke gemeint sind – eigener und fremder Handlungsausführung maximal überlappen zu lassen, müsste die Handlungsausführung aus der 1.- (und nicht wie bisher aus der 2.-) Person-Perspektive gezeigt werden. Nicht nur würde die 1.- im Vergleich zur 2.-Person-Perspektive eine maximale Überlappung sensorischer Informationen eigener und fremder Handlungsausführung darstellen, auch würde (theoretisch) gleichzeitig die Wahrscheinlichkeit der Simulation minimiert, da der soziale Charakter der Situation reduziert würde. Für Menschen stellt die Beobachtung anderer Personen aus der 1.-Person- als einer Selbst-Beobachtungs-Perspektive keine natürliche Situation dar; diese Perspektive ist konstitutiv für die Selbst-Fremd-Differenzierung (z.B. Vogeley, May, Ritzl, Falkai, Zilles & Fink, 2004). Normalerweise werden andere Personen aus der 2.- oder 3.-Person-Perspektive beobachtet (frontale bzw. Seitenansicht, vgl. Abschnitt I.4). Demnach könnte man davon sprechen, dass die motorische Überlappung von Beobachtung und Ausführung nach Beobachtung aus der 1.-Person-Perspektive minimiert sein sollte, da keine oder kaum Simulation stattfinden sollte¹³. Mit dem Begriff „motorische Überlappung“ ist in diesem Zusammenhang also nicht die Überlappung der sichtbaren (Hand- und Arm-) Bewegungen, die ja bei Beobachtung aus der (für die eigene Ausführung typischen) 1.-Person-Perspektive gerade maximal sein sollte, gemeint, sondern die Überlappung motorischer Repräsentationen eigener und fremder Ausführung, die durch die Aktivität von Spiegelneuronen bei der Beobachtung eines Gegenübers hervorgerufen werden. Konkret sollen also Videos gezeigt werden, die in der Perspektive variieren (1.- vs. 2.-Person-Perspektive). Allerdings gibt es für die zugrundeliegende, aus dem theoretisch-empirischen Hintergrund ableitbare perspektivenspezifische Simulationsannahme keine fundierte empirische Evidenz. Es existiert eine Vielzahl an Befunden zur Simulation bei der Beobachtung eines Gegenübers (siehe z.B. die Metaanalyse von Grèzes & Decety, 2001), aber lediglich eine Untersuchung hat 1.- und 2.-Person-Perspektive hinsichtlich ihres Potentials zur Induzierung von Simulation verglichen. Maeda und KollegInnen (2002) fanden dabei sogar das gegenteilige Aktivationsmuster: einen größeren Effekt bei der Beobachtung aus der 1.- vs. 2.-Person-Perspektive. Allerdings nutzten sie keine typische Gegenüber-, sondern eine rotierte

¹³ Dies stellt einen gewissen Widerspruch zu den Befunden von Schütz-Bosbach et al. (2006) dar, die eine Aktivität des Spiegelneuronensystems nach Beobachtung aus der 1.-Person-Perspektive fanden. Allerdings war dies nur dann der Fall, wenn die beobachtete Hand explizit als fremde Hand klassifiziert wurde; zudem wurde keine Vergleichsperspektive realisiert.

Perspektive und zudem eine einfache Fingerbewegung und keine objektbezogene Handlung; daher können die Ergebnisse nicht auf die vorliegende Situation übertragen werden. Theoretisch sollte Simulation in sozialen Situationen, d.h. bei der Beobachtung Anderer aus der 2.- oder 3.-Person-Perspektive, maximal angestoßen werden, und nicht bzw. weniger bei der „Selbst“beobachtung bzw. der Beobachtung einer fremden Person aus der Ich-Perspektive (z.B. Maeda, Kleiner-Fisman & Pascual-Leone, 2002; aber siehe auch Belopolsky, Olivers & Theeuwes, 2008; Schütz-Bosbach et al., 2006). Durch den Vergleich beider Perspektiven sollte also nicht nur die Auswirkung der Variation sensorischer Überlappung, sondern gleichzeitig (jedenfalls theoretisch) die Auswirkung variiertes motorischer Überlappung geprüft werden. Um trotz des Lecks im theoretisch-empirischen Unterbau zu möglichst eindeutigen Schlussfolgerungen gelangen zu können, sollte noch eine weitere Bedingung eingeführt werden, die eine Minimierung sensorischer Überlappung von Beobachtung und Ausführung bei gleich bleibender Wahrscheinlichkeit von Simulation beinhaltet.

Wollte man eine Überlappung von akustischen Informationen eigener und fremder Handlungsausführung ausschalten, ließe sich dies nur über eine stumme Präsentation der Videos umsetzen, da Geräusche bei der Ausführung nicht vermieden werden können. Wollte man eine Überlappung von visuellen Objektinformationen ausschalten, bestünde eine Möglichkeit darin, TeilnehmerInnen zu bitten, die Handlungen mit geschlossenen Augen auszuführen. Beide Maßnahmen zusammengenommen würden bedeuten, dass bei der eigenen Handlungsausführung nur taktile und akustische, bei der Beobachtung der fremden Handlungsausführung nur visuelle Reize enkodiert würden. Eine Überlappung von sensorischen Informationen wäre also minimiert bzw. gar eliminiert (siehe auch Hornstein & Mulligan (2004), die die sensorische Überlappung eigener und fremder Handlungsausführung ebenfalls dadurch variierten, dass die Handlungen entweder mit geschlossenen Augen (maximale Distinktheit), mit offenen Augen, oder vor einem Spiegel (maximale Ähnlichkeit) ausgeführt werden sollten). Wenn jedoch die fremde Person wiederum aus der 2.-Person-Perspektive beobachtet würde, wären gleichzeitig die Voraussetzungen für Simulation geschaffen. Diese Bedingung könnte mit der „Standardbedingung“ – Handlungsausführung mit offenen Augen, Präsentation der Videos ebenfalls aus der 2.-Person-Perspektive, aber mit Ton – verglichen werden, um selektiv den Einfluss der Elimination sensorischer Überlappung bei gleichbleibender motorischer Überlappung zu untersuchen. Sicherlich sind andere Varianten der Elimination sensorischer Überlappung von Beobachtung und Ausführung denkbar. Z.B. hätte man die AkteurInnen bei der Aufzeichnung bitten können, die Handlungen mit Objekten bzw. ohne Objekte auszuführen, so dass allein die Motorik gefilmt würde, akustische und visuelle Objektinformationen vollständig eliminiert wären. Allerdings erschien es als im vorliegenden Zusammenhang zu aufwändig, AkteurInnen so zu schulen, dass sie die Handlungen mit und ohne Objekt exakt identisch aus-

führten, so dass sich beide Bedingungen nur durch die An-/ Abwesenheit des Objektes unterschieden und problemlos miteinander verglichen werden könnten. Auch andere denkbare Varianten – z.B. die TeilnehmerInnen zu bitten, die Handlungen ohne Objekte auszuführen, – scheiterten an ähnlichen Überlegungen zur Vergleichbarkeit (siehe auch Senkfor, 2008). Daher wurde letztlich die beschriebene Operationalisierung als derzeit bestmögliche gewählt, um den Effekt der Elimination der Überlappung sensorischer Eindrücke bei gleichbleibender Überlappung motorischer Eindrücke aufzeigen zu können.

Die Überlegungen sind in der nachstehenden Tabelle zusammengefasst. Dabei wird im Folgenden nicht von eliminiertes, sondern nur von geringer sensorischer Überlappung die Rede sein, da nicht ausgeschlossen werden kann, dass die TeilnehmerInnen selbst sensorische Eindrücke generierten, sich z.B. das Aussehen der Objekte vorstellten (siehe Abschnitt I.7.4).

Tab. 26: Charakterisierung der Untersuchungsbedingungen des dritten Experiments.

	Art der Handlungsausführung in Phase 1; Präsentation des Videos in Phase 2		
	Augen offen; 1.-Person-Perspektive mit Ton	Augen offen; 2.-Person-Perspektive mit Ton	Augen geschl.; 2.-Person-Perspektive ohne Ton
Überlappung sensor. Informationen eigener und fremder Handlungsausführung	hoch	mittel	gering
Sozialer Charakter der Situation/ Überlappung motorischer Informationen (?)	gering	hoch	hoch

Hierbei ist zu beachten, dass die gewählten Versprachlichungen (hoch, mittel, gering) nur sehr grobe Klassifizierungen darstellen. So soll bspw. nicht nahegelegt werden, dass die drei Stufen sensorischer Überlappung gleichabständig wären – sicherlich sind sich die beiden Bedingungen sog. hoher und mittlerer Überlappung ähnlicher als die Bedingungen sog. mittlerer und geringerer Überlappung. Synonym zu den Ausprägungen gering und hoch hätten auch die Begriffe minimiert und maximiert verwendet werden können. Da es zu der mittleren Ausprägung dann allerdings keine passende Begrifflichkeit gegeben hätte, wurde von diesen Bezeichnungen abgesehen.

Ein Ablaufschema ist der folgenden Tabelle zu entnehmen.

Tab. 27: Ablauf des dritten Experiments.

Phase	Variable	Ausprägungen	Art der Variation
Phase 1: Enkodierung	UV-A: Art der Enkodierung	ausgeführt gelesen nicht enkodiert	intraindividuell
Phase 2: Verarbeitung*	UV-B: Frequenz der Verarbeitung	5x 1x 0x	intraindividuell
	UV-C: Sensorische Überlappung/ sozialer Charakter	hoch/niedrig mittel/hoch gering/hoch***	interindividuell
Phase 3: Abruf**	AV-1: Rekognition	alt neu	--
	AV-2: Quellengedächtnis	ausgeführt nicht ausgeführt	--
	AV-3: Sicherheit	siebenstufige Skala	--

Anmerkungen: * 5 Minuten nach Phase 1; ** 2 Wochen nach Phase 2; *** Die geringe Überlappung kam (neben der Videopräsentation ohne Ton in Phase 2) auch durch das Ausführen mit geschlossenen Augen in Phase 1 zustande.

Auch wenn sich die gewählte Operationalisierung der höheren bzw. geringeren Überlappung motorischer Informationen von Ausführung und Beobachtung durch die Beobachtungsperspektive aus dem theoretisch-empirischen Hintergrund ableiten lässt, weist sie zwei klare Schwachstellen auf. Im Gegensatz z.B. zur Manipulation des Antwortkriteriums in Experiment II, die auf bewährten Operationalisierungen beruhte, und zur Manipulation der sensorischen Überlappung, die ebenfalls auf bewährten Operationalisierungen beruht sowie direkt beobachtbar und nachvollziehbar ist, konnte bei der Manipulation des Ausmaßes motorischer Simulation nicht auf bewährte Operationalisierungen zurückgegriffen werden. Auch gibt es (im Rahmen der vorliegenden Studie) weder direkte noch indirekte Indikatoren für eine gelungene Manipulation – abgesehen natürlich vom vorhergesagten Ergebnismuster, das (je nach vermutetem primären Mechanismus) stark unterschiedlich bzw. nahezu gegenläufig ausfällt (siehe auch Tabelle 26). Dies soll bei der Hypothesenformulierung dadurch berücksichtigt werden, dass der zweite Teil der Hypothese, der sich speziell auf die als kritisch angesehenen motorischen Repräsentationen bezieht, vorsichtiger formuliert werden soll als bisher. Da eine gelungene Manipulation motorischer Repräsentationen mit dem vorliegenden Design nicht überprüfbar ist und zudem bewährte Operationalisierungen fehlen, soll an die Stelle der Formulierung „die auf Simulation beruhenden motorischen Repräsentationen stellen jedoch zumeist den kritischeren Faktor dar“ die vorsichtiger Formulierung „die auf Simulation beruhenden motorischen Repräsentationen *scheinen* jedoch zumeist den kritischeren Faktor darzustellen“

treten. Vor diesem Hintergrund sollen die folgenden Bedingungen erfüllt sein, um die theoretische Hypothese

Der observation-inflation-Effekt kann sowohl auf einer Fehlattribution sensorischer als auch auf einer Fehlattribution motorischer Repräsentationen beruhen; die auf Simulation beruhenden motorischen Repräsentationen scheinen jedoch zumeist den kritischeren Faktor darzustellen.

annehmen zu können (die Notation bezieht sich auf den im folgenden Abschnitt präsentierten Versuchsplan):

- (1) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der ersten Phase die Handlungen mit offenen Augen ausführten und in der zweiten Phase die Handlungsausführung aus der 1.-Person-Perspektive mit Ton beobachteten.

$$H_1: \mu_{A2B1C1} > \mu_{A2B3C1}; H_0: \mu_{A2B1C1} \leq \mu_{A2B3C1}$$

- (2) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der ersten Phase die Handlungen mit offenen Augen ausführten und in der zweiten Phase die Handlungsausführung aus der 2.-Person-Perspektive mit Ton beobachteten.

$$H_1: \mu_{A2B1C2} > \mu_{A2B3C2}; H_0: \mu_{A2B1C2} \leq \mu_{A2B3C2}$$

- (3) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die TeilnehmerInnen in der ersten Phase die Handlungen mit geschlossenen Augen ausführten und in der zweiten Phase die Handlungsausführung aus der 2.-Person-Perspektive ohne Ton beobachteten.

$$H_1: \mu_{A2B1C3} > \mu_{A2B3C3}; H_0: \mu_{A2B1C3} \leq \mu_{A2B3C3}$$

- (4) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist signifikant höher, wenn die TeilnehmerInnen in der ersten Phase die Handlungen mit offenen Augen ausführten und in der zweiten Phase die Handlungsausführung aus der 2.-Person-Perspektive mit Ton beobachteten, als wenn sie in der ersten Phase die Handlungen mit offenen Augen ausführten und in der zweiten Phase die Handlungsausführung aus der 1.-Person-Perspektive mit Ton beobachteten.

$$H_1: \mu_{A2B1C2} - \mu_{A2B3C2} > \mu_{A2B1C1} - \mu_{A2B3C1}; H_0: \mu_{A2B1C2} - \mu_{A2B3C2} \leq \mu_{A2B1C1} - \mu_{A2B3C1}$$

- (5) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist nicht signifikant höher, wenn die TeilnehmerInnen in der ersten Phase die Handlungen mit offenen Augen ausführten und in der zweiten Phase die Handlungsausführung aus der 2.-Person-Perspektive mit Ton beobachteten, als wenn sie in der ersten Phase die Handlungen mit geschlossenen Augen ausführten

und in der zweiten Phase die Handlungsausführung aus der 2.-Person-Perspektive ohne Ton beobachteten.

$$H_0: \mu_{A2B1C2} - \mu_{A2B3C2} \leq \mu_{A2B1C3} - \mu_{A2B3C3}; H_1: \mu_{A2B1C2} - \mu_{A2B3C2} > \mu_{A2B1C3} - \mu_{A2B3C3}$$

I.7.1.2 Versuchsplananlage und Versuchsplan

Der Untersuchung lag also wieder die intraindividuell variierte, dreistufige Variable „Art der Enkodierung in der 1. Phase“ (UV-A: ausgeführt, gelesen, nicht enkodiert) und die ebenfalls intraindividuell variierte, dreistufige Variable „Frequenz der Präsentation in der 2. Phase“ (UV-B: 5x, 1x, 0x) zugrunde. Es wurden drei Gruppen benötigt, um die anvisierten Untersuchungsziele zu erreichen: Zwei Gruppen sollten die Handlungen – wie in den Vorgängerstudien – in der ersten Phase mit geöffneten Augen ausführen und die Videos anschließend mit Ton präsentiert bekommen. Dabei sollte die eine Gruppe Videos aus der 1.-, die andere aus der 2.-Person-Perspektive zu sehen bekommen (UV-C: hohe sensorischer Überlappung und geringer sozialer Charakter der Beobachtungssituation vs. mittlere sensorische Überlappung und hoher sozialer Charakter der Beobachtungssituation). Eine dritte Gruppe sollte die Videos ebenfalls aus der 2.-Person-Perspektive beobachten, jedoch ohne Ton, und die Handlungen in der ersten Phase mit geschlossenen Augen ausführen (UV-C: niedrige sensorische Überlappung und hoher sozialer Charakter der Beobachtungssituation). Als abhängige Variablen wurden wiederum Rekognitions- und Quellengedächtnisurteile sowie Sicherheitsratings bzgl. der Erinnerungsurteile erhoben. Es resultierte also die im Folgenden dargestellte dreifaktorielle Versuchsplananlage (VPL-A3, Hussy & Jain, 2002; siehe Tabelle 28, folgende Seite).

Die Umsetzung dieser Versuchsplananlage in den konkreten Versuchsplan (VPL3-RRR) basierte auf denselben Überlegungen wie in Studie I. Die neu hinzugekommene UV-C wurde interindividuell variiert, da Übertragungseffekte vermieden werden sollten.

Tab. 28: Versuchsplan des dritten Experiments.

				UV-C: Sensorische Überlappung von Ausführung (Phase 1) und Beobachtung (Phase 2)/ Sozialer Charakter der Beobachtungssituation (Phase 2)		
				Hoch/ Gering	Mittel/ Hoch	Gering/ Hoch
UV-A: Art der Enkodierung in Phase 1	ausgeführt	UV-B: Frequenz der Präsentation in Phase 2	5x	A ₁ B ₁ C ₁	A ₁ B ₁ C ₂	A ₁ B ₁ C ₃
			1x	A ₁ B ₂ C ₁	A ₁ B ₂ C ₂	A ₁ B ₂ C ₃
			0x	A ₁ B ₃ C ₁	A ₁ B ₃ C ₂	A ₁ B ₃ C ₃
	gelesen		5x	A ₂ B ₁ C ₁	A ₂ B ₁ C ₂	A ₂ B ₁ C ₃
			1x	A ₂ B ₂ C ₁	A ₂ B ₂ C ₂	A ₂ B ₂ C ₃
			0x	A ₂ B ₃ C ₁	A ₂ B ₃ C ₂	A ₂ B ₃ C ₃
	nicht enkodiert		0x	A ₃ B ₃ C ₁	A ₃ B ₃ C ₂	A ₃ B ₃ C ₃

I.7.1.3 Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang

Über die Größe der Effekte konnte nur spekuliert werden, da derartige Manipulationen bisher noch nicht vorgenommen worden sind. In Ermangelung anderer Anhaltspunkte wurde daher von derselben (hohen) Effektgröße ausgegangen wie in Experiment I bzw. II. Folgende Parameter wurden der Berechnung der optimalen Stichprobengröße zugrunde gelegt: $\eta^2=.14$; $\alpha=.05$ und $\beta=.20$. Vor diesem Hintergrund wurde ein optimaler Umfang von 18 Personen pro Gruppe angestrebt, so dass insgesamt Daten von 54 TeilnehmerInnen in die Auswertung einfließen sollten. Wiederum basierte die Versuchsplanung auf dem größtmöglichen Effekt, dem Vergleich der Frequenzen 5x und 0x, so dass auch diesmal die Bedingung 1x nicht in die hypothesenprüfende Auswertung einfluss.

I.7.1.4 Kontrolle von allgemeinen Störeffekten

Siehe Experiment I.

I.7.1.5 Kontrolle von speziellen Stör- und von Materialeffekten

Siehe Experiment I.

I.7.1.6 Ethische Gesichtspunkte

Siehe Experiment I.

I.7.2 Versuchsdurchführung

I.7.2.1 Stichprobe

Die TeilnehmerInnen wurden über die üblichen Kanäle (Aushänge, Mailinglisten usw.) auf die Untersuchung aufmerksam gemacht. 66 an der Universität zu Köln eingeschriebene Bachelor-Studierende der Psychologie nahmen an der ersten Untersuchungssitzung teil. Drei von ihnen erschienen nicht zur zweiten Sitzung. Von den verbliebenen 63 vollständigen Datensätzen gingen jedoch nur 54 in die Auswertung ein. Gründe für den Ausschluss der übrigen neun Personen von der Datenanalyse waren der Einsatz vorläufiger Programmversionen, die noch eine Änderung erfuhren (fünf TeilnehmerInnen), bzw. instruktionswidriges Verhalten (drei TeilnehmerInnen). Ein Datensatz musste aufgrund eines technischen Problems ausgeschlossen werden.

Die Stichprobe von 54 TeilnehmerInnen, 18 pro Gruppe, lässt sich folgendermaßen charakterisieren: Das Alter umfasste eine Spannweite von 19 bis 45 Jahren mit einem Mittelwert von 25.33 Jahren ($SD=6.29$). 39 Personen waren weiblichen, 15 Personen männlichen Geschlechts. Die Verteilung des Alters und des Geschlechts war zwischen den Gruppen vergleichbar.

I.7.2.2 Material

Da das gesamte Untersuchungsmaterial der zweiten Studie in Kanada verblieben war, um dort eine Stichprobe von dementen TeilnehmerInnen zu rekrutieren, wurden die Materialien neu zusammengestellt. Hierdurch ergaben sich einige wenige Änderungen in der Materialauswahl bzw. Formulierung der Handlungsanweisungen (vgl. Anhang C.1); die bereits erfolgreich eingesetzten Pools und Subpools blieben jedoch fast vollständig bestehen. Es wurden ebenfalls dieselben Kombinationen eingesetzt wie in Studie I (siehe Abschnitt I.5.2.2).

Aufgrund der neuen Materialien sowie des Bedarfs an zwei Perspektiven, wurden auch die Videos neu aufgenommen. Es wurde der gleiche männliche Akteur verwendet, der sich schon in den Studien I und II zur Verfügung gestellt hatte. Die Videos, die die Person aus der 2.-Person-Perspektive zeigten, glichen denjenigen aus Studie I bzw. II, d.h. neben einer Arbeits-

fläche und dem Objekt waren die Hände, die Arme sowie Teile des Torsos des Akteurs als eines Gegenübers sichtbar. Hingegen wurden die 1.-Person-Perspektive-Videos so aufgenommen, dass sie die Arbeitsfläche und die Hände sowie einen Teil der Arme aus der Ich-Perspektive des Akteurs zeigten, d.h. bei den Aufnahmen befand sich die Kamera an der Position des Kopfes. Beide Perspektiven wurden gleichzeitig gefilmt, so dass sie sich tatsächlich ausschließlich in der Perspektive unterschieden – die Handlungsausführung war vollkommen identisch (siehe Abbildung 12).



Abb. 12: Ausschnitte aus den Videos zum dritten Experiment. Die Handlungsausführung wurde aus der 1.- (links) bzw. 2.-Person-Perspektive (rechts) gezeigt.

I.7.2.3 Hilfsmittel und Geräte

Siehe Experiment I.

I.7.2.4 Ablauf

Die Untersuchung lief vollständig computergestützt ab.

Die beiden Bedingungen, in denen TeilnehmerInnen die Augen bei der Handlungsausführung in der ersten Phase geöffnet hielten und die Videos mit Ton verfolgten (hohe sensorische Überlappung, geringer sozialer Charakter; mittlere sensorische Überlappung, hoher sozialer Charakter), entsprachen in ihrem Ablauf den bereits genutzten Beobachtungsbedingungen. Die Variation der Videoperspektive stellte keine Änderung des Ablaufs, sondern nur des präsentierten Materials dar. Daher ist das Prozedere identisch mit demjenigen, das unter Studie II geschildert wurde, abgesehen von der Tatsache, dass als Fillertask zwischen Phase 1 und 2 erneut der NFC/FII-Fragebogen zum Einsatz kam und neben einigen untersuchungsbezogenen Aussagen am Ende auch wieder die Items des HZI und des FDS abgefragt wurden (siehe Studie I, Abschnitt I.5.2.4).

Die dritte Bedingung (geringe sensorische Überlappung, hoher sozialer Charakter) unterschied sich hauptsächlich bzgl. der ersten Phase von dem bisherigen Vorgehen: Anstatt – wie bisher – ein Photo des Gegenstandes auf dem Monitor zu präsentieren und selbigen sichtbar zu positionieren, um einen lebendigen Eindruck von dem Objekt zu erwecken, auch wenn dies später nicht gebraucht würde, wurden nur die Objektnamen eingeblendet. Hierdurch sollte jeglicher visueller Eindruck vom Objekt vermieden werden. Die TeilnehmerInnen wurden dann gebeten, die Augen zu schließen und sich – analog zu dem konkreten visuellen Eindruck, der ihnen sonst vermittelt wurde – einen konkreten taktilen Eindruck vom Objekt zu verschaffen, das ihnen durch die Versuchsleiterin in die Hände gelegt wurde. Hierdurch sollte vermieden werden, dass beim Abruf allein die Erinnerung an die Präsenz/ Absenz des Objektes in der Hand als Grundlage der Entscheidung herangezogen werden könnte. Nachdem die TeilnehmerInnen sich so einen Eindruck vom Objekt verschafft hatten, sollten sie dieses zurückgeben und die Versuchsleiterin verberg es wieder. Die Augen durften nun geöffnet werden. Dann wurden die TeilnehmerInnen – analog zum Vorgehen in den anderen Bedingungen – gebeten, einzuschätzen, an wie vielen Tagen einer gewöhnlichen Woche sie das befühlte Objekt nutzen würden. Danach erschien die Handlungsanweisung entweder zusammen mit der Instruktion „Bitte ausführen“ oder „Bitte lesen“. Im ersteren Falle wurden die TeilnehmerInnen gebeten, die Handlungsanweisung einmal zu lesen und dann wieder die Augen zu schließen. Anschließend wurde ihnen das Objekt erneut gereicht, die Handlung wurde einmal ausgeführt und das Objekt wieder außer Sichtweite gebracht, bevor die Augen wieder geöffnet werden durften. Im zweiten Falle wurden die TeilnehmerInnen ebenfalls gebeten, die Handlungsanweisung einmal zu lesen und die Augen zu schließen. Dann allerdings wurden sie mündlich instruiert, eine Kombination bedeutungsloser Handbewegungen durchzuführen. Diese entsprachen exakt den schon in Studie II eingesetzten Handbewegungen. Anstatt sie jedoch zu imitieren – was aufgrund der geschlossenen Augen nicht möglich war –, wurden die TeilnehmerInnen mündlich angeleitet. Des Weiteren wurde die Anzahl an Handbewegungen etwas erhöht; sie waren nun nicht mehr zwischen 3 und 5, sondern zwischen 5 und 7 Bewegungen lang. Diese Maßnahme trug der Tatsache Rechnung, dass die Handlungsausführung mit geschlossenen Augen etwas länger dauerte als mit offenen. Danach durften die Augen wieder geöffnet werden. Abgesehen von diesen Modifikationen der ersten Phase verlief die Untersuchung wie in Experiment II geschildert mit der einzigen weiteren Ausnahme, dass die Videos in der zweiten Phase ohne Ton präsentiert wurden.

Die Instruktionen sind dem Anhang C.2 zu entnehmen.

I.7.3 Ergebnisse

I.7.3.1 Deskriptive Statistiken

Die folgenden Tabellen beinhalten die relativen Häufigkeiten von *alt*- bzw. *ausgeführt*-Antworten, d.h. sowohl Rekognitions- als auch Quellengedächtnisurteile. Die statistische Auswertung entspricht derjenigen des ersten Experiments. Anmerkungen zu den eingesetzten Verfahren sind daher dem entsprechenden Abschnitt (I.5.3.1) zu entnehmen.

Tab. 29: Relative Häufigkeiten von *alt*-Antworten im Rekognitionstest des dritten Experiments.

		Sensorische Überlappung von Ausführung (Phase 1) und Beobachtung (Phase 2)/ Sozialer Charakter der Beobachtungssituation (Phase 2):					
		Hoch/ Gering		Mittel/ Hoch		Gering/ Hoch	
Art der Enkodierung in Phase 1: Ausgeführt							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	1.00	.00	1.00	.00	.97	.08
	1x	.99	.05	.97	.10	.94	.11
	0x	.89	.18	.92	.20	.97	.08
Art der Enkodierung in Phase 1: Gelesen							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.97	.08	.99	.05	.97	.08
	1x	.91	.12	.84	.13	.87	.21
	0x	.41	.24	.40	.27	.52	.34
Art der Enkodierung in Phase 1: Nicht enkodiert							
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	0x	.03	.06	.04	.05	.04	.05

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *alt*-Antworten relativiert an der Anzahl aller Handlungsanweisungen in der entsprechenden Zelle. $N=54$, d.h. 18 pro Gruppe.

Tab. 30: Relative Häufigkeiten von *ausgeführt*-Antworten im Quellengedächtnistest des dritten Experiments.

		Sensorische Überlappung von Ausführung (Phase 1) und Beobachtung (Phase 2)/ Sozialer Charakter der Beobachtungssituation (Phase 2):					
		Hoch/ Gering		Mittel/ Hoch		Gering/ Hoch	
		Art der Enkodierung in Phase 1: Ausgeführt					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.89	.14	.92	.12	.89	.14
	1x	.90	.12	.88	.16	.92	.12
	0x	.77	.20	.78	.29	.78	.23
		Art der Enkodierung in Phase 1: Gelesen					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.17	.21	.28	.21	.20	.24
	1x	.16	.16	.20	.24	.19	.22
	0x	.03	.08	.02	.06	.00	.00
		Art der Enkodierung in Phase 1: Nicht enkodiert					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	0x	.01	.01	.01	.02	.01	.01

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *ausgeführt*-Antworten relativiert an der Anzahl aller Handlungsanweisungen in der entsprechenden Zelle. $N=54$, d.h. 18 pro Gruppe.

I.7.3.2 Manipulation Checks

Da das Vorgehen in der ersten Phase der ersten Sitzung vollständig kontrolliert war, bestand kein Zweifel daran, dass die Handlungen wie instruiert gehandhabt wurden (siehe auch Studie II). Dies wurde auch durch das Vorhandensein eines enactment-Effekts validiert (Rekognition ausgeführter Handlungsanweisungen: .96 (.07) (M(SD)); Rekognition gelesener Handlungsanweisungen .76 (.13) (M(SD)); $t(53)=11.85$, $p<.05$, $\eta^2=.726$).

Diejenigen TeilnehmerInnen, die während der Handlungsausführung in der ersten Phase die Augen geschlossen halten sollten, wurden nach dem Gedächtnistest in der zweiten Sitzung noch einmal gefragt, wie sehr sie der Aussage zustimmten, dass sie sich während der ersten Phase keinen visuellen Eindruck von den Objekten bzw. von der Handlungsausführung bilden konnten. Auf einer siebenstufigen Skala wurden beide Aussagen gleichermaßen im Schnitt mit 6.94 ($SD=.25$) bewertet. Dieser Wert kommt dadurch zustande, dass eine Teilnehmerin eine 6 angab, während alle anderen die 7 wählten. Der Teilnehmerin, die das niedrigere Rating abgab, war bei der Handlungsausführung ein Objekt aus der Hand gefallen, weshalb sie die Augen öffnete. Diese einzige Ausnahme betrifft aber nicht die zentrale abhängige Variable, also die falschen *ausgeführt*-Antworten, da die Handlung ja eben ausgeführt wurde. Bzgl. der zweiten Phase gaben alle TeilnehmerInnen – unabhängig von der speziellen Bedingung – eine sehr hohe und vergleichbare Konzentration auf die Videos an ($M=6.28$, $SD=.86$; $F(2,51)=.52$, $p>.05$).

Insgesamt kann also von einer sehr guten compliance und somit Datenvalidität ausgegangen werden.

I.7.3.3 Hypothesenprüfung

Zunächst wurde überprüft, ob innerhalb der einzelnen Formen der Verarbeitung ein observation-inflation-Effekt auftrat. Dies war sowohl für die beiden Gruppen, die die Handlungen mit offenen Augen ausgeführt und die Videos aus verschiedenen Perspektiven mit Ton präsentiert bekommen hatten (hohe sensorische Überlappung, geringer sozialer Charakter; mittlere sensorische Überlappung, hoher sozialer Charakter), als auch für die Gruppe, die die Handlungen mit geschlossenen Augen ausgeführt und die Videos aus der 2.-Person-Perspektive ohne Ton präsentiert bekommen hatte (geringe sensorische Überlappung, hoher sozialer Charakter), der Fall (hohe sensorische Überlappung, geringer sozialer Charakter: $t(17)=2.75$, $p<.05$, $\eta^2=.308$; mittlere sensorische Überlappung, hoher sozialer Charakter: $t(17)=5.65$, $p<.05$, $\eta^2=.653$; geringe sensorische Überlappung, hoher sozialer Charakter: $t(17)=3.62$, $p<.05$, $\eta^2=.453$).

In einem zweiten Schritt wurde überprüft, ob sich das Ausmaß des observation-inflation-Effekts zwischen den Gruppen unterschied. Hierzu wurden Interaktionen zwischen der Frequenz der Verarbeitung (5x vs. 0x) und den jeweiligen Gruppen berechnet. Ein Vergleich der beiden Perspektiven zeigte, dass – unter ansonsten gleichen Bedingungen (Augen offen, Video mit Ton) – die Beobachtungssituation mit geringem sozialen Charakter (1.-Person-Perspektive) einen deutlich niedrigeren Effekt hervorrief als die Beobachtungssituation mit hohem sozialen

Charakter (2.-Person-Perspektive); $t(34)=1.84$, $p<.05$, $\eta^2=.091$. Demgegenüber hatte die Minimierung überlappender sensorischer Merkmale von Handlungsausführung und Beobachtung bei gleichbleibender Perspektive (mittlere sensorische Überlappung, hoher sozialer Charakter vs. geringe sensorische Überlappung, hoher sozialer Charakter) keine signifikante Reduzierung des Effekts zur Folge; $t(34)=-.73$, $p>.05$.

Da es in dieser Untersuchung gelungen war, jede Materialkombination innerhalb jeder Gruppe und damit auch zwischen den Gruppen gleich häufig einzusetzen, konnte auf eine Analyse des Einflusses der Materialkombination verzichtet werden.

I.7.3.4 Weitere Befunde

Univariate Varianzanalysen der korrekten Erinnerungsleistungen zeigten keinerlei Haupteffekte der Gruppenzugehörigkeit auf die Rekognitions- oder Quellengedächtnisleistung (Rekognition, ausgeführte Items als alt erinnert: $F(2,51)=.02$, $p>.05$; Rekognition, gelesene Items als alt erinnert: $F(2,51)=.45$, $p>.05$; Quellengedächtnis, ausgeführte Items als ausgeführt erinnert: $F(2,51)=.04$, $p>.05$).

Bis auf das beschriebene geringere Ausmaß des observation-inflation-Effekts nach Beobachtung der Handlungsausführung aus der 1.- verglichen mit der 2.-Person-Perspektive, zeigten sich demnach keine signifikanten Unterschiede zwischen diesen beiden Bedingungen. Bei einer differenzierteren Betrachtung derjenigen Gruppen, die die Handlungsausführung aus der 2.-Person-Perspektive beobachteten, ergab sich allerdings eine Besonderheit bzgl. der Rekognition: In bisher allen Studien zeigte sich auch bei der Rekognition der ausgeführten Handlungsanweisungen ein Anstieg mit steigender Frequenz. In der Gruppe, die die Augen während der Ausführung geschlossen hielt und die Videos ohne Ton präsentiert bekam (geringe sensorische Überlappung, hoher sozialer Charakter), war die Quote der rekognizierten ausgeführten Handlungsanweisungen nach fünfmaliger Präsentation jedoch genauso hoch wie nach nullmaliger. Bei einem Vergleich mit der Kontrollbedingung (mittlere sensorische Überlappung, hoher sozialer Charakter) ergab allerdings keine signifikante Interaktion; $t(34)=1.51$, $p>.05$.

Erneut wurde eine Interaktion zwischen der Frequenz (5x vs. 0x) und der Art des Gedächtnistests (Rekognition, Quellengedächtnis) für jede Gruppe einzeln berechnet, um die Möglichkeit auszuräumen, dass der Anstieg in den falschen *ausgeführt*-Antworten lediglich auf den Anstieg in korrekten *alt*-Antworten zurückzuführen sein könnte. Diese Interaktionen wurden durchweg signifikant; der Anstieg der falschen *ausgeführt*-Antworten war also disproportional zum Anstieg der *alt*-Antworten (hohe sensorische Überlappung, geringer sozialer Charakter:

$t(17)=5.58, p<.05, \eta^2=.647$; mittlere sensorische Überlappung, hoher sozialer Charakter: $t(17)=4.30, p<.05, \eta^2=.521$; geringe sensorische Überlappung, hoher sozialer Charakter: $t(17)=2.51, p<.05, \eta^2=.271$).

Auf Personenebene ergab sich folgendes Muster: Von 18 Personen in jeder Gruppe wiesen 8 in der Bedingung „hohe sensorische Überlappung, geringer sozialer Charakter“, 14 in der Bedingung „mittlere sensorische Überlappung, hoher sozialer Charakter“ und 10 in der Bedingung „geringe sensorische Überlappung, hoher sozialer Charakter“ einen observation-inflation-Effekt auf. Dabei wich die Verteilung der TeilnehmerInnen in der Bedingung „hohe sensorische Überlappung, geringer sozialer Charakter“ signifikant von derjenigen in der Bedingung „mittlere sensorische Überlappung, hoher sozialer Charakter“ ab ($\chi^2(1,36)=4.21, p<.05, \Phi^2=.117$). Keine signifikante Abweichung ergab sich beim Vergleich der Bedingungen „geringe sensorische Überlappung, hoher sozialer Charakter“ und „mittlere sensorische Überlappung, hoher sozialer Charakter“ ($\chi^2(1,36)=2.00, p>.05$).

Im Folgenden sind die durchschnittlichen Einschätzungen der Sicherheit bzgl. der falschen *ausgeführt*-Antworten in den einzelnen Bedingungen dargestellt.

Tab. 31: Sicherheitsratings (siebenstufige Skala) bzgl. falscher *ausgeführt*-Antworten im dritten Experiment.

Sensorische Überlappung von Ausführung (Phase 1) und Beobachtung (Phase 2)/ Sozialer Charakter der Beobachtungssituation (Phase 2):									
	Hoch/ Gering			Mittel/ Hoch			Gering/ Hoch		
Frequ.	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
5x	15	6.20	0.94	25	5.96	1.40	18	5.50	1.82
1x	14	5.93	1.00	18	5.28	2.19	17	5.41	1.50
0x	3	5.00	2.00	2	3.00	2.83	0	-	-
∅	32	5.97	1.09	45	5.56	1.88	35	5.46	1.65
	<i>N</i>	<i>f</i> _% (7)	<i>f</i> (7)	<i>N</i>	<i>f</i> _% (7)	<i>f</i> (7)	<i>N</i>	<i>f</i> _% (7)	<i>f</i> (7)
5x	15	53%	8	25	52%	13	18	39%	7

Anmerkungen: *N*=Anzahl an Fehlern; *f*_%(7)=Prozentuale Häufigkeit der höchsten Ratingstufe (7).

Ein Anstieg der Urteilssicherheit mit der Frequenz der Darbietung zeigte sich in allen drei Gruppen und ist in der Kontrollgruppe (Handlungsausführung mit offenen Augen, Videos aus 2.-Person-Perspektive mit Ton) am deutlichsten. Im Durchschnitt liegt die Sicherheit bei maximaler sensorischer Überlappung am höchsten und bei minimaler Überlappung am niedrigsten; bei teilweiser Überlappung liegt sie zwischen diesen beiden, jedoch näher an der Bedin-

gung mit nur geringer Überlappung.

Falsche *ausgeführt*-Antworten verteilten sich im Rahmen der dritten Studie etwas gleichmäßiger über die Handlungsanweisungen als in den ersten beiden Studien. Dies mag v.a. darauf zurückzuführen sein, dass nun alle Handlungsanweisungen gleich häufig eingesetzt worden waren. Die drei am häufigsten fälschlicherweise als ausgeführt erinnerten Handlungsanweisungen lauteten „Stecken Sie die Karte in den Umschlag!“ mit sieben Fehlattritionen und „Nehmen Sie einen Becher vom Stapel!“ bzw. „Reißen Sie einen Zettel vom Block ab!“ mit jeweils sechs Fehlattritionen. Sieben der insgesamt 45 Handlungsanweisungen wurden gar nicht falsch erinnert (z.B. „Mischen Sie die Karten!“, „Spitzen Sie den Bleistift an!“, „Schrauben Sie den Kugelschreiber auseinander!“, „Zeichnen Sie einen Kreis auf das Papier!“). Erneut ließ sich aus meiner Sicht keine Systematik erkennen – weder innerhalb noch zwischen den einzelnen Gruppen. Allerdings wurden die beiden Handlungsanweisungen, die bereits in Studie I und II niemals zu falschen Erinnerungen geführt hatten – nämlich „Mischen Sie die Karten!“ und „Spitzen Sie den Bleistift an!“ –, auch in dieser Studie nicht fälschlich als ausgeführt erinnert. Eine weitere Handlungsanweisung, die weder in Studie II noch III falsche Erinnerungen provozierte, lautete „Zeichnen Sie einen Kreis auf das Papier!“ (diese Aufforderung wurde in Experiment I nur als Beispiel eingesetzt). Die Handlungsanweisung „Nehmen Sie einen Becher vom Stapel!“ hingegen stellte sowohl in Studie II als auch in Studie III eine der am häufigsten fälschlich erinnerten dar.

Wie auch in Studie II lag die Häufigkeit falscher *ausgeführt*-Antworten nach einmaliger Präsentation in allen Gruppen höher als nach nullmaliger. Dieses Ergebnismuster ließ sich in einer explorativen Analyse auch inferenzstatistisch absichern: In allen drei Gruppen war das Ausmaß an falschen Erinnerungen nach einmaliger Präsentation signifikant höher als nach nullmaliger Präsentation; einmalige und fünfmalige Präsentation unterschieden sich jedoch nicht signifikant voneinander (1x vs. 0x: hohe sensorische Überlappung, geringer sozialer Charakter: $t(17)=2.65$, $p<.05$, $\eta^2=.292$; mittlere sensorische Überlappung, hoher sozialer Charakter: $t(17)=2.95$, $p<.05$, $\eta^2=.339$; geringe sensorische Überlappung, hoher sozialer Charakter: $t(17)=3.61$, $p<.05$, $\eta^2=.434$; 5x vs. 1x: hohe sensorische Überlappung, geringer sozialer Charakter: $t(17)=.16$, $p>.05$; mittlere sensorische Überlappung, hoher sozialer Charakter: $t(17)=1.24$, $p>.05$; geringe sensorische Überlappung, hoher sozialer Charakter: $t(17)=.20$, $p>.05$).

Die Quoten falscher *ausgeführt*-Antworten auf neue, d.h. nur im Gedächtnistest präsentierte, Items, waren in allen Gruppen vergleichbar gering. Aufgrund der Bodeneffekte soll auf eine inferenzstatistische Auswertung verzichtet werden.

Die Einschätzung einiger Items, die nach dem Gedächtnistest auf einer siebenstufigen Skala erfolgte, erbrachte die in der Tabelle 32 (folgende Seite) abgetragenen Mittelwerte. Item 4 und 5 bezogen sich – wie in Studie II – auf den vorausgegangenen Gedächtnistest und erforderten eine generelle Einschätzung der Sicherheit, dass eine Handlung tatsächlich ausgeführt worden (Item 4) bzw. nur gelesen worden sei (Item 5). Es sollte wieder überprüft werden, ob sich anhand dieser Items auch auf Ebene des subjektiven Empfindens eine Art enactment-Effekt – also eine größere Sicherheit für ausgeführte im Vergleich zu nicht ausgeführten Handlungen – zeigen würde. Dies war für alle drei Gruppen der Fall (hohe sensorische Überlappung, geringer sozialer Charakter: $t(17)=2.03$, $p<.05$, $\eta^2=.194$; mittlere sensorische Überlappung, hoher sozialer Charakter: $t(17)=3.19$, $p<.05$, $\eta^2=.374$; geringe sensorische Überlappung, hoher sozialer Charakter: $t(17)=1.84$, $p<.05$, $\eta^2=.167$). Zwischen den beiden Gruppen, die sich nur in der Beobachtungsperspektive unterschieden, fanden sich auf deskriptiver Ebene große Unterschiede hinsichtlich folgender Einschätzungen, die z.T. auch inferenzstatistisch abgesichert werden konnten: Nach Beobachtung aus der 1.- im Vergleich zur 2.-Person-Perspektive wurde eher ein Gedächtnistest für die Handlungsanweisungen bzw. eher der speziell eingesetzte Gedächtnistest erwartet. Die TeilnehmerInnen, die die Videos aus der 1.-Person-Perspektive beobachtet hatten, gaben des Weiteren an, sich sicherer zu sein, welche Handlungen sie tatsächlich ausgeführt hatten, und die erste Phase der Untersuchung klarer und lebendiger zu erinnern. Zwischen den beiden Gruppen, die die Videos aus der 2.-Person-Perspektive beobachteten, und eine mittlere vs. geringe Überlappung von sensorischen Merkmalen der eigenen vs. fremden Handlungsausführung aufwiesen, ergaben sich zwei deutliche Unterschiede: Die Videoqualität wurde von den TeilnehmerInnen in der Bedingung mit geschlossenen Augen und der Darbietung ohne Ton als schlechter eingeschätzt, sie gaben jedoch an, sich bei der Beobachtung stärker vorgestellt zu haben, wie sich die Handlungsausführung für die Person anfühlen würde.

Tab. 32: Einschätzung verschiedener Items (siebenstufige Skala) im dritten Experiment.

Sensorische Überlappung von Ausführung (Phase 1) und Beobachtung (Phase 2)/ Sozialer Charakter der Beobachtungssituation (Phase 2):			
	Hoch/ Gering	Mittel/ Hoch	Gering/ Hoch
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Item 1	5.06 (2.01)	3.67 (2.28)	4.39 (2.03)
Ich habe erwartet, dass meine Erinnerung an die Handlungsanweisungen getestet würde.			
Item 2	3.94 (1.98) ¹	2.33 (1.64) ¹	2.83 (1.54)
Ich habe genau diese Art von Gedächtnistest erwartet.			
Item 3	4.72 (1.27)	4.17 (1.20)	4.44 (1.10)
Der Gedächtnistest, den ich gerade gemacht habe, war sehr einfach.			
Item 4	5.39 (1.09)	4.50 (1.72)	4.72 (1.56)
Während ich den Gedächtnistest gemacht habe, war ich mir sehr sicher, welche Handlungen ich tatsächlich ausgeführt habe.			
Item 5	4.61 (1.69)	3.61 (1.94)	4.22 (1.59)
Während ich den Gedächtnistest gemacht habe, war ich mir sehr sicher, welche Handlungen ich nur gelesen habe.			
Item 6	5.56 (0.92) ¹	4.67 (1.19) ¹	5.06 (1.51)
Meine Erinnerung an die erste Phase der Untersuchung (als ich die Handlungsanweisungen ausgeführt bzw. nur gelesen habe) ist sehr klar und lebendig.			
Item 7	5.22 (1.70)	5.28 (1.56)	4.33 (1.84)
Während ich den Gedächtnistest gemacht habe, habe ich an die Videos gedacht, die ich gesehen habe.			
Item 8	4.78 (2.26)	4.83 (1.92)	4.33 (1.88)
Während ich den Gedächtnistest gemacht habe, habe ich befürchtet, dass meine Erinnerung an die Handlungsausführung durch das Beobachten der Videos verfälscht sein könnte.			
Item 9	4.56 (1.79)	4.72 (1.71)	4.11 (1.94)
Während ich den Gedächtnistest gemacht habe, war ich mir manchmal nicht sicher, ob ich selber eine Handlung ausgeführt oder nur beobachtet habe, wie die Person in den Videos die Handlung ausgeführt hat.			
Item 10	6.28 (0.89)	6.50 (0.79)	5.78 (1.52)
Die Videos, die in der Video-Beobachtungsaufgabe gezeigt wurden, waren visuell (und akustisch) von guter Qualität.			
Item 12	4.44 (2.06)	4.94 (1.63)	4.72 (1.60)
Die Video-Beobachtungsaufgabe war äußerst anstrengend.			
Item 13	2.61 (1.24)	2.22 (1.40)	2.83 (1.20)
Während ich die Videos beobachtet habe, habe ich mich in die Person hineinversetzt.			
Item 14	3.06 (1.76)	2.33 (1.28)	3.27 (1.60)
Während ich die Videos beobachtet habe, habe ich mir vorgestellt, wie sich die Handlungsausführung für die Person anfühlt.			
Item 15	3.88 (2.53) (<i>n</i> =8)	4.00 (2.83) (<i>n</i> =7)	4.67 (1.94)
Während ich die Videos beobachtet habe, habe ich meine eigene Handlungsausführung mit derjenigen der Person in dem Video abgeglichen.			

Fortsetzung Tab. 32: Einschätzung verschiedener Items (siebenstufige Skala) im dritten Experiment.

Sensorische Überlappung von Ausführung (Phase 1) und Beobachtung (Phase 2)/ Sozialer Charakter der Beobachtungssituation (Phase 2):		
Hoch/ Gering	Mittel/ Hoch	Gering/ Hoch
<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Item 16		3.28 (1.49)
Während ich die Videos beobachtet habe, habe ich mir vorgestellt, wie sich die Handlungsausführung anhört.		
Item 18		6.44 (0.78)
In der ersten Phase der Untersuchung (als ich die Handlungsanweisungen ausgeführt bzw. nur gelesen habe) habe ich mir einen visuellen Eindruck von den Objekten gebildet bzw. mir vorgestellt, wie die Objekte aussehen.		
Item 20		5.83 (1.50)
In der ersten Phase der Untersuchung (als ich die Handlungsanweisungen ausgeführt bzw. nur gelesen habe) habe ich mir während des Ausführens der Handlungen einen visuellen Eindruck von der Handlungsausführung gebildet bzw. mir vorgestellt, wie die Handlungsausführungen aussehen würden.		
Item 21		5.94 (1.30) (<i>n</i> =17)
Während ich die Videos beobachtet habe, habe ich meine Vorstellung von den Objekten mit den im Video gezeigten Objekten abgeglichen.		

Anmerkungen: *N*=54, d.h. 18 pro Gruppe, falls nicht anders angegeben. Die Items 16 bis 21 wurden allein der Gruppe „geringe sensorische Überlappung und hoher sozialer Charakter“ vorgelegt. 1: $p < .05$ (zweiseitig) beim Vergleich der Gruppen „hohe sensorische Überlappung, geringer sozialer Charakter“ vs. „mittlere sensorische Überlappung, hoher sozialer Charakter“.

Wiederum fanden sich keine systematischen und konsistenten Zusammenhänge zwischen der Einschätzung der berichteten Items sowie der Einschätzungen der Items des NFC/FII, HZI, FDS und dem Ausmaß des observation-inflation-Effekts.

I.7.4 Diskussion

I.7.4.1 Diskussion unter methodischen Gesichtspunkten

Der observation-inflation-Effekt trat zwar nach Beobachtung aus der 1.-Person-Perspektive auf, d.h. in einer Beobachtungssituation mit geringem sozialen Charakter und maximierter sensorischer Überlappung von Ausführung und Beobachtung, war jedoch im Vergleich zur Beobachtung aus der 2.-Person-Perspektive, d.h. in einer Beobachtungssituation mit hohem sozialen Charakter und nur partieller sensorischer Überlappung, reduziert. Wurden die TeilnehmerInnen gebeten, die Augen bei der Ausführung der Handlung zu schließen und die Videos in der 2. Phase aus der 2.-Person-Perspektive ohne Ton präsentiert, d.h. unter Bedin-

gungen minimierter, gar (objektiv) eliminierter sensorischer Überlappung von Ausführung und Beobachtung bei hohem sozialem Charakter der Beobachtungssituation, fand sich ebenfalls ein Effekt; dieser war vergleichbar mit demjenigen, der unter Bedingungen partieller sensorischer Überlappung von Ausführung und Beobachtung bei ebenfalls hohem sozialen Charakter der Beobachtungssituation auftrat.

Im Folgenden sollen verschiedene Aspekte der Validität der Hypothesenprüfung beleuchtet werden, bevor auf theoretische und praktische Implikationen dieses Befundes eingegangen werden soll.

Die theoretische Hypothese wurde vor dem existierenden theoretisch-empirischen Hintergrund adäquat operationalisiert (für Einschränkungen siehe unten), in Testhypothesen überführt und statistisch valide überprüft. Die Testhypothesen können allesamt als angenommen gelten, da die vorhergesagten Unterschiede statistisch nachgewiesen werden konnten bzw. sich keine Unterschiede ergaben, wenn diese auch nicht erwartet worden waren. Dabei waren die erwarteten Effekte signifikant ($p < .05$). Substantielle Effekte ($\eta^2 \geq .14$) zeigten sich bzgl. 3 der 4 aufgestellten Alternativhypothesen. Eine Interaktion erreichte nicht die vorgegebene Effektgröße von $\eta^2 \geq .14$, jedoch mit $\eta^2 = .091$ einen als groß zu klassifizierenden Effekt. Da bei der Ermittlung der optimalen Stichprobengröße auf keinerlei Vorläuferstudien zurückgegriffen werden konnte, konnte auch die erwartete Effektgröße nicht empirisch fundiert werden. Auch weichen gefundene und erwartete Effektgröße nur graduell voneinander ab. Aus diesen Gründen soll die (leichte) Abweichung die Interpretation der theoretischen Hypothese, die auf allen fünf Vergleichen fußt, nicht einschränken.

Bei der Planung und Durchführung der Untersuchung wurde das Ziel einer möglichst strengen, validen und präzisen Hypothesenprüfung verfolgt; so wurden bspw. die Materialpools noch einmal modifiziert, um eine höchstmögliche Parallelität zu erzielen. Dennoch lassen sich problematische Aspekte benennen, die im Folgenden beleuchtet werden sollen.

Es ist zunächst positiv festzuhalten, dass es mit der dritten Studie aufgrund geringerer Ausfallquoten und einer niedrigeren Gesamtstichprobengröße erstmals gelungen war, jede der neun Materialkombinationen sowohl innerhalb als auch zwischen den Bedingungen gleich häufig einzusetzen. Daher können Materialeffekte ausgeschlossen werden. Dies validiert auch im Nachhinein die sehr ähnlich ausgeprägten observation-inflation-Effekte innerhalb der Experimente I und II, bei denen ein möglicher Einfluss der Materialkombination nur argumentativ ausgeräumt und statistisch kontrolliert werden konnte. Allerdings erschienen insgesamt 6 TeilnehmerInnen einen Tag zu früh bzw. zu spät zur zweiten Sitzung. Vier von ihnen gehörten

der Gruppe an, die die Videos aus der 1.-Person-Perspektive beobachtete (hohe sensorische Überlappung, geringer sozialer Charakter), zwei der entsprechenden Kontrollbedingung (mittlere sensorische Überlappung, hoher sozialer Charakter). Da jeweils die Hälfte der TeilnehmerInnen zu früh bzw. zu spät kam und es sich außerdem nur um eine eintägige Verschiebung bei einem insgesamt zweiwöchigen Intervall handelte, wurde diese Variation als vernachlässigbar angesehen. Die Daten weisen auch auf keinerlei Zusammenhang dieses Faktors mit dem Ausmaß des inflation-Effekts hin (hohe sensorische Überlappung, geringer sozialer Charakter: $r(18)=-.118$, $p>.05$; mittlere sensorische Überlappung, hoher sozialer Charakter: $r(18)=.179$, $p>.05$). Des Weiteren waren die Bedingungen untereinander gut vergleichbar und variierten nur minimal, so dass es keiner Rechtfertigung der vorgenommenen Vergleiche bedarf. Allein bzgl. der beiden Gruppen, die die Videos aus der 2.-Person-Perspektive sahen, ließe sich argumentieren, dass eine bessere Vergleichbarkeit erzielt worden wäre, wenn nicht Modifikationen beider Phasen (Augen geöffnet/ geschlossen in Phase 1, Videos mit/ ohne Ton in Phase 2), sondern nur innerhalb einer Phase vorgenommen worden wären. Dies war allerdings nicht möglich, da sich bei eigener Ausführung nur die visuellen, nicht jedoch die akustischen Reize ausschalten ließen, und dasselbe Videomaterial genutzt werden sollte wie in der Vergleichsgruppe. Da sich jedoch keine signifikanten Haupteffekte der Bedingung hinsichtlich der korrekten Erinnerungsleistungen ergaben, kann von einer Vergleichbarkeit ausgegangen werden. Die Einschätzung der compliance zeigte, dass in allen Bedingungen die Instruktionen zur zweiten Phase wie gewünscht befolgt wurden. Auch unterstreichen die Einschätzungen für die Bedingung „Handlungsausführung mit geschlossenen Augen“ hinsichtlich der ersten Phase auf subjektiver Ebene eine gelungene Umsetzung – die TeilnehmerInnen gaben an, in dieser Phase keinerlei visuelle Eindrücke der Objekte bzw. der Handlungsausführung gehabt zu haben. Die bereits für Studie I beschriebenen Maßnahmen zur Kontrolle von allgemeinen und speziellen Störeffekten wurden auch im Rahmen der dritten Studie umgesetzt bzw. noch optimiert. Daher kann davon ausgegangen werden, dass die interne Validität gesichert war.

Bzgl. der Überlegungen zur Konstrukt-, Populations- und Situationsvalidität kann auf das entsprechende Kapitel der ersten Studie zurückgegriffen werden (siehe Abschnitt I.5.4). Ob durch die Realisierung der drei Gruppen tatsächlich die gewünschten (sensorischen vs. motorischen) Informationen manipuliert werden konnten, soll im Rahmen der theoretischen Diskussion erörtert werden. An dieser Stelle soll nur festgehalten werden, dass zwar die Manipulation motorischer Repräsentationen weder direkt noch (abgesehen natürlich vom Ergebnismuster) indirekt erfasst und zudem nicht auf bewährte Operationalisierungen basiert werden konnte, die Manipulation der sensorischen Überlappung von Beobachtung und Ausführung hingegen direkt beobachtbar und an bereits existierende Manipulationen angelehnt war (Hornstein & Mulligan, 2004). Die Einschätzungen der TeilnehmerInnen, auf die auch im Rahmen der theoretischen

Diskussion noch eingegangen werden wird, zeigten, dass alle drei Bedingungen der Videobebachtung als gleich anstrengend empfunden wurden. Des Weiteren wurde die Qualität derjenigen Videos, die aus den beiden unterschiedlichen Perspektiven mit Ton präsentiert wurden, als vergleichbar eingeschätzt. Hingegen wurden die Videos, die ohne Ton präsentiert wurden, qualitativ schlechter empfunden als dieselben Videos, die mit Ton dargeboten wurden; die objektiven Unterschiede spiegelten sich demnach auf subjektiver Ebene wider.

Insgesamt kann also angenommen werden, dass die interne Validität der Untersuchung gegeben war; bzgl. der externen Validität bestehen auch für Studie III die üblichen Einschränkungen eines Laborexperiments, das an einer angefallenen Stichprobe von Psychologie-Studierenden durchgeführt wurde. Die Hypothese

Der observation-inflation-Effekt kann sowohl auf einer Fehlattribution sensorischer als auch auf einer Fehlattribution motorischer Repräsentationen beruhen; die auf Simulation beruhenden motorischen Repräsentationen scheinen jedoch zumeist den kritischeren Faktor darzustellen.

kann daher vor dem Hintergrund der getroffenen Einschränkungen als bewährt angesehen werden – allerdings bleibt zu diskutieren, inwiefern durch die gewählte Operationalisierung tatsächlich die in der Hypothese erwähnten motorischen Repräsentationen manipuliert wurden.

I.7.4.2 Diskussion unter theoretischen und praktischen Gesichtspunkten

Die dritte Studie zeigte, dass unter Bedingungen maximierter sensorischer Überlappung von Ausführung und Beobachtung zwar ein observation-inflation-Effekt auftrat; dieser war jedoch kleiner als in einer Kontrollbedingung mit geringerer sensorischer Überlappung, aber (mutmaßlich) stärkerer motorischer Überlappung. Zusätzlich fand sich ein Effekt auch nach minimierter, gar objektiv eliminiertes sensorischer, aber (vermeintlich) maximierter motorischer Überlappung von Ausführung und Beobachtung. Der Effekt unterschied sich nicht von der Kontrollbedingung, in der – unter ansonsten gleichen Bedingungen – eine stärkere sensorische Überlappung von Ausführung und Beobachtung existierte. Dieses Ergebnismuster stimmt perfekt mit dem vorhergesagten Ergebnismuster überein und ist von daher mit der Hypothese vereinbar, dass der observation-inflation-Effekt sowohl auf einer Fehlattribution von sensorischen als auch motorischen Gedächtnisspuren beruhen kann, wobei die motorischen die kritischere Rolle zu spielen scheinen. Bevor diese Interpretation jedoch abschließend bewertet

werden soll, sollen alternative Interpretationsmöglichkeiten und Aspekte der Konstruktvalidität diskutiert werden.

Zunächst sollen die beiden Gruppen, die sich bei gleicher Handhabung der Ausführung in Phase 1 (Durchführung mit offenen Augen) allein im sozialen Charakter der Beobachtungssituation bzw. in der Perspektive der Beobachtung (1.- vs. 2.-Person-Perspektive) unterschieden, miteinander verglichen werden. Wäre eine Fehlattribution sensorischer Erinnerungsspuren ausschlaggebend für die Gedächtnistäuschung, sollte eine größere perzeptuelle Ähnlichkeit der Quellen mehr Fehler bei der Quellenzuordnung verursachen (siehe Abschnitte I.1.3 und I.6.5.3). Demnach sollte die Beobachtung aus der 1.- verglichen mit der 2.-Person-Perspektive zu einem stärkeren observation-inflation-Effekt führen. Allerdings ist zu beachten, dass sich die Perspektivität im vorliegenden Falle nur in den visuellen Eindrücken entfaltet, während die Überlappung akustischer Informationen unabhängig von der Perspektive gleich bleibt. Auf visueller Ebene lassen sich jedoch zwei Dimensionen unterscheiden – Objekt- und Personenperspektive. Während für einige Objekte die Perspektive keine Rolle spielt, da sie keine klar erkennbare Vorder- und Rückseite haben (wie z.B. ein weißer Plastikbecher oder ein Gummiband), unterscheiden sich 1.- und 2.-Person-Perspektive bei anderen Objekten deutlich (z.B. bei einem Buch oder einem Notizblock). Die Handlung selbst bzw. das Agieren mit dem Objekt hingegen wird einerseits so beobachtet, wie es aussähe, wenn die Handlung selbst ausgeführt würde, andererseits so, wie wenn es eine andere, gegenüberstehende Person tun würde. Es lässt sich schwer ausmachen, welcher Aspekt der Perspektivität hier besonders bedeutsam ist; in der Literatur zum imagination-inflation-Effekt wird jedoch die Bedeutung objektbezogener Informationen betont. Wären diese Objekteindrücke, die sich ja in beiden Perspektiven nicht durchgängig unterscheiden, von besonderer Relevanz, wäre es vielleicht übertrieben, zu erwarten, dass die 1.- im Vergleich zur 2.-Person-Perspektive zu einem deutlichen Anstieg an *ausgeführt*-Antworten führen sollte. Mindestens allerdings sollten beide Perspektiven zu einem vergleichbaren Ausmaß an falschen Erinnerungen führen. Dies ist allerdings eindeutig nicht der Fall. Zwar fand sich ein observation-inflation-Effekt auch nach Beobachtung aus der 1.-Person-Perspektive, dieser war jedoch schwächer ausgeprägt als derjenige, der nach Beobachtung aus der 2.-Person-Perspektive auftrat! Demnach führte eine maximierte sensorische Überlappung von Ausführung und Beobachtung zu einem geringeren Ausmaß an falschen Erinnerungen. Vor dem geschilderten theoretisch-empirischen Hintergrund wird dies darauf zurückgeführt, dass zusammen mit der Perspektive vermutlich auch die Wahrscheinlichkeit von bzw. die Voraussetzungen für motorische Simulation variiert wurden. Simulation stellt ein Phänomen dar, das nachweislich in sozialen Situationen auftritt, also bspw. bei der Beobachtung eines Gegenübers. Demgegenüber sollte die Beobachtung aus der 1.-Person-Perspektive nicht (oder in geringerem Maße) die Voraussetzungen für Simulation erfüllen, da durch die

Selbst-Beobachtungs-Perspektive der soziale Charakter der Situation verringert und eine Objektfokussierung angestoßen wird. Diese Annahme muss allerdings spekulativ bleiben, da im Rahmen der vorliegenden Studie keine neurophysiologischen Daten erhoben wurden und keine anderen Studien existieren, die genau diese Fragestellung anhand vergleichbaren Materials untersucht hätten. Sie wird jedoch durch das Ergebnismuster plausibilisiert – die Maximierung der Überlappung von Handlungsausführung und Beobachtung auf sensorischer Ebene führte nicht zu einem erhöhten oder zumindest gleich bleibenden Ausmaß falscher Erinnerungen, sondern zu einem geringeren Fehler. Dieses Ergebnismuster kann nur durch das Wirken eines weiteren Mechanismus erklärt werden. Allerdings sind – neben motorischer Simulation – auch andere Mechanismen denkbar. So könnte man z.B. mutmaßen, dass die TeilnehmerInnen, die die Handlungen aus der 1.-Person-Perspektive beobachteten, aufgrund der ungewöhnlichen Perspektive eher antizipierten, dass die Untersuchung auf eine Diskrimination selbst- und fremdausgeführter Handlungen abzielte. Daher mögen sie im Gegensatz zu den Personen, die die Handlungen aus der 2.-Person-Perspektive beobachteten, die Videos voreingenommener und wachsender enkodiert haben. So zeigen auch die a posteriori erhobenen Ratings, dass die TeilnehmerInnen, die die Bedingung „hohe sensorische Überlappung, geringer sozialer Charakter“ durchliefen, einen Gedächtnistest (Item 2) bzw. den speziellen Gedächtnistest (Item 3) stärker erwartet hatten, als diejenigen der Bedingung „mittlere sensorische Überlappung, hoher sozialer Charakter“. Allerdings ergaben sich keine signifikanten, negativen Korrelationen zwischen der Einschätzung der Items und dem Ausmaß des observation-inflation-Effekts innerhalb dieser Gruppe, wie bei einer diesbzgl. Sensibilisierung erwartbar gewesen wäre (Item 2: $r(18) = .106$, $p > .05$; Item 3: $r(18) = -.164$, $p > .05$). In ähnlicher Weise könnte auch angenommen werden, dass die ungewöhnliche Perspektive als befremdlich erlebt und von daher aufmerksamer enkodiert worden sein könnte als die Vergleichsperspektive. Dem widerspricht jedoch das Muster der korrekten *ausgeführt*-Antworten, das exakt demjenigen in der Vergleichsbedingung entspricht. Demnach gibt es keine Anhaltspunkte, die diese Interpretationsmöglichkeiten unterstreichen würden. Im Gegenteil könnte gerade ein aus der ungewöhnlichen Perspektive resultierendes Fremdheitsgefühl dazu beitragen, dass die beobachteten Handlungen möglicherweise qualitativ anders – nämlich mit weniger motorischer Resonanz – enkodiert werden (siehe *uncanny valley*, Mori, 1982; siehe aber auch z.B. Schütz-Bosbach et al., 2006). Auch andere zur Simulationsannahme alternative Erklärungsmöglichkeiten werden durch die gefundenen Ergebnisse nicht unterstützt. So wurde z.B. das Beobachten aus der 1.-Person-Perspektive nicht als anstrengender empfunden als das Beobachten aus der 2.-Person-Perspektive. Es kann also davon ausgegangen werden, dass die Videos in beiden Bedingungen ähnlich konzentriert enkodiert wurden. Auch ergab sich kein Zusammenhang zwischen dem Geschlecht der BeobachterInnen und dem Ausmaß des observation-inflation-Effekts ($r_{pbis} = .083$, $n = 18$, $p > .05$). Daher kann ebenfalls ausgeschlossen werden, dass der reduzierte

Effekt allein darauf zurückzuführen sei, dass die weiblichen Teilnehmerinnen, die den Großteil der Gruppe darstellten, eine Distanz zu dem aus der Selbst-Beobachtungs-Perspektive gezeigten gegengeschlechtlichen Akteur aufbauten. Auch die Annahme eines strengeren Antwortkriteriums in der Gruppe, die aus der 1.-Person-Perspektive beobachtete, kann aufgrund des perfekt mit der Kontrollbedingung vergleichbaren Ergebnismusters korrekter *ausgeführt*-Antworten ausgeschlossen werden. Zudem zeigte schon Studie II, dass eine Warnung, die die TeilnehmerInnen aufgrund des Abfrageformats nach Beobachtung aus der 1.-Person-Perspektive möglicherweise selbst generiert haben könnten, nicht zu einer substantiellen Reduktion des Effekts führt. Demnach können auch keine Strategien bzw. Antwortkriterien beim Abruf für die Reduktion des Effekts verantwortlich gemacht werden. Zusammengefasst lassen sich keine Erklärungsansätze finden, die das gefundene Ergebnismuster ebenso plausibel erklären könnten, wie die Simulationsannahme. Diese wird auch durch die Art der fehlattribuierten Handlungsanweisungen unterstützt. So waren die nach Beobachtung aus der 2.-Person-Perspektive mit am häufigsten fehlattribuierten Handlungsanweisungen solche, deren Objektperspektive bei eigener und fremder Ausführung deutlich divergiert, die jedoch gleichzeitig sehr geräuscharm sind („Blättern Sie in dem Buch!“ „Stecken Sie die Karte in den Umschlag!“; siehe auch Abschnitt I.6.5). Akzeptiert man auf dieser Basis die Annahme, dass das reduzierte Ausmaß falscher Erinnerungen in dieser Bedingung auf fehlenden oder nur sehr schwach ausgeprägten motorischen Repräsentationen beruhen könnte, muss jedoch zugestanden werden, dass der observation-inflation-Effekt auch durch eine Fehlattribution vorrangig sensorischer Erinnerungsspuren hervorgerufen werden kann – denn schließlich konnte selbiger in einem substantiellen Ausmaß nachgewiesen werden. Der Effekt war in dieser Gruppe sogar von dem höchsten Ausmaß subjektiver Sicherheit begleitet, was wiederum plausibel erscheint, wenn die falschen Erinnerungen in dieser Gruppe tatsächlich primär auf erinnerte qualitative Merkmale zurückzuführen wären.

Wenn TeilnehmerInnen die Handlungen in der ersten Phase mit geschlossenen Augen ausführten, so dass sie nur taktile und akustische, jedoch keine visuellen Eindrücke des Objektes oder der Handlung enkodierten, und die Videos aus der 2.-Person-Perspektive ohne Ton beobachteten, also nur die visuellen Eindrücke aufnahmen, dann zeigten diese TeilnehmerInnen dennoch einen observation-inflation-Effekt. Da nun de facto keine Überlappung von Ausführung und Beobachtung auf sensorischer Ebene existierte, müssen Fehlattributionen andersartiger Qualitäten dieses Ergebnismuster hervorgerufen haben. Bevor diese Möglichkeit diskutiert werden soll, lässt sich jedoch zunächst in Frage stellen, ob Ausführung und Beobachtung auf sensorischer Ebene tatsächlich so distinkt waren, wie angenommen, d.h. ob eine Überlappung sensorischer Merkmale tatsächlich eliminiert war. Bspw. ist denkbar, dass die TeilnehmerInnen selbst visuelle Eindrücke generierten, indem sie sich die Objekte oder die Hand-

lungsausführung vorstellten. Allerdings könnte dies auch einen gegenteiligen Effekt – nämlich eine leichtere Diskriminierbarkeit – nach sich ziehen, da vorgestellte Objekte nicht mit den nachher im Video gezeigten übereinstimmen müssen. So könnte eine Person sich ein rotes Handtuch vorstellen, müsste aber im Video erkennen, dass das Handtuch tatsächlich blau war. Dies sollte dann eher zu einem gegenteiligen Effekt – also sogar zu einer besseren Diskriminierbarkeit beider Quellen – führen. Allerdings waren unter den Gegenständen auch solche, bei denen keine bzw. kaum Abweichungen zwischen vorgestelltem und tatsächlichem Aussehen aufgetreten sein können, da sie über ein sehr standardisiertes Äußeres verfügen (z.B. Tennisball). Auch ist denkbar, dass die TeilnehmerInnen bei der Beobachtung selbst sensorische Qualitäten hinzufügten, sich z.B. taktile oder akustische Eindrücke vorstellten. In der Tat zeigen die a posteriori erhobenen Ratings, dass beides der Fall war – sowohl während der Ausführung als auch während der Beobachtung generierten die TeilnehmerInnen selbst sensorische Eindrücke. Wichtig ist aber, dass die Korrelationen zwischen dem Ausmaß, in dem zusätzliche sensorische Eindrücke generiert wurden, und dem Ausmaß des observation-inflation-Effekts nicht konsistent und nicht substantiell bzw. signifikant waren (Phase 1: Objekt visuell vorgestellt: $r(18)=.324$, $p>.05$; Phase 1: Ausführung visuell vorgestellt: $r(18)=.371$, $p>.05$; Phase 2: Anfühlen vorgestellt: $r(18)=-.195$, $p>.05$; Phase 2: Geräusche vorgestellt: $r(18)=-.110$, $p>.05$). Demnach lässt sich nicht unterstreichen, dass die falschen Erinnerungen vorrangig auf einer Fehlattribution selbst-generierter sensorischer Eindrücke basierten. Einzig die Tatsache, dass korrekte *ausgeführt*-Antworten nach fünf- im Vergleich zu nullmaliger Präsentation auch in dieser Gruppe häufiger auftraten, macht etwas skeptisch. Es wäre erwartbar gewesen, dass sich hier die Frequenz der Präsentation weniger stark auswirken würde, da das wiederholte Beobachten auf sensorischer Ebene ja v.a. mit lebendigen visuellen Eindrücken einhergeht, die aber gerade in der betrachteten Gruppe nicht notwendig als diagnostisch für die eigene Handlungsausführung angesehen werden sollten. Demnach hätte das Ergebnismuster dieser Gruppe keinen Frequenzeinfluss hinsichtlich korrekter *ausgeführt*-Urteile aufweisen sollen, da die eigene Handlungsausführung eher aufgrund von taktilen oder akustischen Eindrücken beurteilt werden sollte, die ja durch die Videopräsentation nicht beeinflusst wurden (siehe auch Warnungsbedingung, Experiment II, Abschnitt I.6.5). Dem ist allerdings entgegen zu setzen, dass der Frequenzeffekt nicht zwangsweise aufgrund von lebendigen visuellen Eindrücken zustande gekommen sein muss. Denn durch die wiederholte Beobachtung werden vermutlich auch die motorischen Repräsentationen gestärkt. Demnach könnte das Muster auch dafür sprechen, dass motorische cues als Basis der Erinnerungsurteile fungierten. Interessanterweise findet sich nämlich in dieser Gruppe – wie auch für die korrekten *ausgeführt*-Urteile erwartet – bzgl. der rekognizierten Items kein Gefälle zwischen fünf- und nullmaliger Präsentation, obwohl dies in allen anderen Bedingungen der Fall ist. Auch sind die Rekognitionsquoten fünf- bzw. nullmalig präsentierter Handlungsanweisungen, die in Phase 1 *nicht*

ausgeführt wurden, weniger unterschiedlich als in den anderen Bedingungen. Dies mag damit zusammenhängen, dass die Objekte, die ja jeweils vor Erscheinen der Handlungsanweisungen befühlt wurden, bei der Ausführung mit geschlossenen Augen tiefer enkodiert wurden, da die taktile im Vergleich zur visuellen Enkodierung ungewöhnlicher ist und somit eine tiefere Gedächtnisspur hinterlässt, die durch die nachfolgende Videodarbietung nicht so stark beeinflussbar ist. Insgesamt scheint also auch die Möglichkeit, dass die TeilnehmerInnen beim Abruf auf aus den Videos stammende visuelle cues zurückgriffen, obwohl diese nicht notwendig diagnostisch für die Handlungsausführung waren, nicht überzeugend. Das Ergebnismuster ist hingegen gut mit der Idee vereinbar, dass motorische Repräsentationen, die durch Simulation hervorgerufen wurden, eine kritische Rolle spielen. Diese Interpretation wird insbesondere durch den Vergleich mit der Kontrollbedingung unterstrichen (offene Augen, 2.-Person-Perspektive mit Ton). Beide Gruppen unterscheiden sich nur im Ausmaß sensorischer Überlappung von Ausführung und Beobachtung, die in dem einen Fall partiell, in dem anderen nicht gegeben ist. Aufgrund der gleichen Perspektive kann jedoch davon ausgegangen werden, dass in beiden Fällen motorische Simulation gleichermaßen angestoßen wurde. Da nun die (objektive) Elimination sensorischer Überlappung keine signifikante Reduktion des Effekts zur Folge hatte, scheinen die motorischen Repräsentationen die kritische Rolle zu spielen. Die Art der fehlattribuierten Handlungsanweisungen trägt diesbzgl. allerdings nicht viel zur Klärung bei, da v.a. (objekt-)perspektivisch unspezifische Handlungen fehlattribuiert wurden. Es ist jedoch anzumerken, dass viele der fälschlich als ausgeführt erinnerten Handlungen Objekte involvierten, deren Aussehen nicht standardisiert ist, so dass das vorgestellte mit großer Sicherheit von dem tatsächlichen Aussehen abweichen würde (z.B. bzgl. der drei am häufigsten fehlattribuierten Handlungsanweisungen: eine grüne Tasse mit Rosen, eine orangene Schachtel, rote Würfel). Dabei ist erstaunlich, dass die Erinnerungsurteile beider Gruppen mit einer relativ hohen Sicherheit einhergingen, wobei partielle sensorische Überlappung etwas höhere Werte hervorbrachte. Dies ist möglicherweise darauf zurückzuführen, dass in jener Gruppe auch qualitative Merkmale als Basis der Erinnerungsurteile dienten (siehe auch Gruppe „hohe sensorische Überlappung, geringer sozialer Charakter“). Zusätzlich konvergiert das Ergebnis mit den Befunden aus der zweiten Studie – aufgrund des automatischen und unbewussten Charakters der beobachtungsinduzierten Simulation scheint es (auch bei Minimierung der sensorischen Überlappung von Ausführung und Beobachtung) ausgeschlossen, dass falsche Erinnerungen aufgrund derartiger motorischer Erinnerungsspuren vermeidbar sein könnten. Kritisch ist allerdings anzumerken, dass der observation-inflation-Effekt in der Bedingung „geringe sensorische Überlappung, hoher sozialer Charakter“ deskriptivstatistisch reduziert war. Möglicherweise wäre dieser Unterschied bei größerem Stichprobenumfang statistisch abzuschließen gewesen. Nichtsdestotrotz hätte – auch bei größerer Stichprobe – in dieser Bedingung

überhaupt ein substantieller, nicht auf sensorische Überlappung zurückführbarer Effekt bestanden.

Zusammenfassend lässt sich also festhalten, dass im Rahmen der vorliegenden Studie keine direkte Evidenz für die Bedeutsamkeit motorischer cues bei falschen Erinnerungen an die Handlungsausführung vorliegt, da rein behaviorale Daten erfasst wurden. Daher soll die Hypothese „Der observation-inflation-Effekt kann sowohl auf einer Fehlattribution sensorischer als auch auf einer Fehlattribution motorischer Repräsentationen beruhen; die auf Simulation beruhenden motorischen Repräsentationen scheinen jedoch zumeist den kritischeren Faktor darzustellen.“ stets im Zusammenhang mit der gewählten Operationalisierung gesehen werden.

Für den Alltag bedeutet dies – bei aller Vorsicht bei der Übertragung des Laborparadigmas in die Feldsituation –, dass falsche Erinnerungen an alltägliche, einfache Handlungen sowohl durch sensorische Eindrücke, die bei der Beobachtung einer anderen Person oder auch bei der eigenen Ausführung zu einem früheren Zeitpunkt enkodiert wurden, als v.a. auch durch motorische Repräsentationen, die wiederum auf die Beobachtung einer fremden oder die eigene Ausführung zu einem früheren Zeitpunkt zurückgehen können, hervorgerufen werden könnten. Damit erscheint die Möglichkeit, derartige Gedächtnistäuschungen durch metakognitive Strategien a posteriori vermeiden zu können, fast ausgeschlossen (vgl. auch Abschnitt I.5.4).

I.7.4.3 Résumé und Ausblick

Anhand von Studie III konnte gezeigt werden, dass der observation-inflation-Effekt sowohl auf einer Fehlattribution von sensorischen als (v.a.) auch motorischen Merkmalen basieren könnte. Wurde die motorische Überlappung von Ausführung und Beobachtung (jedenfalls vermeintlich) verringert, reduzierte sich – trotz maximierter sensorischer Überlappung – auch der Effekt. Hingegen hatte die Elimination sensorischer Überlappung keine Reduktion des Effekts zur Folge, wenn die motorische Überlappung nicht variiert wurde. Die Ergebnisse sind demnach mit der Annahme vereinbar, dass die durch Simulation erzeugten motorischen Repräsentationen eine kritischere Rolle spielen als die sensorischen Repräsentationen.

Dabei wurde die Manipulation motorischer Repräsentationen anhand einer Operationalisierung realisiert, die empirisch nicht abgesichert war und anhand der behavioralen Daten auch nicht abgesichert werden konnte. Um die Simulationsannahme direkt empirisch untermauern zu können, wären bildgebende oder elektrophysiologische Studien vonnöten, die eine Erfassung der neuronalen Aktivität motorischer Areale erlauben. Den Einfluss sensorischer und motorischer Merkmale auf Verhaltensebene weiter einzugrenzen, ist hingegen eine sehr schwierige

Aufgabe. In derartigen verhaltensbasierten Studien könnten möglicherweise Videos eingesetzt werden, auf denen die Handlungen nur imitiert und nicht mit den Objekten ausgeführt werden. Ein ähnlicher Effekt könnte durch den Einsatz der Lichtpunkt-Technik (Johansson, 1973) erzielt werden. Mit dieser Standardmethode ließe sich allein die Motorik aufzeichnen, während keinerlei sensorische Eindrücke (auch keine konkreten Personeneindrücke) hervorgerufen würden. Aufgrund der minimalen, jedoch sicherlich sehr bedeutsamen Bewegungsunterschiede bei der Ausführung verschiedener einfacher Handlungen (z.B. „Schrauben Sie den Kugelschreiber auf!“ vs. „Schrauben Sie den Deckel vom Deo!“), sind die vorgeschlagenen Möglichkeiten verhaltensbasierter Manipulationen jedoch unter Umständen inadäquat und bedürfen einer sehr kritischen Abwägung (z.B. Senkfor, 2008). Die Trennung beider Mechanismen sollte jedenfalls unbedingt Gegenstand weiterer Untersuchungen sein, um ihren Einfluss genauer ein- und abgrenzen zu können.

Wie genau die Umsetzung motorischer Repräsentationen in Erinnerungsurteile aussehen könnte, bleibt unklar – dennoch ist anzunehmen, dass die vermuteten Gedächtnis- und Urteilsprozesse eher unbewusst ablaufen (vgl. z.B. Schütz-Bosbach et al., 2006).

I.8 Zwischenrésumé und Ausblick zu den Experimenten I bis III

Die Studien I bis III haben gezeigt, dass analog zum *imagination-inflation-* (Garry, Manning, Loftus & Sherman, 1996; Goff & Roediger, 1998) ein *observation-inflation-Effekt* existiert: Nicht nur nach wiederholter Vorstellung, sondern auch nach wiederholter Beobachtung der Handlungsausführung erinnerten TeilnehmerInnen, einige Handlungen ausgeführt zu haben, die sie de facto nicht ausführten. Bei diesem Effekt handelt es sich weder um ein methodologisches Artefakt, noch um das Resultat einer Antworttendenz oder eines bestimmten Antwortkriteriums, sondern um eine echte Gedächtnistäuschung, die a posteriori nahezu unvermeidlich aufzutreten scheint. Als Mechanismen scheiden eine Fehlattribution von Vertrautheit mit den Handlungsanweisungen sowie eine Fehlattribution des Informationsursprungs aus. Vielmehr scheint die Gedächtnistäuschung sowohl durch eine Fehlattribution von sensorischen Eindrücken als auch motorischen Repräsentationen verursacht zu werden, wobei letztere den kritischeren Faktor darstellen könnten. Allerdings ist einschränkend anzumerken, dass keine direkte Evidenz für eine gelungene Manipulation motorischer Repräsentationen erbracht werden konnte.

Die genannten Befunde erweitern bisherige Erkenntnisse zu falschen Erinnerungen an die Ausführung einfacher Handlungen ganz erheblich – und zwar auf verschiedenen Ebenen. Zunächst wurden derartige falsche Erinnerungen bislang ausschließlich anhand von Vorstellungen untersucht, die gegenüber Beobachtungen jedoch zwei Nachteile aufweisen: Mangelnde Überprüfbarkeit der compliance und interindividuelle Variabilität der Vorstellungsbilder. Durch den Einsatz von Beobachtung können demnach zwei methodologische Schwachstellen bisheriger Untersuchungen umgangen werden. Darüber hinaus konnten hinsichtlich der Eingrenzung des Mechanismus, der dem inflation-Effekt unterliegt, wesentliche Fortschritte gegenüber existierenden Studien erzielt werden. Einerseits wurde in vorausgegangenen Untersuchungen beim Vergleich verschiedener Verarbeitungsprozesse nicht die Dimension des Informationsursprungs bzw. der Verarbeitungstiefe (Craik & Lockhart, 1972) beachtet. Daher konnte bis dato nicht vollkommen sichergestellt werden, dass sowohl Vertrautheit als auch Informationsherkunft *keinen* Einfluss auf die Gedächtnistäuschung ausüben – genau dies konnte jedoch im Rahmen der vorliegenden Untersuchungsserie ausgeschlossen werden. Zukünftige Untersuchungen könnten jedoch zusätzlich zu den hier umgesetzten weitere Operationalisierungen von Verarbeitungstiefe bzw. internalem Informationsursprung realisieren und so den Geltungsbereich der Erkenntnisse ausweiten. Darüber hinaus konnte zumindest für den Prozess des Beobachtens gezeigt werden, dass es sich beim inflation-Effekt um eine genuine Gedächtnistäuschung und nicht um eine bloße Antworttendenz oder ein methodologisches Artefakt handelt. Aufgrund der ausgeprägten Ähnlichkeiten beider Prozesse könnte sich dies auf den

imagination-inflation-Effekt übertragen lassen; diese Annahme müsste jedoch ebenfalls im Rahmen zukünftiger Untersuchungen empirisch überprüft werden. Die entscheidende Erweiterung besteht allerdings darin, dass die gefundene Gedächtnistäuschung nicht nur – wie bisher – ausschließlich aus kognitionspsychologischer Perspektive betrachtet wurde, sondern auch neurowissenschaftliche Befunde integriert wurden. Hierdurch war es möglich, auf einen weiteren Mechanismus aufmerksam zu werden, der möglicherweise sogar die kritischere Rolle bei derartigen falschen Erinnerungen spielen könnte: eine Fehlattribution motorischer Repräsentationen, die auf die Simulation der Handlungsausführung zurückgehen. Um die Bedeutung und Wirkungsweise dieses Mechanismus näher eingrenzen zu können, sind jedoch weitere Studien vonnöten. Insbesondere bieten sich hier bildgebende bzw. elektrophysiologische Verfahren an, die direkte Evidenz für die Aktivität motorischer Hirnareale liefern könnten. Inwiefern diese Erinnerungsqualität auch für den imagination-inflation-Effekt bedeutsam ist, sollte ebenfalls in zukünftigen Studien geklärt werden, denn möglicherweise beruhen sowohl falsche Erinnerungen aufgrund von Beobachtung als auch aufgrund von Vorstellung auf sehr ähnlichen Mechanismen, da beide Prozesse als Formen der Handlungssimulation begriffen werden können (vgl. für eine Diskussion der Unterschiede Abschnitt I.4).

Es ist offensichtlich, dass die gewonnenen Erkenntnisse nicht einfach auf Alltagssituationen übertragen werden können, denn außerhalb der Laborsituation wirken eine Menge weiterer Variablen ein, die einen entscheidenden Einfluss auf die Erinnerungsleistung ausüben, wie die Aufmerksamkeitszuwendung oder das Intervall zwischen Enkodierung und Abruf. Allerdings ist zu konstatieren, dass es sich bei dem verwendeten Material um ganz alltägliche Aktivitäten handelt, so dass die gewonnenen Erkenntnisse unter ansonsten vergleichbaren Bedingungen auf den Alltag übertragbar sein dürften. Neben den genannten Faktoren variiert im Alltag aber auch der Erinnerungsinhalt – die Handlungen können z.B. viel komplexer sein als die in der Untersuchung realisierten. Auf derartige Handlungen bzw. Ereignisse lassen sich die Erkenntnisse jedoch sicherlich nicht übertragen, da bei der Erinnerung an solche Handlungen andersartige Erinnerungsqualitäten eine Rolle spielen bzw. die untersuchten Erinnerungsqualitäten anderweitig gewichtet werden dürften – darauf weisen z.B. die Erkenntnisse zum imagination-inflation-Effekt für Kindheitsereignisse hin.

Auch wenn die Übertragbarkeit bzgl. komplexerer Handlungen nicht gegeben ist – schon allein die (potentielle) Übertragbarkeit der Erkenntnisse bzgl. einfacher Handlungen auf die Alltagssituation besitzt eine besondere Relevanz. Denn gerade auch falsche Erinnerungen an derartige einfache Handlungen – z.B. daran, den Herd abgestellt, eine Tablette genommen oder die Tür abgeschlossen zu haben – können gravierende Konsequenzen nach sich ziehen, im schlimmsten Falle z.B. einen Wohnungsbrand, eine Untermedikation oder ein Eindringen Fremder in die

eigene Wohnung. Das Ausmaß solcher Erinnerungsfehler dürfte dabei nicht nur von den bereits erörterten situationalen Bedingungen und dem Erinnerungsinhalt, sondern (neben weiteren Faktoren) auch von dem Alter der erinnernden Person abhängen; die erläuterten Erkenntnisse gelten daher nur für die untersuchte Altersgruppe junger Erwachsener. So ist bspw. davon auszugehen, dass derartige Lapsi älteren Personen häufiger unterlaufen – es ist bekannt, dass ältere Menschen gerade hinsichtlich der untersuchten episodischen Gedächtnisprozesse Einschränkungen aufweisen. In der Studie von Thomas und Bulevich (2006) fand sich dann auch für die älteren im Vergleich zu den jüngeren TeilnehmerInnen ein stärker ausgeprägter imagination-inflation-Effekt. Um ein vollständigeres Bild des gefundenen observation-inflation-Effekts entwickeln zu können, sollen die bisherigen Befunde daher durch Untersuchungen an älteren TeilnehmerInnen ergänzt werden. Hierdurch sollen nicht nur mögliche Mechanismen weiter eingegrenzt bzw. die Populationsvalidität ausgeweitet werden, sondern es soll auch gefragt werden, welche Strategien besonders älteren TeilnehmerInnen an die Hand gegeben werden könnten, um derartige Gedächtnistäuschungen zu vermeiden bzw. minimieren.

II ERWEITERUNG DER EMPIRISCHEN BASIS: OBSERVATION INFLATION BEI ÄLTEREN ERWACHSENEN

II.1 Veränderungen von Gedächtnisleistungen im Alter

II.1.1 Übersicht über Veränderungen und deren vermutete Ursachen

Kognitionspsychologische Untersuchungen verschiedener Altersgruppen zeigen einhellig, dass nicht alle Aspekte des Gedächtnisses im Alter gleichermaßen beeinträchtigt sind (z.B. Balota, Dolan & Duchek, 2000; Anderson & Craik, 2000). Während sich z.B. auf Ebene des Langzeitgedächtnisses das prozedurale bzw. implizite Gedächtnis jüngerer und älterer TeilnehmerInnen fast gar nicht unterscheiden, treten deutliche Unterschiede in den episodischen Gedächtnisleistungen zu Tage. Diese sind allerdings auch von der Abrufsituation abhängig – so zeigen sich bei Aufgaben, die eine freie Reproduktion des Enkodierten erfordern, große Differenzen, die jedoch mit zunehmender Vorgabe von Abrufhinweisen verschwinden, so dass sich bspw. bei der Rekognition kaum noch Alterseffekte finden (z.B. Mayr, 2003). Auch auf Ebene des semantischen Gedächtnisses ist zu differenzieren zwischen Aufgaben, die eine automatische Ausbreitung der Aktivierung beinhalten und keine Altersunterschiede hervorbringen, und solchen, die mehr Aufmerksamkeitsressourcen beanspruchen bzw. den kontrollierten Abruf von Wissen erfordern und eine Beeinträchtigung der älteren TeilnehmerInnen zu Tage fördern (*tip-of-the-tongue-Erlebnis*; z.B. Balota, Dolan & Duchek, 2000). Bzgl. des sensorischen Registers (Sperling, 1960) und der sog. Gedächtnisspanne zeigen sich ebenfalls keine bzw. kaum Unterschiede zwischen den Altersstufen; Leistungen des Arbeitsgedächtnisses (genauer der sog. zentralen Exekutive; Baddeley, 1986) sind jedoch bei älteren im Vergleich zu jüngeren TeilnehmerInnen deutlich reduziert. Im Übrigen werden unter dem Begriff „jüngere TeilnehmerInnen“ meist Personen im Alter von ca. 20 Jahren verstanden, unter dem Begriff „ältere TeilnehmerInnen“ Personen im Alter ab ca. 60 Jahren, also solche, die bereits berentet sind (z.B. Balota, Dolan & Duchek, 2000).

Die skizzierten Altersunterschiede in vorwiegend episodischen, aber teilweise auch semantischen Erinnerungsleistungen sowie im Arbeitsgedächtnis (der zentralen Exekutive) wurden auf verschiedene Mechanismen zurückgeführt, von denen im Folgenden nur die beiden zentralsten erwähnt werden sollen (z.B. Anderson & Craik, 2000; für eine Übersicht siehe z.B. Balota, Dolan & Duchek, 2000). Der erste Ansatz führt die Unterschiede in der Erinnerungsleistung auf Unterschiede in der Verarbeitungsgeschwindigkeit zurück (Salthouse, 1996a). So fand bspw. Salthouse (1996b), dass die Verarbeitungsgeschwindigkeit als Mediator zwischen Alter

und Gedächtnisleistung fungierte; wurde der Faktor herauspartialisiert, brach der Zusammenhang zwischen Alter und Gedächtnisleistung zusammen. Der zweite Ansatz sieht die begrenzten Aufmerksamkeitsressourcen als ursächlich für die schlechteren Gedächtnisleistungen an. Aufgrund reduzierter Aufmerksamkeitsressourcen sind ältere TeilnehmerInnen nicht in der Lage, kognitiv aufwändigere Enkodierstrategien anzuwenden, die eine elaborierte Verarbeitung gewährleisten würden (Craik & Byrd, 1982). Stattdessen würden Informationen unspezifischer und automatischer enkodiert, so dass beim Abruf weniger Erinnerungsspuren generiert werden könnten. Dieser Erklärungsansatz wird durch Untersuchungen plausibilisiert, die die Daten älterer TeilnehmerInnen mit denjenigen jüngerer TeilnehmerInnen vergleichen, die während des Enkodierens eine Zweitaufgabe bearbeiten sollen (z.B. Naveh-Benjamin, Craik, Gavrilescu & Anderson, 2000; Naveh-Benjamin, 2006). In derartigen Untersuchungen erzielen ältere TeilnehmerInnen, die unter Bedingungen voller Aufmerksamkeit enkodierten, i.d.R. Ergebnisse, die mit denjenigen von jüngeren TeilnehmerInnen, die unter Bedingungen geteilter Aufmerksamkeit enkodierten, vergleichbar sind. Dies gilt interessanterweise nicht nur für die Erinnerungsleistung, sondern auch für die Reaktionszeiten (z.B. Anderson, Craik & Naveh-Benjamin, 1998). Craik (2006) zufolge könnte daher auch die häufig gefundene Verlangsamung älterer TeilnehmerInnen (siehe Ansatz 1) letztlich das Ergebnis eines übergeordneten Defizits an Verarbeitungsressourcen darstellen. Dabei wird der Verlust derartiger Ressourcen auf mit dem Alter einhergehende neurobiologische Veränderungen zurückgeführt (Craik, 2006); auf einige dieser Veränderungen wird im folgenden Abschnitt näher eingegangen. Beide Erklärungsansätze wurden kritisiert, da sie aufgrund ihres unspezifischen Charakters das spezifische Muster beeinträchtigter Erinnerungsleistungen nicht vollständig zu erklären vermögen; ein Argument, das jedoch durch Weiterentwicklungen entschärft werden konnte (Balota, Dolan & Duchek, 2000).

Die gefundenen Altersunterschiede in episodischen Erinnerungsleistungen können grundsätzlich auf drei Stagen des Erinnerungsprozesses zurückgehen: auf die Enkodierung, das Behalten oder den Abruf der Informationen. Aus den bisher referierten Befunden geht bereits hervor, dass der ersten und letzten Stage eine besondere Bedeutung zukommt.

Auf ein Problem älterer Menschen im Rahmen der Enkodierstage wurde bereits hingewiesen: Ihnen stehen mutmaßlich nicht genug Verarbeitungskapazitäten zur Verfügung, um Informationen ausreichend tief zu enkodieren bzw. verschiedene Informationen zu einem kohärenten und differenzierten Bild zusammenzufügen (*binding*; z.B. Craik, 2006). Demnach sind die beim Abruf verfügbaren Erinnerungsspuren bereits „blasser“ und weniger distinkt als diejenigen jüngerer TeilnehmerInnen; dies spiegelt sich auch in der Einschätzung von Erinnerungsqualitäten durch ältere Menschen wider (z.B. Mäntylä, 1993; Parkin & Walter, 1992; Spencer &

Raz, 1995). Daraus folgt bzgl. der letzten Stage, dass Erinnerungstests, die stark die Aktivität und Selbstinitiative des Erinnernden erfordern (z.B. freier Abruf), deutliche Beeinträchtigungen älterer Menschen aufzeigen müssen. Es lässt sich ebenfalls ableiten, dass mit zunehmender Verfügbarkeit von Abrufhinweisen Unterschiede zwischen den Altersgruppen verschwinden sollten. Die Gedächtnisbeeinträchtigungen älterer TeilnehmerInnen scheinen aber nicht nur mit der Qualität der Erinnerungsspuren zusammenzuhängen. Hinzu kommen ineffiziente Abrufstrategien, die insb. den selbstständigen Abruf, aber auch den Abruf unter Vorlage von Hinweisreizen, negativ beeinflussen. So scheinen ältere TeilnehmerInnen häufig nicht das ganze Spektrum und die volle Qualität ihrer Erinnerungsspuren mit einzubeziehen (z.B. Buckner & Logan, 2002), sondern auf Basis eines reduzierten Spektrums, insb. auf Basis der Vertrautheit von Informationen, zu urteilen (z.B. Jacoby & Hay, 1998). Insgesamt scheinen also bzgl. des Abrufs dreierlei Probleme zu existieren: Die verminderte Qualität spezifischer Erinnerungsspuren, das reduzierte Spektrum berücksichtigter spezifischer Erinnerungsspuren sowie die übermäßige Nutzung globaler Erinnerungscharakteristika, wie Vertrautheit.

II.1.2 Veränderungen des Quellengedächtnisses und deren neuronale Korrelate

Aus dem geschilderten Muster alterskorrelierter Gedächtnisbeeinträchtigungen lässt sich schließen, dass ältere TeilnehmerInnen gerade in Aufgaben, die eine nachträgliche Diskrimination von Quellen erfordern, besonders schlecht abschneiden müssten – schließlich handelt es sich um episodische Erinnerungen, deren Kontext im Detail zu reproduzieren ist. In der Tat weisen ältere TeilnehmerInnen besonders starke Beeinträchtigungen in Quellengedächtnistests auf und sind daher auch besonders anfällig für verschiedene Formen falscher Erinnerungen (z.B. Cohen & Faulkner, 1989; Ferguson, Hashtroudi & Johnson, 1992; McIntyre & Craik, 1987; Mitchell, Johnson & Mather, 2003).

Dies ist nicht erstaunlich, wenn man die neuronalen Korrelate des Quellengedächtnisses in Betracht zieht. Wenn Menschen gebeten werden, Quellenzuordnungen vorzunehmen, scheint sich dies auf neuronaler Ebene unter Beteiligung der Frontallappen abzuspielen. So wird bspw. berichtet, dass PatientInnen mit fokalen Schädigungen dieser Gehirnstrukturen zwar ein mit Gesunden vergleichbares Item- oder Inhaltsgedächtnis aufweisen, jedoch große Probleme haben, die Quellen dieser Inhalte korrekt zuzuschreiben (z.B. Janowsky, Shimamura & Squire, 1989; Johnson et al., 1997a,b). Bildgebende Verfahren (fMRT) und ERP (ereigniskorrelierte Potentiale) unterstützen diese Annahme bei Gesunden (z.B. Senkfor & Van Petten, 1998; Noldé, Johnson & D'Esposito, 1998). Der frontale Cortex ist aber nicht nur beim Abruf, sondern auch bei der Enkodierung aktiv und interagiert mit dem medialen Temporallappen: „During encoding PFC [prefrontal cortex] is assumed to implement the strategies that organize the

input to the MTL [medial temporal lobes], and during retrieval, it is assumed to control search and monitoring operations" (Cabeza, 2006, S. 597; siehe auch Buckner & Logan, 2002). Neuronale Aktivität dieser Gehirnstruktur scheint also mit Prozessen der Organisation, des Überprüfens und Manipulierens von Informationen zusammenzuhängen; daher wurden auch Arbeitsgedächtnisprozesse mit einer Aktivität der Frontallappen in Verbindung gebracht (z.B. Buckner & Logan, 2002). Genau diese Hirnstrukturen, die Frontallappen, atrophieren besonders stark im Alter (z.B. Buckner & Logan, 2002; Cabeza, 2006; Raz, 2000; West, 1996). Demgegenüber wird davon ausgegangen, dass das sog. Itemgedächtnis eher auf einer Aktivität des medialen Temporallappens basiert, der im Alter einer weniger starken Atrophie unterliegt (Shimamura & Squire, 1987; vgl. für eine alternative Sichtweise Cabeza, 2006). Diese Annahmen werden durch eine Studie von Glisky, Polster und Routhieaux (1995) unterstützt. Die AutorInnen stellten sich die Frage, ob beide Arten der Erinnerung – Inhalt und Quelle – doppelt dissoziiert sein könnten. Sie teilten eine Gruppe älterer TeilnehmerInnen nach ihrem Abschneiden in einer Testbatterie zur Erfassung sog. frontaler Hirnfunktionen und nach ihrem Abschneiden in einer Testbatterie zur Erfassung sog. medial-temporalen (bzw. diencephalen) Hirnfunktionen in jeweils zwei Gruppen (gute und schlechte Leistung) ein. Es zeigte sich, dass die Gruppe, die in der „frontalen Testbatterie“ gut abgeschnitten hatte, im Vergleich zu der schlechten Gruppe bei gleichen Inhaltsgedächtnisleistungen besser im Quellengedächtnis abschnitt. Umgekehrt erreichte die Gruppe, die in der „medial-temporalen Testbatterie“ gut abgeschnitten hatte, im Vergleich zu der schlechten Gruppe bei gleichen Quellengedächtnisleistungen bessere Ergebnisse im Inhaltsgedächtnis. Demnach scheinen also tatsächlich beide Formen der Erinnerung auf (zumindest partiell) unterschiedlichen Gehirnstrukturen zu basieren, die durch das Alter differentiell beeinflusst werden (siehe auch Cabeza, 2006). Im Übrigen haben Fletcher et al. (1995) gezeigt, dass auch das Teilen der Aufmerksamkeit, das – wie erwähnt – häufig genutzt wird, um die Leistungen jüngerer mit denjenigen älterer TeilnehmerInnen vergleichbar zu machen, zu einer Reduzierung der neuronalen Aktivität im frontalen Cortex führt (Fletcher, Frith, Grasby, Shallice, Frackowiak & Dolan, 1995). Insgesamt lässt sich mit Buckner und Logan (2002) resümieren: „First, older adults fail to appropriately recruit frontal regions associated with encoding. Second, to the degree that frontal regions are recruited, older adults do so in a less selective manner than younger adults. [...] Moreover, these principles are not limited to studies of episodic encoding and working memory. Reduced and non-selective activation of frontal regions have also been evident in studies investigating memory retrieval" (S. 75; siehe auch Morcom, Li & Rugg, 2007).

Die altersassoziierten neuronalen Veränderungen scheinen also (mit)ursächlich dafür zu sein, dass ältere TeilnehmerInnen schlechter im Quellengedächtnis als im Item- oder Inhaltsgedächtnis abschnitten (z.B. Raz, 2000). Sie erinnern bspw. schlechter, welche Stimme eine

Information vorgetragen hat, in welcher Farbe oder Schriftart Items präsentiert wurden, oder ob eine Information einem Videoband oder einer Photographie entstammte (Kausler & Puckett, 1981; Ferguson, Hashtroudi & Johnson, 1992; Light, LaVoie, Valencia-Laver, Owens & Mead, 1992; Schacter, Koutstaal, Johnson, Gross & Angell, 1997). Das letzte Beispiel weist bereits darauf hin: Ältere TeilnehmerInnen sind aufgrund der spezifischen Gedächtnisbeeinträchtigungen auch in stärkerem Ausmaß als junge TeilnehmerInnen anfällig für den Falschinformationseffekt (z.B. Mitchell, Johnson & Mather, 2003). Darüber hinaus unterlaufen ihnen auch mehr Fehler im DRM-Paradigma (z.B. Norman & Schacter, 1997; Balota et al., 1999). Letztere sind vermutlich darauf zurückzuführen, dass das semantische Gedächtnis bzw. die Erinnerung an die globale Bedeutung der Information, intakt ist, während die spezifischen Details bzw. der genaue Wortlaut nicht abgerufen werden können (siehe auch Dennis, Kim & Cabeza, 2008). Auf die Leistungen älterer TeilnehmerInnen im imagination-inflation-Paradigma wird im nächsten Kapitel genauer eingegangen.

II.1.3 Imagination inflation bei älteren TeilnehmerInnen

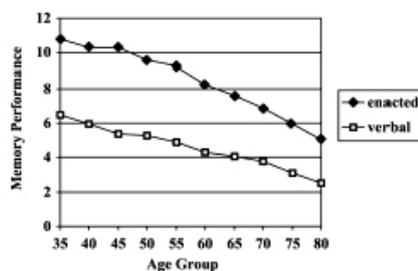


Abb. 13: Ausmaß des enactment-Effekts in verschiedenen Altersgruppen, erfasst anhand des freien Abrufs (nach Nyberg et al., 1996; abgeb. nach Nyberg et al., 2002, S. 836).

Der enactment-Effekt, die bessere Erinnerbarkeit ausgeführter im Vergleich zu nicht ausgeführten (häufig: gelesenen) Handlungsanweisungen, ließ sich auch bei Stichproben älterer TeilnehmerInnen nachweisen. Während einige Studien sogar zu dem Ergebnis kamen, dass nach motorischer Enkodierung keine Unterschiede in der Erinnerungsleistung von jüngeren vs. älteren TeilnehmerInnen nachzuweisen seien (Bäckman & Nilsson, 1984, 1985; Dick, Kean & Sands, 1989), kamen andere Untersuchungen zum gegenteiligen Ergebnis (Guttentag & Hunt, 1988; Nilsson & Craik, 1990; Nyberg, Nilsson & Bäckman, 1992). Im Rahmen einer groß angelegten Studie mit 1000 TeilnehmerInnen im Alter von 35 bis 80 Jahren haben schließlich Rönnlund, Nyberg, Bäckman und Nilsson (2003) zeigen können, dass letzteres Ergebnismuster die Regel zu sein scheint: Eine mit dem Alter einhergehende kontinuierliche Abnahme in der Erinnerungsleistung zeigte sich gleichermaßen für die ausgeführten wie auch für die nur verbal enkodierten Handlungsanweisungen und zwar unter unterschiedlichen Abrufbedingungen (free und cued recall, siehe Abbildung 13). Ein Vorteil der motorischen im Vergleich zur ausschließlich verbalen Enkodierung ließ sich jedoch in allen Altersstufen nachweisen und war im Ausmaß über die Altersstufen hinweg vergleichbar. Die AutorInnen merken allerdings an, dass der Zusammenhang von Erinnerungsleistungen und Alter teilweise verschwindet, wenn der Faktor Bildung aus den Daten herauspar-

tialisiert wird. Davon nicht betroffen war die Gruppe der 60-80 Jährigen, in der sich auch zwischen den Regressionsresiduen ein enger Zusammenhang ergab. Demnach scheint mit Sicherheit für diese Gruppe der sog. jungen und alten Alten ein enger Zusammenhang von Erinnerungsleistung und Alter zu bestehen, der sich nicht auf andere Faktoren zurückführen lässt. Auf diesen Daten aufbauend konnten Nyberg, Persson und Nilsson (2002) zeigen, dass TeilnehmerInnen, die nicht von motorischer Enkodierung profitierten, mutmaßlich über beeinträchtigte motorische Systeme verfügen. Dabei waren in dieser Gruppe der „Nicht-Profitierer“ ältere Personen überproportional repräsentiert. Der Befund lässt sich gut in die skizzierte Rahmenvorstellung integrieren, und untermauert die Vorstellung, dass die Erinnerung von Handlungen mit einer Reaktivierung von motorischen Informationen einhergeht.

Tab. 33: Ergebnisse des ersten Experiments von Thomas und Bulevich (2006). *Ausgeführt*-Antworten im Quellengedächtnistest (abgeb. nach Thomas & Bulevich, 2006, S. 383).

Variable and no. of imaginings	2-week retention interval				2-day retention interval			
	Older		Younger		Older		Younger	
	M	SE	M	SE	M	SE	M	SE
Not presented in Session 1								
0	.04	.03	.03	.01	.01	.01	.00	.00
1	.17	.05	.10	.02	.08	.01	.01	.01
5	.41	.07	.20	.06	.17	.05	.01	.01
Imagined in Session 1								
0	.14	.04	.04	.02	.04	.02	.00	.00
1	.39	.06	.11	.03	.16	.05	.02	.01
5	.46	.06	.23	.05	.27	.05	.01	.01
Performed in Session 1								
0	.48	.06	.68	.06	.92	.02	.94	.02
1	.76	.05	.87	.04	.96	.02	.98	.01
5	.89	.02	.89	.04	.98	.02	.97	.04

Auf Basis des skizzierten Musters altersassoziierter Veränderungen von Erinnerungsleistungen schlussfolgerten Thomas und Bulevich (2006), dass ältere verglichen mit jüngeren Personen auch einen stärkeren imagination-inflation-Effekt aufweisen müssten, also anfälliger für falsche Erinnerungen an die eigene Handlungsausführung durch Vorstellen sein sollten. In einem ersten Experiment durchliefen zwei Gruppen älterer und zwei Gruppen jüngerer TeilnehmerInnen das „klassische“ imagination-inflation-

Paradigma. In einer ersten Phase wurden sie gebeten, Handlungen entweder auszuführen oder nicht (was in dieser Studie bedeutete, dass die TeilnehmerInnen sich die Ausführung vorstellen sollten). In einer zweiten Phase, die einen Tag später stattfand, wurden sie gebeten, sich die Ausführung einiger alter und neuer Handlungsanweisungen einmal oder fünfmal vorzustellen. Variiert wurde das Intervall zwischen dieser zweiten Phase und dem Abruf (Phase 3). Jeweils eine der beiden Altersgruppen wurde gebeten, zwei Tage später wiederzukommen; die andere Gruppe sollte (wie üblich) erst nach zwei Wochen zur dritten Phase erscheinen. In dieser dritten Phase erhielten die ProbandInnen dann die Aufgabe, für jede der präsentierten Handlungsanweisungen zu entscheiden, ob diese aus der ersten Phase stammte und ausgeführt wurde, ob diese aus der ersten Phase stammte und nicht ausgeführt wurde, oder ob diese neu war. Wie erwartet trat bei den älteren TeilnehmerInnen ein höherer imagination-inflation-Effekt auf als bei den jüngeren (siehe Tabelle 33). Es zeigte sich auch, dass

das verkürzte Intervall zwischen Enkodierung und Abruf den imagination-inflation-Effekt bei den älteren TeilnehmerInnen entscheidend reduzierte; es eliminierte gar falsche Erinnerungen an die Handlungsausführung bei den jüngeren. In einer zweiten Studie variierten Thomas und Bulevich neben dem Intervall auch die Instruktion beim Abruf. Die TeilnehmerInnen erhielten entweder nur die bisher stets eingesetzte Standardinstruktion, oder eine zusätzliche Instruktion zur Quellenprüfung, die sie veranlassen sollte, verstärkt auf Hinweise für bzw. gegen die eigene Handlungsausführung zu achten (siehe Abschnitt I.5.4.3). Es zeigte sich, dass beide Gruppen älterer TeilnehmerInnen, die diese Instruktion erhalten hatten, signifikant weniger falsche Erinnerungen produzierten als diejenigen, die keine derartige Abrufhilfe bekommen hatten (siehe Tabelle 34). Bei zweiwöchigem Intervall näherte sich das Ausmaß zwar demjenigen jüngerer TeilnehmerInnen an, lag aber immer noch über deren Quote. Bei zweitägigem Intervall erreichte der Prozentsatz falscher *ausgeführt*-Antworten nach Instruktion zur Quellenprüfung bei den älteren TeilnehmerInnen einen Tiefstand; dennoch war ein imagination-inflation-Effekt nachweisbar, während die jüngeren TeilnehmerInnen bei dem verkürzten Intervall auch ohne zusätzliche Instruktion keinerlei falsche Erinnerungen produzierten (siehe Experiment 1). Die jüngeren TeilnehmerInnen, die die zweite Sitzung nach zwei Wochen absolvierten, profitierten in keiner Weise von den zusätzlichen Hinweisen beim Abruf.

Eine mögliche Erklärung dieses Ergebnismusters könnte darin bestehen, dass die älteren TeilnehmerInnen nach Lesen der Instruktion zur Quellenprüfung einfach ein strikteres Antwortkriterium ansetzten. Diese Vorstellung lässt sich allerdings insofern nicht mit den vorhandenen Daten vereinbaren, als die Quote korrekter *ausgeführt*-Antworten in den Gruppen mit und ohne Quellenprüfung konstant blieb. Denkbar wäre jedoch, dass die TeilnehmerInnen, die keine Instruktion zur Quellenprüfung erhielten, globale Kriterien, insb. Vertrautheit, nutzten, während diejenigen TeilnehmerInnen, die zur Quellenprüfung aufgefordert wurden, (wie instruiert) nach spezifischen Kriterien suchten und ihre Erinnerungsurteile auf diese basierten. Es wäre aber auch vorstellbar, dass die älteren TeilnehmerInnen in beiden Fällen zwar spezifi-

Tab. 34: Ergebnisse des zweiten Experiments von Thomas und Bulevich (2006). *Ausgeführt*-Antworten im Quellengedächtnistest (abgeb. nach Thomas & Bulevich, 2006, S. 385).

Variable and no. of imaginings	2-week older		2-week younger		2-day older	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
No retrieval support						
Not presented in Session 1						
0	.04	.03	.02	.01	.02	.02
1	.19	.05	.10	.02	.11	.03
5	.46	.07	.20	.06	.20	.04
Imagined in Session 1						
0	.07	.02	.04	.01	.02	.01
1	.24	.04	.12	.02	.14	.02
5	.50	.05	.24	.05	.30	.04
Performed in Session 1						
0	.47	.04	.69	.04	.95	.01
1	.75	.04	.90	.03	.97	.01
5	.92	.02	.91	.03	.99	.01
Retrieval support						
Not presented in Session 1						
0	.04	.03	.02	.01	.01	.01
1	.13	.03	.09	.02	.04	.02
5	.31	.06	.19	.06	.06	.02
Imagined in Session 1						
0	.06	.02	.04	.01	.03	.01
1	.18	.03	.13	.02	.06	.01
5	.37	.04	.24	.05	.12	.02
Performed in Session 1						
0	.50	.04	.68	.04	.94	.02
1	.73	.04	.87	.03	.97	.01
5	.89	.03	.89	.03	.99	.03

sche, qualitative Kriterien nutzten, diese aber ohne Instruktion zur Quellenprüfung nicht in vollem Ausmaß evaluierten (siehe z.B. auch McDaniel, Lyle, Butler & Dornburg, 2008). So folgern Thomas und Bulevich aus ihren Ergebnissen dann auch, dass die Probleme, die ältere TeilnehmerInnen in Quellengedächtnistests aufweisen, auf eine ineffiziente cue-Nutzung beim Abruf zurückzuführen sind. Während die jüngeren TeilnehmerInnen offensichtlich automatisch nach spezifischen Hinweisen auf die Quellenzugehörigkeit suchen und diese in ihrer ganzen Bandbreite einbeziehen, scheinen ältere TeilnehmerInnen eine solche Strategie nicht zu verfolgen und ihre Erinnerungsurteile eher auf globale Kriterien oder auf einen Ausschnitt spezifischer Erinnerungsspuren zu basieren.

McDaniel et al. (2008) unterzogen in einer erst kürzlich publizierten Studie die drei erläuterten möglichen Mechanismen, die ältere TeilnehmerInnen beim Abruf zu falschen Erinnerungsurteilen verleiten können, einer empirischen Überprüfung anhand des imagination-inflation-Paradigmas (Fehlattritionen aufgrund 1. der verminderten Qualität spezifischer Erinnerungsspuren, 2. des reduzierten Spektrums berücksichtigter spezifischer Erinnerungsspuren, 3. der übermäßigen Nutzung globaler Erinnerungscharakteristika, besonders von Vertrautheit). Durch eine Abwandlung des Designs entfiel die Unterscheidung zwischen einer ersten und zweiten Phase; im Rahmen einer einzigen Phase wurden Handlungsanweisungen entweder nur ausgeführt, nur vorgestellt oder beides, und zwar mit unterschiedlichen Häufigkeiten. In einer zweiten Sitzung, die zwei Wochen später stattfand, wurden den TeilnehmerInnen alte und neue Handlungsanweisungen zusammen mit zwei Skalen, einer *ausgeführt*- und einer *vorgestellt*-Skala, präsentiert. Sie wurden gebeten, auf beiden Skalen anzugeben, wie häufig (0 bis maximal 8 Mal) sie eine Handlung ausgeführt und/ oder sich vorgestellt hatten. Das Ergebnismuster ließ sich nicht mit der Annahme einer Fehlattrition von Vertrautheit vereinbaren. Vielmehr sprach es dafür, dass die älteren – wie auch die jüngeren – TeilnehmerInnen ihre Erinnerungsurteile auf sensorische Details basieren. Da ältere Personen sensorische Details jedoch weniger elaboriert enkodieren, sind die Erinnerungsspuren von Vorstellung und Ausführung ähnlicher und daher schwieriger zu unterscheiden. Auch wenn das Ergebnismuster nicht vollständig mit der Hypothese zu vereinbaren ist, dass die älteren TeilnehmerInnen kognitive Operationen, die mit Vorstellungen assoziiert sein sollten, nicht in ausreichendem Maße in ihre Erinnerungsurteile einfließen ließen, merken die AutorInnen an, dass ein derartiger Mechanismus eine zusätzliche Rolle spielen könnte und auch mit den Befunden von Thomas und Bulevich (2006) vereinbar wäre. Demnach könnte der Effekt nicht nur auf die verminderte Qualität der Erinnerungsspuren zurückzuführen sein, sondern – wie bereits ausgeführt – auch partiell auf eine ineffiziente cue-Nutzung beim Abruf, speziell darauf, dass Erinnerungen an kognitive Operationen, die mit Vorstellungen assoziiert sein sollten, nicht berücksichtigt werden.

McDaniel et al. teilten die Stichprobe ältere TeilnehmerInnen unter Nutzung der Testbatterie zur Erfassung sog. frontaler Hirnfunktionen von Glisky et al. (1995) nach ihrem Abschneiden in zwei Gruppen ein (McDaniel, Butler & Dornburg, 2006). Sie erwarteten, dass die Gruppe mit schlechteren „frontalen Leistungen“ auch anfälliger für den imagination-inflation-Effekt sein würde. Dies war nicht der Fall. Beide Gruppen wiesen äußerst ähnliche Ergebnismuster auf. Lediglich nach mehrmaliger Handlungsausführung schrieb die Gruppe mit den besseren „frontalen Leistungen“ die entsprechenden Handlungsanweisungen seltener der Kategorie „vorgestellt“ zu. Dieser Befund ist vor dem skizzierten Hintergrund einer häufig nachgewiesenen Beziehung zwischen frontaler Gehirnaktivität und Quellengedächtnisurteilen äußerst erstaunlich. Die AutorInnen diskutieren mehrere Erklärungsmöglichkeiten und kommen letztlich zu dem Schluss, dass möglicherweise die Art der Quelle entscheidend ist. Demnach könnte man periphere Quellen (z.B. die Stimme, die einen Satz sprach) und zentrale Quellen (z.B. die Modalität, in der eine Handlungsanweisung verarbeitet wurde) unterscheiden. Während für die Enkodierung, den Prozess des bindings und den (aufwändigeren) Abruf peripherer Quellen frontale Prozesse notwendig sind, könnten diese für zentrale Quellen nicht notwendig sein, da diese Bestandteil des Erinnerungsinhaltes sind. Gerade bei der Unterscheidung von Ausführung und Vorstellung einer Handlung, also der Enkodiermodalität als zentraler Qualität, könnte daher vorwiegend der mediale Temporallappen aktiviert sein.

Neben der Studie von Thomas und Bulevich (2006) liefern auch andere Untersuchungen Hinweise darauf, dass ältere Menschen unter bestimmten Bedingungen in der Lage sind, falsche Erinnerungen im Sinne von Fehlern bei der Quellenzuordnung zu reduzieren. Neben Hinweisen beim Abruf haben sich auch explizite Hinweise auf die Quelle im Rahmen der Enkodierung (z.B. Glisky, Rubin & Davidson, 2001; Naveh-Benjamin & Craik, 1995) als fruchtbar erwiesen. Solche Manipulationen wurden jedoch noch nicht im Rahmen des imagination-inflation-Paradigmas eingesetzt. Derartige Studien helfen, die Natur des Problems näher zu charakterisieren sowie Möglichkeiten zu dessen Überwindung aufzuzeigen.

II.2 Befunde zu motorischer Simulation bei älteren TeilnehmerInnen

Da das Forschungsgebiet motorischer Simulation noch recht jung ist, wurden differentielle Untersuchungen bisher kaum durchgeführt. Es existiert jedoch eine Studie, die sich mit Altersunterschieden in der corticalen Aktivierung im Kontext von motorischer Simulation auseinandergesetzt hat. Léonard und Tremblay (2007) nutzten die Methode der TMS, um motorisch evozierbare Potentiale hervorzurufen. Gemessen wurden die MEPs in zwei Handmuskeln, dem ersten musculus interosseus dorsalis (first dorsal interosseous, FDI), der den aufgabenspezifischen Muskel darstellte, und dem musculus abductor digiti minimi (abductor digiti minimi, ADM). Die ProbandInnen hatten vier verschiedene Aufgaben: Beobachtung einer Handlung, Vorstellung der Handlung, Ausführung der Handlung und keine Aktivität (Kontrollbedingung). Signifikante motorische Aktivität konnte für beide Altersgruppen in allen drei Aktivitätsbedingungen nachgewiesen werden, was die bereits berichteten Befunde zu motorischer Simulation repliziert. Ein Alterseffekt zeigte sich jedoch in der Selektivität der Aktivierung: Während die jüngeren TeilnehmerInnen v.a. selektiv den FDI aktivierten, wurde bei den älteren TeilnehmerInnen häufig eine Aktivität von FDI und ADM zusammen registriert. Die Autoren schlussfolgern „This loss [in selectivity] probably reflects age-related alterations at the central level affecting the level of cortical inhibition leading to more widespread facilitation when actions are either executed or internally simulated“ (S. 174). Demnach könnten also alterskorrelierte Veränderungen in der Funktionsweise des corticospinalen Systems Ursache der reduzierten Selektivität sein; ein Befund, der mit den Ergebnissen der Analyse von Nyberg, Persson und Nilsson (2002) konvergiert. Die Forscher hatten gezeigt, dass TeilnehmerInnen, die nicht von motorischer Enkodierung profitierten, mutmaßlich über beeinträchtigte motorische Systeme verfügen, und dass hiervon besonders ältere Personen betroffen waren.

II.3 Résumé und Ableitung der Fragestellung

Aus dem beschriebenen theoretisch-empirischen Hintergrund wird deutlich, dass der observation-inflation-Effekt, der bisher ausschließlich an Stichproben jüngerer TeilnehmerInnen untersucht wurde, bei älteren TeilnehmerInnen auf andersartigen Mechanismen beruhen könnte, da diese vor dem Hintergrund neurobiologischer Veränderungen altersassoziierte Beeinträchtigungen von Gedächtnisleistungen aufweisen (z.B. Craik, 2000). Während der observation-inflation-Effekt bei jüngeren Gesunden auf eine Fehlattribution von sensorischen Erinnerungsspuren und – vermutlich sogar in stärkerem Ausmaß – auch auf eine Fehlattribution motorischer Erinnerungsspuren zurückzuführen ist, könnten Spektrum und Gewichtung der Mechanismen bei älteren TeilnehmerInnen anders gelagert sein. Aus Untersuchungen zu altersassoziierten Veränderungen der Gedächtnisleistung geht einerseits hervor, dass ältere TeilnehmerInnen ihre Erinnerungsurteile häufig nicht auf die ganze Bandbreite verfügbarer spezifischer Erinnerungsspuren basieren und sich gleichzeitig eher auf globale Erinnerungsspuren, insb. die Vertrautheit von Informationen, stützen. Andererseits scheinen die Erinnerungsspuren älterer im Vergleich zu denjenigen jüngerer TeilnehmerInnen deutlich blässer bzw. schwächer zu sein, vermutlich, da diese aufgrund begrenzter Verarbeitungsressourcen bzw. reduzierter Aktivität des frontalen Cortex nicht in vollem Ausmaß enkodiert und/ oder zu einem Gesamtbild zusammengefügt werden können. Demnach stehen den älteren TeilnehmerInnen beim Abruf weniger spezifische Erinnerungsspuren, z.B. sensorische Details, zur Verfügung, was wiederum eine unvollständige Evaluation der Erinnerungsspuren und einen übermäßigen Einbezug von Vertrautheit begünstigt und eine eigenständige Reproduktion von Erlebtem im Vergleich zu einer Erinnerung mit Abrufhilfen erschwert. (z.B. McDaniel et al., 2008).

Da im klassischen imagination-inflation Paradigma durch die zahlreichen Wiederholungen innerhalb der zweiten Phase sowohl die Vertrautheit der Informationen als auch die spezifischen sensorischen Eindrücke gestärkt werden, zudem aufgrund der Einschränkungen bei der Enkodierung/ dem binding die Repräsentation von ausgeführten und vorgestellten Handlungen ähnlicher sein sollte, müssten ältere TeilnehmerInnen auch in stärkerem Ausmaß dem imagination-inflation-Effekt unterliegen. Thomas und Bulevich (2006) haben nachgewiesen, dass dies der Fall ist. Zudem konnten sie zeigen, dass die älteren TeilnehmerInnen – wenn ihnen eine spezielle Instruktion zur Quellenprüfung gegeben wurde – in der Lage waren, das Ausmaß des Effekts signifikant zu reduzieren. Möglicherweise urteilten sie nach Lesen der Instruktion eher aufgrund der vorhandenen spezifischen und weniger aufgrund von globalen Erinnerungsspuren. Die Studien von McDaniel et al. (2006, 2008) weisen demgegenüber auf eine stärkere Ähnlichkeit des Vorgehens jüngerer und älterer TeilnehmerInnen hin. Demnach könnten ältere TeilnehmerInnen ihre Erinnerungsspuren auch ohne Instruktion zur Quellenprüfung

auf spezifischen sensorischen Informationen und weniger auf Vertrautheit basieren, die Erinnerungsspuren von ausgeführten und vorgestellten Handlungen sind jedoch ähnlicher als dies bei den jüngeren TeilnehmerInnen der Fall ist. Dieser Mechanismus vermag allerdings nicht zu erklären, warum die TeilnehmerInnen in der Studie von Thomas und Bulevich (2006) in der Lage waren, nach Lesen einer Instruktion zur Quellenprüfung das Ausmaß falscher Erinnerungen zu reduzieren – schließlich werden die Erinnerungsspuren durch die nachträglichen Instruktion in keinster Weise tangiert. Eine weitere Erklärungsmöglichkeit, die auch den von Thomas und Bulevich gefundenen Effekt erklären könnte, besteht daher darin, dass ältere TeilnehmerInnen, die keine spezielle Anleitung beim Abruf erhalten, nicht die volle Bandbreite der Erinnerungsspuren evaluieren. Speziell kognitive Operationen, die mit Vorstellungen assoziiert sein sollten, könnten nicht ausreichend berücksichtigt werden.

Insgesamt lässt sich also festhalten, dass die Frage, welcher Mechanismus falschen Erinnerungen an einfache Handlungen bei älteren TeilnehmerInnen zugrunde liegt, noch nicht vollkommen geklärt ist. Die bisher in Betracht gezogenen möglichen Mechanismen beziehen sich einerseits auf die Qualität der Erinnerungsspuren selbst, andererseits auf die Art der Beurteilung dieser (z.B. McDaniel et al., 2008).

Erstaunlicherweise konnten McDaniel et al. (2006) das Ausmaß der Quellenverwechslungen älterer TeilnehmerInnen im Rahmen des imagination-inflation-Paradigmas jedoch nicht mit dem Abschneiden in verschiedenen Testverfahren in Zusammenhang bringen, die sog. frontale Gehirnfunktionen erfassen; Gehirnfunktionen, die normalerweise in Zusammenhang mit der Enkodierung und dem Abruf von Quelleninformationen stehen.

Hinsichtlich der Prozesse der Simulation bzw. Repräsentation von motorischen Informationen lässt sich festhalten, dass ältere Menschen zwar – genau wie jüngere – die Handlungen anderer automatisch zu simulieren scheinen, dieser Prozess allerdings (möglicherweise aufgrund beeinträchtigter motorischer Systeme) weniger selektiv ausfällt (Léonard & Tremblay, 2007). Der mögliche Einfluss motorischer Repräsentationen auf falsche Erinnerungen für die Handlungsausführung wurde – genau wie für die jungen TeilnehmerInnen – bei der Untersuchung älterer Personen bisher nicht berücksichtigt.

Die Frage, ob eine einfache Handlung ausgeführt wurde oder nicht, ist von großer alltäglicher Relevanz – man denke nur daran, welche Konsequenzen es hätte, fälschlich zu erinnern, eine Tablette schon genommen zu haben (Untermedikation), bzw. umgekehrt, fälschlich zu erinnern, eine Tablette noch nicht genommen zu haben (Übermedikation). Daher ist die Eingrenzung der Mechanismen, die solchen Erinnerungsfehlern zugrunde liegen, gerade im Hinblick

auf ältere Menschen, die besonders anfällig für derartige Verwechslungen sind, äußerst wichtig. Wären nämlich die Mechanismen eingegrenzt, ließen sich möglicherweise Strategien ableiten, um derartigen falschen Erinnerungen entgegenzuwirken. Dabei würde der Einsatz von Beobachtung im Gegensatz zur Vorstellung der Handlungsausführung eine methodologische Verbesserung darstellen, da er eine bessere Kontrollierbarkeit und Vergleichbarkeit gewährleistet.

Vor dem geschilderten Hintergrund ist anzunehmen, dass sich – analog zum imagination-inflation-Effekt – auch ein stärkerer observation-inflation-Effekt der älteren im Vergleich zu den jüngeren TeilnehmerInnen zeigen wird. Im Hinblick auf diesen lässt sich zusammengefasst vermuten, dass er auf die größere Ähnlichkeit sensorischer Erinnerungsspuren der Quellen Beobachtung und Ausführung, auf eine reduzierte Evaluation der Bandbreite und Art der Erinnerungsspuren und/ oder auf eine zu starke Berücksichtigung von Vertrautheit zurückgehen könnte. Inwiefern eine Fehlattribution bisher nicht beachteter motorischer Repräsentationen bei älteren TeilnehmerInnen eine Rolle spielt, ist vor dem Hintergrund eingeschränkter Funktionalität motorischer Systeme schwerlich abzuschätzen. Auch wurde der Informationsursprung als weitere potentielle Quelle von Fehlattributionen bisher nicht berücksichtigt.

Um das Spektrum und die Gewichtung möglicher Mechanismen eingrenzen zu können, wäre demnach analog zu den Experimenten I bis III eine ganze Untersuchungsserie notwendig, die u.a. auch bisher fehlende Kontrollgruppen realisieren müsste, um Einflüsse des Informationsursprungs und der Verarbeitungstiefe bzw. Vertrautheit zu erfassen. Aus Kapazitätsgründen kann im Rahmen der vorliegenden Arbeit jedoch nur eine einzige Untersuchung realisiert werden. Zunächst soll daher durch Nutzung des modifizierten Paradigmas zur Untersuchung falscher Erinnerungen an die Handlungsausführung aufgrund von Beobachtung überprüft werden, ob der Untersuchungsansatz bei älteren TeilnehmerInnen einsetzbar ist und ob sich überhaupt der vermutete Alterseffekt zeigt. Daher soll eine Gruppe von älteren TeilnehmerInnen anhand des in Studie II entwickelten verschlankten Designs untersucht und hinsichtlich des observation-inflation-Effekts mit einer Gruppe jüngerer TeilnehmerInnen verglichen werden. Eine weitere Gruppe älterer TeilnehmerInnen soll dasselbe Paradigma durchlaufen, jedoch mit auf wenige Minuten verkürztem Intervall zwischen Enkodierung und Abruf. Da bereits Thomas und Bulevich (2006) auch nach Verkürzung des Intervalls ein zwar gegenüber dem langen Intervall reduziertes, aber immer noch substantielles Ausmaß an falschen Erinnerungen aufgrund von Vorstellung fanden, wird davon ausgegangen, dass sich ein ähnliches Muster auch bzgl. des observation-inflation-Effekts zeigen wird (obwohl es sich bei der anvisierten um eine stärkere Verkürzung des Intervalls handelt als die von Thomas und Bulevich realisierte). Da in derselben Studie gezeigt werden konnte, dass jüngere TeilnehmerInnen bei

verkürztem Intervall keinerlei Probleme bei der Quellendiskrimination aufweisen, soll auf eine entsprechende Kontrollgruppe verzichtet werden.

Die gewählte Variation des Intervalls bietet v.a. zwei praktische Vorteile: Einerseits kann so überprüft werden, ob – für weitere Untersuchungen älterer TeilnehmerInnen – durch eine einmalige anstatt der aufwändigeren zweimaligen Untersuchungsteilnahme derselbe Effekt erzielt werden kann. Andererseits war geplant, die Daten gesunder in weiteren Untersuchungen mit denjenigen (leicht) dementer älterer TeilnehmerInnen zu vergleichen. Der sinnvolle Einschluss dieser Population, die insb. durch beeinträchtigte Gedächtnisleistungen gekennzeichnet ist, ist allerdings nur denkbar, wenn Enkodierung und Abruf relativ zeitnah aufeinander folgen.

Die Variation des Intervalls führt jedoch nicht zu einer eindeutigen Eingrenzung potentieller Mechanismen, da die drei Prozesse durch die experimentelle Manipulation nicht selektiv beeinflusst werden: Durch die Verkürzung des Behaltensintervalls wird die Qualität der spezifischen Erinnerungsspuren deutlich gesteigert. Gleichzeitig sollte dies die Tendenz reduzieren, Vertrautheit als indikativ für die Handlungsausführung anzusehen. Vermutlich wird durch die Verfügbarkeit einer detailreichen Erinnerung dann auch die Evaluation der ganzen Bandbreite dieser Erinnerungsspuren angestoßen. Demnach könnte der Effekt nach kurzem Intervall wiederum auf anderen Mechanismen beruhen als derjenige nach langem Intervall. Es ist jedoch anzunehmen, dass die Betrachtung des Ergebnismusters über alle drei Gruppen hinweg Hinweise darauf liefern kann, welche Mechanismen für die Gruppe älterer Gesunder eine über- und welche nur eine untergeordnete Rolle spielen könnten. Diesen Hinweisen könnte dann in weiteren Untersuchungen nachgegangen werden.

Vor diesem Hintergrund lassen sich folgende Hypothesen aufstellen, die sich nicht auf die zugrundeliegenden Mechanismen, sondern auf den Nachweis des Phänomens an sich beziehen:

Ältere Personen weisen (sowohl nach langem als auch nach kurzem Intervall) zumeist einen observation-inflation-Effekt auf. Bei gleichem Intervall produzieren ältere Personen zumeist mehr falsche Erinnerungen an die Handlungsausführung als jüngere Personen. Bei verkürztem Intervall produzieren ältere Personen weniger falsche Erinnerungen an die Handlungsausführung als bei langem Intervall.

Um zusätzlich explorativ Zusammenhänge zwischen dem Ausmaß an falschen Erinnerungen für die Handlungsausführung und sog. frontalen vs. medial-temporalen Hirnfunktionen bestimmen und so möglicherweise ebenfalls die relevanten Mechanismen eingrenzen zu können, sollen diese in Anlehnung an McDaniel et al. (2006) anhand verschiedener Testverfahren erfasst und mit dem Ausmaß der Gedächtnistäuschung korreliert werden (siehe auch Glisky et al., 1995). So fanden bspw. Henkel et al. (1998) nach einem kurzen Intervall keinen Zusammenhang der Quellengedächtnisleistung älterer TeilnehmerInnen mit Maßen für sog. frontale Funktionen, nach langem Intervall jedoch schon (zusätzlich ergaben sich Zusammenhänge zu medial-temporalen Funktionen). Demnach könnten frontale Hirnareale eher in den aufwändigeren, strategischen Abruf bzw. in komplexere kognitive Bewertungsprozesse involviert sein, während für die noch gut zugängliche Quellenzuordnung nach kurzem Intervall die Aktivität v.a. medial-temporalen Regionen ausreicht.

II.4 Experiment IV

II.4.1 Versuchsplanung

II.4.1.1 Operationalisierung und Variablenvalidität

In Experiment IV sollte das bereits mehrfach bei gesunden jungen TeilnehmerInnen eingesetzte Paradigma genutzt werden, um Effekte des Lebensalters auf den observation-inflation-Effekt zu untersuchen. Dabei sollten zwei Altersgruppen gegenübergestellt werden: jüngere und ältere TeilnehmerInnen. In Anlehnung an den Großteil kognitionspsychologischer Studien zu Alterseffekten sollte es sich bei den jüngeren TeilnehmerInnen um Personen im Alter von etwa 20 Jahren handeln; die älteren TeilnehmerInnen sollten über 60 Jahre alt sein (z.B. Balota, Dolan & Duchek, 2000). Daneben sollten – in Anlehnung an Thomas und Bulevich (2006) – zwei Abrufintervalle realisiert werden: 2 Wochen und wenige Minuten. In letzterem Fall konnte also das gesamte Paradigma innerhalb von einer Sitzung durchlaufen werden. Dabei soll für das verkürzte Intervall kein Altersvergleich angestellt werden, da vor dem Hintergrund der Ergebnisse von Thomas und Bulevich davon ausgegangen werden kann, dass die jüngeren TeilnehmerInnen fehlerlos abschneiden würden.

Die Untersuchung wurde ebenfalls an der University of Alberta, Edmonton, Kanada, durchgeführt. Durchführung, Material und Ablauf entsprechen vollständig der Standardbedingung des zweiten Experiments (Beobachten in der zweiten Phase und Abruf ohne Abrufinstruktion). Die dort berichtete Stichprobe von 14 Studierenden stellt daher auch gleichzeitig die Referenzgruppe für die Stichprobe älterer TeilnehmerInnen (langes Intervall) dar.

Vor dem Hintergrund der gewählten Operationalisierungen ergeben sich bzgl. der drei aufgestellten Hypothesen

- (1) Ältere Personen weisen (sowohl nach langem als auch nach kurzem Intervall) zumeist einen observation-inflation-Effekt auf.
- (2) Bei gleichem (langem) Intervall produzieren ältere Personen zumeist mehr falsche Erinnerungen an die Handlungsausführung als jüngere Personen.
- (3) Bei verkürztem Intervall produzieren ältere Personen weniger falsche Erinnerungen an die Handlungsausführung als bei langem Intervall.

folgende Vorhersagen (die Notation bezieht sich auf den im folgenden Abschnitt präsentierten Versuchsplan):

- (1.1) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die älteren TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und den Erinnerungstest nach einem Intervall von zwei Wochen durchliefen.

$$H_1: \mu_{A2B1C2D1} > \mu_{A2B3C2D1}; H_0: \mu_{A2B1C2D1} \leq \mu_{A2B3C2D1}$$

- (1.2) Der Mittelwert der falschen *ausgeführt*-Antworten ist nach fünfmaliger im Vergleich zu nullmaliger Präsentation signifikant höher, wenn die älteren TeilnehmerInnen in der zweiten Phase die Handlungsausführung beobachteten und den Erinnerungstest nach einem Intervall von zwanzig Minuten durchliefen.

$$H_1: \mu_{A2B1C2D2} > \mu_{A2B3C2D2}; H_0: \mu_{A2B1C2D2} \leq \mu_{A2B3C2D2}$$

- (2.) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist signifikant höher, wenn die älteren TeilnehmerInnen den Erinnerungstest nach zwei Wochen durchliefen, als wenn die jüngeren TeilnehmerInnen den Erinnerungstest nach zwei Wochen durchliefen.

$$H_1: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} > \mu_{A2B1C1D1} - \mu_{A2B3C1D1};$$

$$H_0: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} \leq \mu_{A2B1C1D1} - \mu_{A2B3C1D1}$$

- (3.) Die Differenz der Mittelwerte falscher *ausgeführt*-Antworten nach fünfmaliger und nullmaliger Präsentation ist signifikant höher, wenn die älteren TeilnehmerInnen den Erinnerungstest nach einem zweiwöchigen Intervall durchliefen, als wenn sie ihn nach einem zwanzigminütigen Intervall durchliefen.

$$H_1: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} > \mu_{A2B1C2D2} - \mu_{A2B3C2D2};$$

$$H_0: \mu_{A2B1C2D1} - \mu_{A2B3C2D1} \leq \mu_{A2B1C2D2} - \mu_{A2B3C2D2}$$

Im Folgenden sollen nur die Unterschiede, die sich bei der Durchführung der Untersuchung mit den älteren im Vergleich zu den jüngeren TeilnehmerInnen ergaben, dargestellt werden.

II.4.1.2 Versuchsplananlage und Versuchsplan

Der Untersuchung lag – wie auch in Studie II – die intraindividuell variierte, dreistufige Variable „Art der Enkodierung in der 1. Phase“ (UV-A: *ausgeführt*, *gelesen*, *nicht enkodiert*) und die ebenfalls intraindividuell variierte, dreistufige Variable „Frequenz der Präsentation in der 2. Phase“ (UV-B: 5x, 1x, 0x) zugrunde. Es wurden drei Gruppen benötigt, um die anvisierten Untersuchungsziele zu erreichen: Eine Gruppe jüngerer und eine Gruppe älterer TeilnehmerInnen, die das Paradigma in zwei Sitzungen mit zweiwöchigem Abstand durchliefen. Eine weitere Gruppe älterer TeilnehmerInnen sollte die gesamte Prozedur innerhalb einer Sitzung durchlaufen. Die beiden interindividuell variierten, zweistufigen Variablen lauteten also Altersgruppe (UV-C: jüngere, ältere TeilnehmerInnen) und Intervall zwischen Enkodierung und Abfrage (UV-D: 20 Minuten, 2 Wochen). Aus Gründen mangelnder Verfügbarkeit älterer Teil-

nehmerInnen musste auf Kontrollgruppen, die in der zweiten Phase die Handlungsanweisungen nur lasen, aber nicht die Handlungsausführung beobachteten, verzichtet werden. Als abhängige Variablen wurden wiederum Rekognitions- und Quellengedächtnisurteile sowie Sicherheitsratings bzgl. der Erinnerungsurteile erhoben. Somit resultierte eine vierfaktorielle Versuchsplananlage (VPL-A4, Hussy & Jain, 2002; siehe Tabelle 35). Es ist zu beachten, dass es sich beim Alter um ein organismisches Merkmal handelt; daher erscheint in der Bezeichnung des Versuchsplans erstmals die Deklaration einer quasi-experimentellen Variable.

Die Umsetzung dieser Versuchsplananlage in den konkreten Versuchsplan (VPL4-RRQR) basierte auf denselben Überlegungen wie in Studie I.

Tab. 35: Versuchsplan des vierten Experiments.

				UV-C: Altersgruppe		
				jüngere TeilnehmerInnen	ältere TeilnehmerInnen	
				UV-D: Intervall zwischen Enkodierung (Phase 1) und Abfrage (Phase 3)		
				zwei Wochen	zwei Wochen	zwanzig Minuten
UV-A: Art der Enkodierung in Phase 1	ausgeführt	UV-B: Frequenz der Präsentation in Phase 2	5x	A ₁ B ₁ C ₁ D ₁	A ₁ B ₁ C ₂ D ₁	A ₁ B ₁ C ₂ D ₂
			1x	A ₁ B ₂ C ₁ D ₁	A ₁ B ₂ C ₂ D ₁	A ₁ B ₂ C ₂ D ₂
			0x	A ₁ B ₃ C ₁ D ₁	A ₁ B ₃ C ₂ D ₁	A ₁ B ₃ C ₂ D ₂
	gelesen		5x	A ₂ B ₁ C ₁ D ₁	A ₂ B ₁ C ₂ D ₁	A ₂ B ₁ C ₂ D ₂
			1x	A ₂ B ₂ C ₁ D ₁	A ₂ B ₂ C ₂ D ₁	A ₂ B ₂ C ₂ D ₂
			0x	A ₂ B ₃ C ₁ D ₁	A ₂ B ₃ C ₂ D ₁	A ₂ B ₃ C ₂ D ₂
	nicht enkodiert		0x	A ₃ B ₃ C ₁ D ₁	A ₃ B ₃ C ₂ D ₁	A ₃ B ₃ C ₂ D ₂

II.4.1.3 Signifikanzniveau, Betafehler, Effektgröße und optimaler Stichprobenumfang

In der Vorgängerstudie im engeren Sinne (Thomas & Bulevich, 2006) wurden zwar keine Effektgrößen berichtet, es konnte aber auf der Basis der verfügbaren Daten geschätzt werden, dass es sich um große Effekte handelte. Von einer Übertragbarkeit der Ergebnisse zum imagination- auf den observation-inflation-Effekt wurde ausgegangen. Daher wurde – analog zu den Studien an jungen TeilnehmerInnen – vor dem Hintergrund der Parameter $\eta^2 = .14$, $\alpha = .05$ und $\beta = .20$ ein Umfang von $n = 18$ Personen pro Gruppe angestrebt.

II.4.1.4 Kontrolle von allgemeinen und speziellen Stör- sowie von Materialeffekten

Siehe Experiment II.

II.4.1.5 Ethische Gesichtspunkte

Siehe Experiment II. Auch eine Durchführung der Untersuchung mit älteren Gesunden wurde ohne Einschränkungen durch das Reviewboard genehmigt. Die TeilnehmerInnen unterzeichneten vor Untersuchungsbeginn einen Vertrag und erhielten nach Abschluss eine Zusammenfassung des Studienhintergrundes (siehe Anhänge D.1 und D.2). Als Aufwandsentschädigung wurde ihnen pro Sitzung ein Betrag von 10 CAD bezahlt.

II.4.2 Versuchsdurchführung

II.4.2.1 Stichprobe

Die beiden Stichproben älterer Personen wurden v.a. über eine bereits im psychologischen Department der University of Alberta existierende Liste rekrutiert, die Namen und Kontaktdaten von potentiellen InteressentInnen enthielt. Darüber hinaus wurden TeilnehmerInnen über Aushänge, Anzeigen u.ä. gewonnen. Insgesamt konnten so 34 ältere TeilnehmerInnen rekrutiert werden. Drei TeilnehmerInnen durchliefen das Paradigma zunächst probenhalber innerhalb von zwei Sitzungen im Abstand von zwei Tagen; ihre Ergebnisse blieben bei der Analyse unberücksichtigt. Eine weitere Teilnehmerin konnte den Termin zur zweiten Sitzung nicht wahrnehmen. Zwei TeilnehmerInnen erhielten versuchsweise die auch bei einer Gruppe jüngerer TeilnehmerInnen eingesetzte Instruktion zur Quellenprüfung, so dass ihre Daten nicht in die berichtete Auswertung einfließen. Alle weiteren regulären 28 Datensätze (14 pro Gruppe) gingen in die Analyse ein und werden unter dem Abschnitt Ergebnisse berichtet. Bzgl. der Rekrutierung der jüngeren TeilnehmerInnen vgl. Abschnitt I.6.1.3.

Die Stichprobe der jüngeren TeilnehmerInnen umfasste 5 männliche und 9 weibliche Studierende im Alter von 16 bis 20 Jahren ($M=18.93$, $SD=.73$; siehe Abschnitt I.6.3.1). Die beiden Stichproben älterer TeilnehmerInnen umfassten insgesamt 10 männliche und 18 weibliche TeilnehmerInnen; die Verteilung des Geschlechts war zwischen allen Gruppen vergleichbar. Das Alter der beiden Gruppen älterer TeilnehmerInnen reichte von 61 bis 85 Jahren ($M=75.00$, $SD=6.01$). Es existierte kein signifikanter Altersunterschied zwischen den beiden Gruppen älterer TeilnehmerInnen; $t(26)=.685$, $p>.05$.

II.4.2.2 Material, Hilfsmittel und Geräte

Der experimentelle Teil der Untersuchung wurde wie in Abschnitt I.6.3.4 beschrieben sowie mit denselben Hilfsmitteln und Geräten durchgeführt. Nach Abwägung der Vor- und Nachteile wurde auch dasselbe Videomaterial genutzt, wie in Studie II, da es kaum möglich erschien, vollkommen identische Handlungsausführungen zweier (alters)verschiedener Akteure aufzuzeichnen. Der Akteur war mit einem Alter von fast 40 Jahren zwar dem durchschnittlichen Alter der jüngeren TeilnehmerInnen näher, lag aber immerhin zwischen beiden Gruppen. Die Befragung einiger älterer TeilnehmerInnen zeigte zudem, dass diese dem Alter des Akteurs keine besondere Aufmerksamkeit schenkten. Zusätzlich zum experimentellen Paradigma wurden neuropsychologische Testverfahren eingesetzt, die im Folgenden kurz beschrieben werden sollen.

Die Testbatterie orientierte sich im Wesentlichen an der Zusammenstellung von Glisky, Polster und Routhieaux (1995). Folgende Testverfahren können als Maße für sog. frontale Gehirnfunktionen gelten und luden in der Untersuchung von Glisky et al. (1995) auf einem Faktor:

- Controlled Oral Word Association (FAS) (Benton & Hamsher, 1976)
- Wechsler Memory Scale (WMS) – digit spans backward (Wechsler, 1987)
- Wechsler Memory Scale (WMS) – mental control (Wechsler, 1987)
- Wechsler Adult Intelligence Scale (WAIS) – mental arithmetic (Wechsler, 1981)
- Wisconsin Card Sorting Test (WCST) (Hart, Kwentus, Wade & Taylor, 1988)

Da die einzelnen Tests unterschiedliche Aspekte frontaler Funktionen erfassen (Inhibition, Kontrolle, Flexibilität usw.) und gerade der Zusammenhang dieser verschiedenen Funktionen mit dem Ausmaß der Gedächtnistäuschung erfasst werden sollte, wurden alle genannten Verfahren eingesetzt. Darüber hinaus sollten auch Maße für sog. medial-temporale Hirnfunktionen einbezogen werden. Aus Gründen der Testökonomie konnte jedoch nur ein Verfahren ausgewählt werden:

- Wechsler Memory Scale (WMS) – logical memory (Wechsler, 1987).

Die Testbatterie von Glisky et al. umfasst zusätzlich folgende Testverfahren zur Erfassung sog. medial-temporalen Funktionen, die aus Kapazitätsgründen nicht berücksichtigt werden konnten:

- Wechsler Memory Scale (WMS) – verbal paired associates I und II (Wechsler, 1987)
- California Verbal Learning Test (CVLT) – long-delay cued recall (Delis, Kramer, Kaplan & Ober, 1987).

Zusätzlich wurde als weithin akzeptiertes Screening-Verfahren zur Bestimmung des allgemeinen kognitiven Zustandes, insb. zur Bestimmung des Schweregrades einer Demenz, der Mini-Mental-State-Test (MMSE; Folstein, Folstein & McHugh, 1975) eingesetzt. Zur Erfassung de-

pressiver Tendenzen diene die Depressionsskala des Center for Epidemiologic Studies (CES-D; Radloff, 1977). Explorativ wurden zudem noch der Test für kognitives Schätzen (TKS; Brand, Kalbe, Fujiwara, Huber & Markowitsch, 2003) und eine Zahlenkodieraufgabe aus dem DemTect (Kalbe et al., 2004) verwendet. Neben dem Zahlennachsprechen rückwärts, das Teil der Testbatterie zur Erfassung frontaler Funktionen war, wurde auch das Zahlennachsprechen vorwärts durchgeführt.

Im Folgenden werden die eingesetzten Testverfahren kurz beschrieben:

Controlled Oral Word Association (FAS) (Benton & Hamsher, 1976): Aufgabe ist es, zu drei vorgegebenen Anfangsbuchstaben (F, A, S) innerhalb von einer Minute jeweils so viele Wörter wie möglich zu generieren. Es kann jede beliebige Wortart verwendet werden, ausgenommen sind jedoch Wörter, die mit einem Großbuchstaben beginnen (, was im Englischen im Wesentlichen auf Eigennamen zutrifft). Des Weiteren darf derselbe Wortstamm nicht mehrfach gebraucht werden. Daneben wird das Generieren von Wörtern zu einer Kategorie (in diesem Falle „Tiere“) erbeten. Gezählt wird die Anzahl an gültigen Wörtern.

Wechsler Memory Scale (WMS) – digit spans forward/ backward (Wechsler, 1987): Es werden zufällige Zahlenreihen, die aus einstelligen Ziffern bestehen, vorgegeben. Diese sollen zunächst in der vorgelesenen Reihenfolge wiedergegeben werden. Dabei werden die Reihen mit jedem Durchgang um eine Zahl länger, es sei denn, den ProbandInnen unterlaufen Fehler. In diesem Fall wird erneut eine Zahlenreihe gleicher Länge vorgegeben. Die Aufgabe wird abgebrochen, wenn die TeilnehmerInnen zwei Zahlenreihen gleicher Länge nicht wiedergeben können. Ist dieser Teil abgeschlossen, werden die TeilnehmerInnen gebeten, die nächsten Zahlenreihen, die vorgelesen werden, in umgekehrter Reihenfolge, also rückwärts, zu wiederholen. Hierbei besteht dasselbe Vorgehen und Abbruchkriterium. Alle Zahlen der korrekt wiedergegebenen Reihen werden aufaddiert und als Maß verwendet.

Wechsler Memory Scale (WMS) – mental control (Wechsler, 1987): Diese Aufgabe erfordert Schnelligkeit und Flexibilität: TeilnehmerInnen werden zunächst recht einfache Aufgaben gegeben, die so schnell und korrekt wie möglich erledigt werden sollen, z.B. die Wiedergabe der Monate des Jahres in der gewöhnlichen Abfolge, d.h. angefangen mit Januar. Während der zweiten Hälfte des Tests besteht die Aufgabe dann in der Wiedergabe der Inhalte in der umgekehrten Reihenfolge: Z.B. sollen die Monate des Jahres in der umgekehrten Reihenfolge genannt werden, d.h. angefangen mit Dezember. Das Punktesystem berücksichtigt die benötigte Zeit und die Fehleranzahl.

Wechsler Adult Intelligence Scale (WAIS) – mental arithmetic (Wechsler, 1981): Hierbei handelt es sich um eine Sammlung von mathematischen Textaufgaben, die im Kopf zu lösen sind

und in der Schwierigkeit ansteigen. Auch hier werden bei der Auswertung Zeit und Fehler berücksichtigt.

Wisconsin Card Sorting Test (WCST) (Hart, Kwentus, Wade & Taylor, 1988): Bei diesem computergestützt eingesetzten Verfahren kommen Karten zum Einsatz, deren Symbole drei Dimensionen enthalten: Anzahl, Farbe und Form. Jede Dimension umfasst dabei vier Ausprägungen (Anzahl: 1, 2, 3, 4; Farbe: blau, grün, rot, gelb; Form: Kreis, Stern, Dreieck, Quadrat). Den ProbandInnen liegen vier Karten vor, die alle Dimensionen und Ausprägungen abdecken. Sie sollen nun eine Karte, die der Computer „aufdeckt“, einer der vier offen vor sich liegenden Karten zuordnen; demnach können sie alle drei Kriterien für die Zuordnung nutzen. Die computervermittelte Rückmeldung zeigt an, ob die getroffene Zuordnung richtig oder falsch ist, d.h. mit der vom Computer gewählten übereinstimmt oder nicht. Die Aufgabe der ProbandInnen besteht also darin, herauszufinden, welche Zuordnungsdimension vom Computer akzeptiert wird. Dabei ändern sich die Zuordnungsregeln von Zeit zu Zeit, so dass die ProbandInnen flexibel reagieren müssen, um weiterhin positive Rückmeldungen zu erfahren. Zwei gängige Auswertungsmaße sind der prozentuale Anteil an Perseverationen, also das Ausmaß, in welchem Zuordnungsstrategien auch nach negativer Rückmeldung beibehalten wurden, und die Anzahl an vervollständigten Kategorien, also die Anzahl an gefundenen Zuordnungsdimensionen.

Wechsler Memory Scale (WMS) – Logical Memory (Wechsler, 1987): Bei diesem Gedächtnistest werden zwei Kurzgeschichten vorgelesen, die eine Fülle an Detailinformationen enthalten. Diese sind von den ProbandInnen frei und direkt nach der Enkodierung wiederzugeben, wobei die zweite Geschichte zweimal vorgelesen wird und auch zweimal frei wiedergegeben werden soll. Nach ca. 20 Minuten erfolgt erneut eine freie Wiedergabe sowie eine Rekognitionsaufgabe zu jedem Text. Zur Auswertung der freien Wiedergaben existiert ein standardisiertes Kodierschema. Neben den wörtlichen Detailinformationen (recall unit score) werden auch Punkte für die Wiedergabe der Textbedeutung – unabhängig davon, ob alle Details stimmten oder nicht – vergeben (thematic unit score). Als Maß für die Rekognitionsleistung dient die Anzahl korrekt rekognizierter Items.

Mini-Mental-State-Exam (MMSE); Folstein, Folstein & McHugh, 1975): Dieses Screening-Instrument kognitiver Funktionen umfasst eine ganze Bandbreite von Aufgaben, z.B. zur zeitlichen und örtlichen Orientiertheit, zum Sprachverständnis oder zur Visuokonstruktion. Für jede richtige Antwort wird ein Punkt vergeben. Maximal können 30 Punkte erreicht werden.

Center for Epidemiologic Studies Depression Scale (CES-D); Radloff, 1977): Die CES-D umfasst Aussagen zu möglichen depressiven Symptomen, die auf einer vierstufigen Skala hinsichtlich der Häufigkeit ihres Vorkommens in der vorausgegangenen Woche eingeschätzt werden sollen. Die Einschätzungen werden addiert.

Zahlenkodieraufgabe (DemTect; Kalbe et al., 2004): Aufgabe ist es, zwei Zahlen, die als Ziffern vorgegeben sind, in Worte zu überführen und niederzuschreiben, bzw. umgekehrt zwei ausgeschriebene Zahlen in Ziffern zu überführen und niederzuschreiben. Es wird die Anzahl fehlerhafter Transformationen gezählt. Da nur eine einzige Teilnehmerin bei dieser Aufgabe einen Fehler beging, werden die Daten im Folgenden nicht weiter erwähnt.

Test für kognitives Schätzen (TkS; Brand, Kalbe, Fujiwara, Huber & Markowitsch, 2003): Bei diesem Test werden TeilnehmerInnen bildliche oder verbale Stimuli vorgelegt, die hinsichtlich einer von vier untersuchten Dimension (Länge, Gewicht, Menge, Zeit) einzuschätzen sind. Die Antwort wird anhand von Normen als Über-, Unter- oder korrekte Schätzung gewertet. Die Ergebnisse wurden allerdings noch nicht ausgewertet und werden auch nicht dargestellt, da momentan noch eine kanadische Normstichprobe erhoben wird.

II.4.2.3 Ablauf

Betrag das Intervall zwischen Enkodierung und Abruf *zwei Wochen*, entsprach der Ablauf des experimentellen Teils der Untersuchung vollständig demjenigen des zweiten Experiments (vgl. Abschnitt I.6.3.4). Im Anschluss an den experimentellen Part in der zweiten Sitzung – wenn also die Untersuchung für die jungen TeilnehmerInnen bereits beendet war – wurden die älteren TeilnehmerInnen jedoch zusätzlich einer ausführlichen neuropsychologischen Testung unterzogen, die ca. 1.5 Stunden in Anspruch nahm und die beschriebenen Testverfahren umfasste. Mit den TeilnehmerInnen, die die gesamte Untersuchung in einer Sitzung absolvierten, wurden Teile der neuropsychologischen Untersuchung zwischen Verarbeitung (Phase 2) und Abruf (Phase 3) des experimentellen Paradigmas durchgeführt, um ein Intervall von ca. 20 Minuten zu gewährleisten. Einige TeilnehmerInnen empfanden diese eine Sitzung, die ca. 2.5 Stunden dauerte, als zu anstrengend. Mit zwei dieser TeilnehmerInnen wurden daher einige neuropsychologische Testverfahren in einer zusätzlich anberaumten zweiten Sitzung durchgeführt. Für drei dieser TeilnehmerInnen stellte es allerdings einen zu großen Aufwand dar, erneut an der Universität zu erscheinen; daher wurde auf die Anberaumung einer zweiten Sitzung und damit die Durchführung einiger Tests in diesen Fällen verzichtet.

Der genaue Ablauf ist der folgenden Tabelle zu entnehmen. Die Instruktionen sind im Anhang D.3 bzw. B.5 abgedruckt. Das experimentelle Paradigma lief vollständig computergestützt ab; die neuropsychologische Untersuchung basierte v.a. auf paper-pencil-Verfahren. Ausnahmen stellten der Wisconsin Card Sorting Test und der Test für kognitives Schätzen dar, die beide computergestützt durchgeführt wurden.

Tab. 36: Ablauf des vierten Experiments.

Stage	Tag	Gruppe			Tag
		jüngere TeilnehmerInnen; Intervall: zwei Wochen	ältere TeilnehmerInnen; Intervall: zwei Wochen	ältere TeilnehmerInnen; Intervall: zwanzig Minuten	
ESD	1	<i>Einverständniserklärung</i>	<i>Einverständniserklärung</i>	<i>Einverständniserklärung</i>	1
ESD		<i>Eingabe von soziodemographischen Daten am Computer</i>	<i>Fragebogen zu soziodemographischen Daten, Medikation und Gesundheitsstatus</i>	<i>Fragebogen zu soziodemographischen Daten, Medikation und Gesundheitsstatus</i>	
OIP		Enkodierung (Phase 1)	Enkodierung (Phase 1)	Enkodierung (Phase 1)	
NPT		FAS/ animals	FAS/ animals	FAS/ animals	
OIP		Verarbeitung (Phase 2)	Verarbeitung (Phase 2)	Verarbeitung (Phase 2)	
		2 Wochen Pause		keine Pause	
OIP/NPT	2	Abruf (Phase 3)	Abruf (Phase 3)	WMS: logical memory – direct recall	1
NPT			WMS: logical memory – direct recall	TkS	
NPT			TkS	WMS: digit spans – fore-/ backward	
NPT			WMS: digit spans – fore-/ backward	DemTect: Zahlenkodieraufgabe	
OIP/NPT			DemTect: Zahlenkodieraufgabe	Abruf (Phase 3)	
NPT			MMSE	WMS: logical memory – delayed recall, delayed recognition	
NPT			WAIS: mental arithmetic	MMSE	
NPT			WMS: mental control	WAIS: mental arithmetic	
NPT			WMS: logical memory – delayed recall, delayed recognition	WMS: mental control	
NPT			CES-D	CES-D	
NPT			WCST	WCST	

Anmerkungen: ESD: Einverständniserklärung/ soziodemographische Daten, **OIP: Observation-inflation-Paradigma**, NPT: Neuropsychologische Testbatterie

II.4.3 Ergebnisse

II.4.3.1 Deskriptive Statistiken

Die folgenden Tabellen beinhalten die relativen Häufigkeiten von *alt-* bzw. *ausgeführt-* Antworten, d.h. sowohl Rekognitions- als auch Quellengedächtnisurteile. Die statistische Auswertung entspricht derjenigen des ersten Experiments. Anmerkungen zu den eingesetzten Verfahren sind daher dem entsprechenden Abschnitt (I.5.3.1) zu entnehmen.

Tab. 37: Relative Häufigkeiten von *alt-*Antworten im Rekognitionstest des vierten Experiments.

		Altersgruppe:					
		jüngere TeilnehmerInnen		ältere TeilnehmerInnen			
		Intervall zwischen Enkodierung (Phase 1) und Abfrage (Phase 3):					
		zwei Wochen		zwei Wochen		zwanzig Minuten	
		Art der Enkodierung in Phase 1: Ausgeführt					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.97	.07	.94	.17	1.00	.00
	1x	.97	.07	.89	.15	.99	.05
	0x	.91	.15	.69	.23	1.00	.00
		Art der Enkodierung in Phase 1: Gelesen					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.94	.09	.89	.13	.99	.05
	1x	.84	.26	.70	.20	.97	.07
	0x	.50	.22	.36	.24	.77	.19
		Art der Enkodierung in Phase 1: Nicht enkodiert					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	0x	.06	.06	.11	.09	.00	.01

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *alt-*Antworten relativiert an der Anzahl aller Handlungsanweisungen in der entsprechenden Zelle. *N*=42, d.h. 14 pro Gruppe.

Tab. 38: Relative Häufigkeiten von *ausgeführt*-Antworten im Quellengedächtnistest des vierten Experiments.

		Altersgruppe:					
		jüngere TeilnehmerInnen		ältere TeilnehmerInnen			
		Intervall zwischen Enkodierung (Phase 1) und Abfrage (Phase 3):					
		zwei Wochen		zwei Wochen		zwanzig Minuten	
		Art der Enkodierung in Phase 1: Ausgeführt					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.87	.17	.76	.27	.97	.07
	1x	.87	.17	.74	.21	.99	.05
	0x	.74	.21	.34	.23	.99	.05
		Art der Enkodierung in Phase 1: Gelesen					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	5x	.29	.28	.44	.28	.10	.17
	1x	.23	.17	.24	.27	.01	.05
	0x	.06	.09	.13	.15	.01	.05
		Art der Enkodierung in Phase 1: Nicht enkodiert					
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequenz der Präsentation in Phase 2:	0x	.02	.02	.03	.04	.00	.00

Anmerkungen: Die dargestellten Werte entsprechen der Anzahl an *ausgeführt*-Antworten relativiert an der Anzahl aller Handlungsanweisungen in der entsprechenden Zelle. $N=42$, d.h. 14 pro Gruppe.

II.4.3.2 Manipulation-Checks

Wie bereits für die Gruppe jüngerer TeilnehmerInnen gezeigt, ergab auch die Betrachtung der beiden Gruppen älterer TeilnehmerInnen, dass ausgeführte Handlungsanweisungen besser rekonstruiert wurden als nur gelesene Handlungsanweisungen (enactment-Effekt; nach zweiwöchigem Intervall: Rekonstruktion ausgeführter vs. gelesener Handlungsanweisungen ($M(SD)$): .84 (.13) vs. .65 (.13); $t(13)=6.12$, $p<.05$, $\eta^2=.742$; nach zwanzigminütigem Intervall: Rekonstruktion ausgeführter vs. gelesener Handlungsanweisungen ($M(SD)$): 1.00 (.02) vs. .91 (.08); $t(13)=3.63$, $p<.05$, $\eta^2=.503$). Dabei war der enactment-Effekt der älteren TeilnehmerInnen bei zweiwöchigem Intervall vergleichbar mit demjenigen der jüngeren TeilnehmerInnen bei gleichem Intervall; $t(26)=.00$, $p>.05$. Die älteren TeilnehmerInnen, die bereits nach zwanzig Minuten den Erinnerungstest durchliefen, wiesen (wie zu erwarten) einen weniger stark ausgeprägten enactment-Effekt auf als diejenigen, die den Erinnerungstest erst nach zwei Wochen absolvierten; $t(26)=2.68$, $p<.05$, $\eta^2=.216$. Bei der Videobeobachtung wurden – so zeigten die a posteriori erhobenen Ratings – die Instruktionen in allen drei Gruppen in vergleichbarem Ausmaß befolgt (durchschnittliche Einschätzung über alle drei Gruppen hinweg: $M=6.06$ ($SD=1.16$); $F(2,32)=3.28$, $p>.05$).

II.4.3.3 Hypothesenprüfung

Die Auswertung konzentrierte sich zunächst auf das Vorhandensein eines observation-inflation-Effekts in den beiden Gruppen älterer TeilnehmerInnen; für die Stichprobe der jüngeren TeilnehmerInnen war dieser bereits nachgewiesen worden (siehe Abschnitt I.6.4.3).

In beiden Gruppen älterer TeilnehmerInnen konnte ein erhöhtes Ausmaß falscher *ausgeführt*-Antworten nach fünfmaliger verglichen mit nullmaliger Videopräsentation nachgewiesen werden (zweiwöchiges Intervall: $t(13)=3.67$, $p<.05$, $\eta^2=.508$; zwanzigminütiges Intervall: $t(13)=2.12$, $p<.05$, $\eta^2=.257$).

In einem nächsten Schritt wurde das Ausmaß der inflation-Effekte zwischen den Gruppen verglichen. Einerseits sollten – bei gleichem Intervall von zwei Wochen zwischen Enkodierung und Abfrage – die älteren mit den jüngeren TeilnehmerInnen verglichen werden; andererseits sollten – bei gleicher Altersgruppe – die zwei unterschiedlichen Zeitintervalle verglichen werden.

Es ergab sich keine signifikante Interaktion zwischen der Präsentationsfrequenz in Phase 2 (5x vs. 0x) und der Altersgruppe (jüngere, ältere TeilnehmerInnen); $t(26)=.83$, $p>.05$. Demnach

unterschied sich bei konstantem Intervall das Ausmaß des observation-inflation-Effekts nicht zwischen den untersuchten Altersstufen. Ein Vergleich der beiden Gruppen älterer TeilnehmerInnen zeigte jedoch, dass der inflation-Effekt nach längerem Intervall deutlich stärker ausgeprägt war als nach kürzerem Intervall; $t(26)=2.41, p<.05, \eta^2=.138$.

Erneut konnten die Materialpools weder innerhalb noch zwischen den Gruppen gleich häufig eingesetzt werden. Um sicherzustellen, dass es sich bei den Effekten nicht um Materialeffekte handelte, wurde daher der Einfluss der verwendeten Materialkombinationen überprüft. Die Analyse basiert auf allen drei Gruppen. Es zeigten sich keine signifikanten Effekte der Materialkombination auf falsche *ausgeführt*-Antworten nach fünfmaliger ($F(8,33)=1.20, p>.05$) bzw. nullmaliger ($F(8,33)=1.61, p>.05$) Präsentation.

II.4.3.4 Weitere Befunde

Bevor die falschen Erinnerungen an die Handlungsausführung einer genaueren Betrachtung unterzogen werden, sollen zunächst die korrekten Erinnerungsleistungen hinsichtlich potentieller Alters- bzw. Intervalleffekte analysiert werden.

Dabei zeigte sich, dass ältere im Vergleich zu jüngeren TeilnehmerInnen insgesamt weniger ausgeführte Handlungsanweisungen rekognizierten; $t(18.65)=3.03, p<.05, \eta^2=.261$. Bei näherer Betrachtung, d.h. unter Einbezug der verschiedenen Präsentationsfrequenzen, ging dieser Befund darauf zurück, dass die Handlungsanweisungen, die in Phase 2 nicht wieder präsentiert wurden (0x), von den älteren TeilnehmerInnen deutlich schlechter wiedererkannt wurden, während die erneut (5x vs. 1x) präsentierten gleich gut rekogniziert wurden (5x: $t(26)=.59, p>.05$; 1x: $t(18.70)=1.91, p>.05$; 0x: $t(26)=3.09, p<.05, \eta^2=.269$). Auch korrekte Quellengedächtnisurteile bzgl. ausgeführter Handlungen waren bei den jüngeren TeilnehmerInnen häufiger als bei den älteren; $t(26)=3.69, p<.05, \eta^2=.344$. Dies zeigte sich jedoch ebenfalls nicht nach fünf- und ein-, sondern nur nach nullmaliger Präsentation in Phase 2 (5x: $t(21.60)=1.33, p>.05$; 1x: $t(26)=1.77, p>.05$; 0x: $t(26)=4.79, p<.05, \eta^2=.469$). Es ergab sich kein signifikanter Unterschied bzgl. der Rekognitionsquote der in Phase 1 nur gelesenen Handlungsanweisungen zwischen beiden Altersgruppen ($t(26)=1.98, p>.05$) – die differentielle Betrachtung zeigte, dass dies auch innerhalb der verschiedenen Frequenzstufen galt (5x: $t(26)=1.34, p>.05$; 1x: $t(26)=1.61, p>.05$; 0x: $t(26)=.89, p<.05$).

Besonders auffällig an dem Muster der älteren TeilnehmerInnen, die die dritte Phase nach langem Intervall durchliefen, ist die (im Vergleich zu den jungen TeilnehmerInnen) enorme Differenz der Quoten korrekter *ausgeführt*-Antworten auf Handlungsanweisungen, die in der

zweiten Phase erneut präsentiert wurden, und solchen, die nicht erneut präsentiert wurden – es zeigte sich hier also ganz deutlich auch ein observation-inflation-Effekt für korrekte *ausgeführt*-Urteile. Während sich innerhalb der Gruppe der jungen TeilnehmerInnen keine signifikanten Unterschiede zwischen diesen Frequenzstufen ergaben, zeigten sich deutliche Unterschiede zwischen ein- und nullmaliger Präsentation innerhalb der Gruppe der älteren TeilnehmerInnen (ältere TeilnehmerInnen: 5x-1x: $t(13)=.16$, $p>.05$; 1x-0x: $t(13)=6.36$, $p<.05$; $\eta^2=.757$; 5x-0x: $t(13)=5.20$, $p<.05$; $\eta^2=.675$; jüngere TeilnehmerInnen: 5x-1x: $t(13)=.00$, $p>.05$; 1x-0x: $t(13)=1.88$, $p>.05$; 5x-0x: $t(13)=2.22$, $p<.05$; $\eta^2=.276$). Die älteren TeilnehmerInnen gaben also signifikant weniger korrekte *ausgeführt*-Antworten auf Handlungsanweisungen, die in der zweiten Phase nicht erneut präsentiert wurden, verglichen mit solchen, die in der zweiten Phase (ein- oder fünfmal) präsentiert wurden. Die jüngeren TeilnehmerInnen hingegen zeigten einen signifikanten Unterschied nur im Vergleich der Frequenzen 5x und 0x (siehe Abschnitt I.6.4.3). Dies warf die Frage auf, ob der observation-inflation-Effekt für korrekte Erinnerungen an die Handlungsausführung, d.h. das Ausmaß korrekter *ausgeführt*-Antworten nach fünf- verglichen mit nullmaliger Präsentation, bei den älteren auch signifikant stärker ausgeprägt sein könnte als bei den jüngeren TeilnehmerInnen. Dies war sehr deutlich der Fall: $t(13)=2.90$, $p<.05$, $\eta^2=.244$. Weiterhin zeigen die deskriptiven Statistiken, dass der observation-inflation-Effekt für korrekte *ausgeführt*-Urteile bei den älteren TeilnehmerInnen sogar stärker ausgeprägt war als derjenige für falsche *ausgeführt*-Urteile. Bei den jüngeren TeilnehmerInnen fand sich das umgekehrte Verhältnis. Allerdings konnte weder der Vergleich innerhalb noch zwischen den Gruppen statistisch abgesichert werden (ältere TeilnehmerInnen: $t(13)=.83$, $p>.05$; jüngere TeilnehmerInnen: $t(13)=1.05$, $p>.05$; ältere vs. jüngere TeilnehmerInnen: $F(1,26)=1.70$, $p>.05$). Ein ähnliches Muster (größere Differenz zwischen fünf- bzw. einmalig und nullmalig präsentierten Handlungsanweisungen bei den älteren als bei den jüngeren TeilnehmerInnen) findet sich – allerdings etwas schwächer ausgeprägt – auch für die *Rekognition ausgeführter Handlungsanweisungen* (ältere TeilnehmerInnen: 5x-1x: $t(13)=.89$, $p>.05$; 1x-0x: $t(13)=2.65$, $p<.05$, $\eta^2=.350$; 5x-0x: $t(13)=5.83$, $p<.05$; $\eta^2=.723$; jüngere TeilnehmerInnen: 5x-1x: $t(13)=.00$, $p>.05$; 1x-0x: $t(13)=1.47$, $p>.05$; 5x-0x: $t(13)=1.17$, $p>.05$). Insgesamt ähneln sich die Ergebnismuster beider Altersgruppen bzgl. *nicht ausgeführter Handlungsanweisungen* (sowohl im Rekognitions- als auch im Quellengedächtnistest) stärker als bzgl. ausgeführter Handlungsanweisungen.

Auf einen differenzierten Vergleich der Ergebnisse der älteren TeilnehmerInnen für die beiden Abrufintervalle soll an dieser Stelle verzichtet werden. Die deskriptivstatistischen Ergebnisse zeigen anschaulich genug, dass nach kurzem Intervall deutliche Deckeneffekte auftraten, so dass auch alle möglichen Vergleiche signifikant werden, abgesehen von einem: Die Rekognitionsquote von in Phase 1 ausgeführten Handlungen, die in Phase 2 fünfmalig beobachtet wur-

den, ist in beiden Gruppen vergleichbar; $t(13.00)=1.30$, $p>.05$. Die Ergebnisse zeigen jedoch auch sehr deutlich, dass – wenn überhaupt Fehler innerhalb dieser Gruppe mit kurzem Intervall begangen wurden – diese in der mangelnden Rekognition von in Phase 2 nicht erneut präsentierten Items bzw. in falschen Quellenzuordnungen nicht ausgeführter Handlungsanweisungen, die in Phase 2 mehrfach beobachtet wurden (d.h. im observation-inflation-Effekt), bestanden.

Bzgl. der falschen Erinnerungen wurden – wie schon für die jüngeren TeilnehmerInnen – auch für beide Gruppen älterer TeilnehmerInnen Interaktionen zwischen der Frequenz der Präsentation in Phase 2 (5x vs. 0x) und der Art des Gedächtnistests in Phase 3 (Rekognition, Quellengedächtnis) berechnet. Diese Interaktionen wurden – wie schon für die jüngeren TeilnehmerInnen gefunden – durchweg signifikant (ältere TeilnehmerInnen, zweiwöchiges Intervall: $t(13)=2.21$, $p<.05$, $\eta^2=.272$; ältere TeilnehmerInnen, zwanzigminütiges Intervall: $t(13)=2.22$, $p<.05$, $\eta^2=.276$). Demnach ist der Anstieg der falschen *ausgeführt*-Antworten auch in den beiden Stichproben älterer TeilnehmerInnen disproportional zum Anstieg der *alt*-Antworten auf die entsprechenden Items.

Die Betrachtung auf Personenebene ergab folgendes Bild: Von 14 Personen in jeder der beiden Gruppen älterer TeilnehmerInnen wiesen 4 nach zwanzigminütigem Intervall und 10 nach zweiwöchigem Intervall einen inflation-Effekt auf. In der Gruppe jüngerer TeilnehmerInnen waren es 9 von 14. Dabei weicht die Verteilung der älteren TeilnehmerInnen, die das lange Intervall durchliefen, nicht signifikant von derjenigen der jüngeren TeilnehmerInnen mit gleichem Intervall ab ($\chi^2(1,28)=.16$, $p>.05$); die Verteilung der beiden Gruppen älterer TeilnehmerInnen hingegen war signifikant unterschiedlich ($\chi^2(1,28)=5.14$, $p<.05$, $\Phi^2=.429$).

Im Folgenden sind die Sicherheitsratings in den einzelnen Bedingungen dargestellt.

Tab. 39: Sicherheitsratings (siebenstufige Skala) bzgl. falscher *ausgeführt*-Antworten im vierten Experiment.

	Altersgruppe								
	jüngere TeilnehmerInnen			ältere TeilnehmerInnen					
	Intervall zwischen Enkodierung und Abfrage								
	zwei Wochen			zwei Wochen			zwanzig Minuten		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
5x	20	5.25	1.59	31	5.19	1.64	7	5.86	1.46
1x	16	5.50	1.37	17	4.88	1.73	1	5.00	-
0x	4	4.00	1.54	9	5.11	1.05	1	5.00	-
∅	40	5.23	1.49	57	5.09	1.57	9	5.67	1.33
	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>	<i>N</i>	<i>f_%(7)</i>	<i>f(7)</i>
5x	20	20%	4	31	32%	10	7	57%	4

Anmerkungen: *N*=Anzahl an Fehlern; *f_%(7)*=Prozentuale Häufigkeit der höchsten Ratingstufe (7).

Demnach sind die älteren TeilnehmerInnen, die den Erinnerungstest nach nur zwanzigminütigem Intervall absolvierten, in ihren falschen Erinnerungsurteilen bzgl. der Handlungsausführung sicherer als die älteren und jüngeren TeilnehmerInnen, die den Test nach zweiwöchigem Intervall durchliefen. Ein Einfluss der Frequenz auf die Sicherheit der Erinnerungsurteile lässt sich deutlich für die jüngeren TeilnehmerInnen ablesen, aber auch bei den älteren TeilnehmerInnen zeigt sich eine solche Tendenz. Hierbei ist der Zuwachs an Sicherheit nach fünfmaliger Präsentation im Vergleich zu nullmaliger Präsentation für die Gruppe mit kurzem Intervall deutlicher als für die Gruppe mit langem Intervall; allerdings ist das Ergebnismuster der erstgenannten Gruppe diesbzgl. mit Vorsicht zu interpretieren, da jeweils lediglich eine Handlungsanweisung nach ein- und nullmaliger Präsentation fehlattribuiert wurde. Nach zweiwöchigem Intervall ist der Unterschied zwischen fünf- und nullmaliger Präsentation gering; die niedrigste Sicherheit wird nach einmaliger Präsentation erzielt. Während nach fünfmaliger Präsentation die Sicherheitsurteile älterer und jüngerer TeilnehmerInnen mit langem Intervall vergleichbar ausfallen, weisen die jüngeren TeilnehmerInnen nach einmaliger und die älteren nach nullmaliger Präsentation eine im Vergleich höhere Sicherheit auf.

Ein Blick auf die fälschlich als ausgeführt erinnerten Handlungsanweisungen ergab dasselbe Bild, das schon im Rahmen der zweiten Studie geschildert worden war: eine relativ ungleichmäßige Verteilung, allerdings bei ungleich häufig eingesetzten Kombinationen. Die beiden am häufigsten fälschlicherweise als ausgeführt erinnerten Handlungsanweisungen der älteren TeilnehmerInnen waren „Nehmen Sie einen Becher vom Stapel!“, „Biegen Sie eine Büroklammer auseinander!“ und „Wickeln Sie das Band auf!“ mit jeweils vier Fehlattributionen. Acht der insgesamt 45 Handlungsanweisungen wurden gar nicht falsch erinnert. Auffällig war allein,

dass von den insgesamt nur neun fehlerinnerten Handlungsanweisungen in der Gruppe mit zwanzigminütigem Intervall jeweils zweimal die Handlungsanweisungen „Falten Sie das Handtuch!“ und „Nehmen Sie den Deckel von der Schachtel!“ fehlerinnert wurden. Auch unter den niemals fälschlich erinnerten Handlungen fanden sich zwei bekannte: Niemand erinnerte sich fälschlicherweise, die Karten gemischt oder den Bleistift angespitzt zu haben. Unter den weiteren sechs Handlungsanweisungen, die niemals falsch erinnert wurden, fanden sich jedoch auch solche, die von den jungen TeilnehmerInnen – z.T. sogar sehr häufig – fehlattribuiert wurden.

Während für die jüngeren TeilnehmerInnen nachgewiesen werden konnte, dass schon das einmalige Beobachten der Videos zu einem Anstieg falscher *ausgeführt*-Antworten führte und sich dieses Ausmaß durch die fünfmalige Präsentation nicht signifikant steigern ließ, zeigte sich Umgekehrtes für die älteren TeilnehmerInnen nach zweiwöchigem Intervall. In dieser Gruppe unterschied sich das Ausmaß falscher *ausgeführt*-Antworten nicht nach einmaliger im Vergleich zu nullmaliger, sondern nur bei einmaliger im Vergleich zu fünfmaliger Präsentation. Innerhalb der Gruppe, bei der nur 20 Minuten zwischen Enkodierung und Abruf vergingen, zeigte sich hingegen kein Unterschied zwischen nullmaliger und einmaliger sowie einmaliger und fünfmaliger Präsentation (1x vs. 0x: zweiwöchiges Intervall: $t(13)=1.38$, $p>.05$, zwanzigminütiges Intervall: $t(13)=.00$, $p>.05$; 5x vs. 1x: zweiwöchiges Intervall: $t(13)=2.31$, $p<.05$, $\eta^2=.292$; zwanzigminütiges Intervall: $t(13)=1.71$, $p>.05$).

Alle drei Gruppen wiesen (im Sinne eines Bodeneffekts) äußerst geringe Quoten an falschen *ausgeführt*-Antworten auf neue, d.h. erst im Gedächtnistest präsentierte, Items auf; in der Gruppe älterer TeilnehmerInnen mit kurzem Intervall wurden sogar überhaupt keine Fehlattribuierungen dieser Items vorgenommen.

Nach dem Gedächtnistest in der zweiten Sitzung wurde die Einschätzung einiger Items auf einer siebenstufigen Skala erbeten (siehe Tabelle 40, folgende Seite). Dabei lag die Sicherheit, mit der eine Handlungsanweisung als in Phase 1 ausgeführt versus gelesen erinnert wurde (Item 4 vs. 5) in allen drei Gruppen für das Ausführen höher als für das Lesen. Wie auch schon für die Gruppe der Studierenden gezeigt, war dieser Unterschied jedoch auch in den Gruppen älterer TeilnehmerInnen nicht signifikant. Jüngere und ältere TeilnehmerInnen unterschieden sich dadurch, dass die jüngeren angaben, während des Gedächtnistests mehr an die Videos gedacht und eine stärkere Beeinflussung durch die Videos befürchtet zu haben (Items 7 und 8). Die beiden Gruppen älterer TeilnehmerInnen unterschieden sich demgegenüber hinsichtlich (fast) aller anderen Einschätzungen: Die Items 2 bis 6 wurden von den TeilnehmerInnen nach kürzerem Intervall signifikant höher eingeschätzt. Allerdings ist zu beachten,

dass die beiden Stichproben unterschiedlich groß sind, da die Ratings erst im Laufe der Untersuchung eingeführt wurden.

Tab. 40: Einschätzung verschiedener Items (siebenstufige Skala) im vierten Experiment.

	Altersgruppe		
	jüngere TeilnehmerInnen	ältere TeilnehmerInnen	
	Intervall zwischen Enkodierung und Abfrage		
	zwei Wochen	zwei Wochen	zwanzig Minuten
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Item 1: I expected that my memory for the action statements would be tested.	4.93 (1.98)	4.21 (2.55)	4.71 (2.75)
Item 2: I expected exactly this kind of memory test.	3.21 (1.58)	2.50 (1.91)	4.29 (2.43)
Item 3: The memory task I just fulfilled was very easy	4.57 (1.34)	3.93 (1.64) ²	5.57 (1.62) ²
Item 4: While I was doing the memory task, I was very sure about which actions I actually performed.	4.14 (1.41)	4.50 (1.09) ²	6.57 (0.79) ²
Item 5: While I was doing the memory task, I was very sure about which actions I merely read.	3.86 (1.56)	4.21 (1.25) ²	6.14 (1.97) ²
Item 6: My memory for the first phase of the study (i.e. when I performed or merely read the action statements) is very clear and vivid.	4.21 (1.31)	4.00 (1.30) ²	5.86 (1.86) ²
Item 7: I thought about the video-observation task while I was doing the memory task.	6.00 (0.78) ¹	4.29 (1.86) ¹	5.43 (2.22)
Item 8: While I was doing the memory task, I feared that my memory for action performance could be distorted by the videos I observed later.	4.79 (1.67) ¹	3.36 (1.91) ¹	2.86 (2.61)

Anmerkungen: $N=35$; $n=14$ für die Bedingungen „jüngere TeilnehmerInnen, zweiwöchiges Intervall“ und „ältere TeilnehmerInnen, zweiwöchiges Intervall“, $n=7$ für die Bedingung „ältere TeilnehmerInnen, zwanzigminütiges Intervall“. 1 bzw. 2: $p < .05$ (zweiseitig) beim Vergleich der Gruppen „jüngere vs. ältere TeilnehmerInnen, zweiwöchiges Intervall“ (1) bzw. „ältere TeilnehmerInnen, zweiwöchiges Intervall“ vs. „ältere TeilnehmerInnen, zwanzigminütiges Intervall“ (2).

In der folgenden Tabelle sind für beide Gruppen älterer TeilnehmerInnen die Korrelationen zwischen dem Ausmaß des observation-inflation-Effekts und den Leistungen in den erhobenen Testverfahren wiedergegeben. Neben der Differenz falscher *ausgeführt*-Antworten nach fünf- und nullmaliger Präsentation (observation inflation für falsche *ausgeführt*-Antworten) wurde auch die Differenz für die korrekten *ausgeführt*-Antworten (observation inflation für korrekte *ausgeführt*-Antworten) – die ja wie beschrieben auffällig hoch war – explorativ mit den erhobenen Maßen korreliert.

Tab. 41: Korrelationen von Leistungen in Testverfahren zur Erfassung sog. frontaler bzw. medial-temporalen Funktionen und dem observation-inflation-Effekt für falsche bzw. korrekte *ausgeführt*-Antworten im vierten Experiment.

		ältere TeilnehmerInnen, zweiwöchiges Intervall		ältere TeilnehmerInnen, zwanzigminütiges Intervall	
		observation inflation -falsch	observation inflation -korrekt	observation inflation -falsch	observation inflation -korrekt
		<i>r</i> (<i>p</i>)	<i>r</i> (<i>p</i>)	<i>r</i> (<i>p</i>)	<i>r</i> (<i>p</i>)
frontal	FAS	.079 (.789)	.577 (.031)	-.280 (.333)	-.337 (.239)
	animals	-.091 (.756)	.220 (.450)	-.121 (.680)	-.322 (.262)
	WMS: digit spans backward	-.452 (.105)	.445 (.111)	-.510 (.063)	-.320 (.264)
	WMS: mental control	.346 (.226)	.598 (.024)	.352 (.262, <i>n</i> =12)	.273 (.390, <i>n</i> =12)
	WAIS: mental arithmetic	.225 (.439)	.764 (.001)	.245 (.444, <i>n</i> =12)	.174 (.589, <i>n</i> =12)
	WCST:categories completed	-.061 (.835)	.328 (.253)	.404 (.218, <i>n</i> =11)	-.289 (.389, <i>n</i> =11)
	frontal factor	.076 (.797)	.677 (.008)	.135 (.693, <i>n</i> =11)	-.264 (.433, <i>n</i> =11)
	medial- temporal	WMS: logical memory – direct recall, rus	.585 (.028)	-.104 (.724)	-.265 (.360)
WMS: logical memory – direct recall, tus		.584 (.028)	-.200 (.492)	-.227 (.435)	-.202 (.488)
WMS: logical memory – delay. recall, rus		.240 (.409)	.035 (.905)	-.244 (.400)	-.491 (.075)
WMS: logical memory – delay recall, tus		.430 (.125)	-.178 (.543)	-.200 (.493)	-.335 (.182)
WMS: logical memory – delay. recognition		.332 (.246)	.101 (.732)	-.322 (.261)	-.421 (.134)
sonstiges	MMSE	.043 (.885)	.400 (.157)	-.452 (.105)	.021 (.944)
	CES-D	-.020 (.946)	.217 (.456)	.518 (.070, <i>n</i> =13)	.142 (.643, <i>n</i> =13)

Anmerkungen: Fett markiert sind die höchsten ($r > .400$), teilweise signifikanten Korrelationen. $N=28$, d.h. 14 pro Gruppe, falls nicht anderweitig angegeben. rus: recall unit score. tus: thematic unit score (siehe Abschnitt II.4.2.2). Frontal factor: Zusammenfassung aller aufgelisteten Tests zur Erfassung sog. frontaler Funktionen..

II.4.4 Diskussion

II.4.4.1 Diskussion unter methodischen Gesichtspunkten

Studie IV zeigte, dass – wie vorhergesagt – auch die älteren TeilnehmerInnen einen observation-inflation-Effekt aufwiesen. Dies galt sowohl für ein langes, als auch für ein kurzes Intervall, wobei der Effekt nach kurzem Intervall kleiner ausfiel. Entgegen der aufgestellten Hypothese zeigte sich bei langem Intervall jedoch kein Alterseffekt in Bezug auf die falschen Quellenzuordnungen; die älteren TeilnehmerInnen unterlagen dem Effekt also nicht signifikant stärker als die jüngeren.

Die theoretischen und praktischen Implikationen dieses Befundes sollen im nächsten Abschnitt erörtert werden. Zunächst sollen jedoch verschiedene Aspekte der Validität der Hypothesenprüfung beleuchtet werden.

Die Hypothesen wurden adäquat operationalisiert, in Testhypothesen überführt und statistisch valide überprüft. Die Testhypothesen können zum Großteil als angenommen gelten, da die erwarteten Effekte trotz der um vier Personen pro Gruppe verfehlten optimalen Stichprobengröße nicht nur signifikant ($p < .05$), sondern auch substantiell ($\eta^2 \geq .14$) waren. Die Hypothese 3, die sich auf den Vergleich beider Altersgruppen hinsichtlich des observation-inflation-Effekts bezog, konnte statistisch jedoch nicht nachgewiesen werden. Deskriptivstatistisch deutet das Ergebnismuster relativ klar in die erwartete Richtung, daher ist nicht auszuschließen, dass sich der Alterseffekt aufgrund der zu geringen Stichprobengröße nicht statistisch absichern ließ. Ein Hinweis hierauf stellt die im Nachhinein berechnete Teststärke dar, die mit knapp .19 recht niedrig lag.

Es wurde eine möglichst strenge, valide und präzise Hypothesenprüfung angestrebt. Potentiell problematische Aspekte sollen im Folgenden beleuchtet werden.

Zunächst zeigten Ratings, dass sich alle TeilnehmerInnen auch in der zweiten Phase, in der sie im Gegensatz zu Phase 1 nicht durchgehend kontrolliert bzw. beobachtet wurden, stark bemühten, die Instruktionen wie gewünscht zu befolgen. Demnach können die Ergebnisse nicht auf eine mangelnde compliance während der Untersuchungsdurchführung zurückgeführt werden. Auch die Möglichkeit, dass es sich bei den gefundenen Effekten um bloße Materialeffekte handeln könnte, konnte zurückgewiesen werden. Des Weiteren konnte auch für die älteren

TeilnehmerInnen eindeutig gezeigt werden, dass der Anstieg der falschen *ausgeführt*-Antworten kein bloßes Abbild des Anstieges an *a/t*-Antworten ist, der Effekt also mit anderen Worten kein methodologisches Artefakt darstellt. Die auch schon für Studie I beschriebenen Maßnahmen zur Kontrolle von allgemeinen und speziellen Störeffekten stellten die interne Validität sicher; auf einige potentielle Einschränkungen wird im Folgenden eingegangen.

Problematisch für die Interpretation stellt sich die Tatsache dar, dass der Gedächtnistest nach zwanzigminütigem Intervall zu leicht zu fallen schien. Die älteren TeilnehmerInnen produzierten hier fast ausschließlich Decken- bzw. Bodeneffekte, es entstand kaum Varianz. Dies spiegelt sich auch in den a posteriori erhobenen Ratings wider. Die TeilnehmerInnen schätzten nach kurzem Intervall die Qualität ihrer Erinnerungen und die Sicherheit ihrer Urteile im Vergleich zu der Gruppe mit langem Intervall deutlich höher, ja nahezu perfekt, ein. Diese an sich unerwünschten Decken- und Bodeneffekte wurden jedoch in Kauf genommen, da diese Gruppe u.a. als Kontrollgruppe für eine Stichprobe leicht dementer TeilnehmerInnen dienen sollte, deren Einschluss nur bei einem vergleichsweise kurzen Intervall sinnvoll erscheint. Statistisch wurde das Problem dadurch gelöst, dass die Ergebnisse auch anhand von nicht-parametrischen Verfahren abgesichert wurden.

Zwei weitere Aspekte könnten möglicherweise die Interpretation der Ergebnisse beeinflussen und sollen im Folgenden diskutiert werden.

Einerseits wurden zwar signifikante observation-inflation-Effekte in beiden Gruppen älterer TeilnehmerInnen gefunden. Allerdings bleibt unklar, ob diese wirklich spezifisch auf die Beobachtung zurückzuführen sind und nicht bspw. auf die durch das wiederholte Lesen erzeugte Vertrautheit mit den Handlungsanweisungen. Mit anderen Worten fehlt eine Kontrollgruppe, die die Handlungsanweisungen in der zweiten Phase nur las, so wie sie für die jüngeren TeilnehmerInnen eingeführt wurde. Diese konnte aufgrund mangelnder Verfügbarkeit älterer TeilnehmerInnen nicht realisiert werden. Allerdings wird aus mehreren Gründen davon ausgegangen, dass der Effekt tatsächlich spezifisch für die Beobachtung ist und nicht auf bloßer Vertrautheit mit den Handlungsanweisungen beruht. Diese Gründe sollen jedoch im Rahmen der theoretischen Diskussion erörtert werden.

Andererseits wurden zwar die jüngeren und die älteren TeilnehmerInnen innerhalb der jeweiligen Altersgruppen zufällig den verschiedenen Untersuchungsgruppen zugeteilt (Variation des Intervalls zwischen Enkodierung und Abruf bei den älteren TeilnehmerInnen; Variation der Art der Verarbeitung und der Art der Abrufmanipulation bei den jüngeren TeilnehmerInnen, siehe Experiment II); das Alter an sich ist jedoch eine organismische Variable und muss daher als

quasi-experimentell aufgefasst werden. Problematisch wäre diese Eigenschaft allerdings nur, wenn Konfundierungen, z.B. in Form von Kohorteneffekten, angenommen werden müssten. Im vorliegenden Fall waren die Altersgruppen jedoch sorgfältig gewählt und hinsichtlich potentiell relevanter Dimensionen vergleichbar. So waren nicht nur für die Studierenden, sondern auch für die älteren TeilnehmerInnen Universitätsumgebung und Studienteilnahme nicht ungewöhnlich, da diese zum Großteil mit der Universität affiliert waren und schon häufiger an derartigen Untersuchungen teilgenommen hatten. Daher unterschieden sich auch die Ratings beider verglichener Gruppen jüngerer und älterer TeilnehmerInnen hinsichtlich der Gewöhnlichkeit der Untersuchungsteilnahme nicht; $t(26)=.46$; $p>.05$. Viele der älteren TeilnehmerInnen hatten auch studiert, so dass die durchschnittliche Dauer der Ausbildung mit 16.57 ($SD=3.69$) Jahren recht hoch lag. Auch ergab sich kein Zusammenhang des Ausmaßes falscher Erinnerungen an die Handlungsausführung mit dieser Variable ($r(14)=-.007$, $p>.05$), so dass ein bedeutsamer Einfluss ausgeschlossen werden kann. Auch waren fast alle älteren TeilnehmerInnen mit dem Computerumgang vertraut, konnten also – genau wie die jüngeren TeilnehmerInnen – selbstständig ihre Eingaben tätigen. Da auch das eingesetzte Material für beide Gruppen mutmaßlich gleich alltäglich war, ist alles in allem von einer Vergleichbarkeit auszugehen. Dies gilt insbesondere für die falschen *ausgeführt*-Antworten, da sich die Re-kognitionsquoten dieser Handlungsanweisungen zwischen beiden Gruppen entsprechen.

Bzgl. der Überlegungen zur Konstrukt-, Populations- und Situationsvalidität kann auf das entsprechende Kapitel der ersten Studie verwiesen werden. Es ist allerdings positiv anzumerken, dass die hier untersuchten Stichproben älterer TeilnehmerInnen weniger selektiv waren, da es sich nicht nur um ehemalige Studierende diverser Fachrichtungen, sondern auch um solche TeilnehmerInnen, die kein Studium absolviert hatten, handelte. Darüber hinaus ist natürlich die Tatsache zentral, dass erstmals nicht nur jüngere, sondern eben auch ältere Personen eingeschlossen wurden. Da auch für diese ein observation-inflation-Effekt gefunden wurde, stellt dies einen Hinweis auf die Ausweitbarkeit des Befundes auf die Population älterer Erwachsener ohne besondere Einschränkungen bzgl. deren Bildungsweg dar.

Zusammengefasst kann also davon ausgegangen werden, dass die interne sowie die Ableitungsvalidität auf den verschiedenen Ebenen der Untersuchung gegeben waren; bzgl. der externen Validität bestehen auch für Experiment IV die üblichen Einschränkungen eines Laborexperiments, das zwar an einer angefallenen, jedoch nicht allzu selektiven Stichprobe durchgeführt wurde. Demnach können die Hypothesen

Ältere Personen weisen (sowohl nach langem als auch nach kurzem Intervall) zumeist einen observation-inflation-Effekt auf.

Bei verkürztem Intervall produzieren ältere Personen weniger falsche Erinnerungen an die Handlungsausführung als bei langem Intervall.

vor dem Hintergrund der beschriebenen Einschränkungen als vorläufig bewährt gelten. Die Hypothese

Bei gleichem (langem) Intervall produzieren ältere Personen zumeist mehr falsche Erinnerungen an die Handlungsausführung als jüngere Personen.

konnte hingegen nicht nachgewiesen werden.

II.4.4.2 Diskussion unter theoretischen und praktischen Gesichtspunkten

Experiment IV zeigte, dass der für jüngere TeilnehmerInnen wiederholt replizierte und hinsichtlich der Mechanismen eingegrenzte observation-inflation-Effekt auch bei älteren TeilnehmerInnen gefunden werden konnte; dies galt sogar dann, wenn die Abfrage nur wenige Minuten nach der Enkodierung stattfand.

Dennoch ist bzgl. der Gruppe mit kurzem Intervall zu konstatieren, dass der gefundene observation-inflation-Effekt lediglich auf den falschen Erinnerungen von vier TeilnehmerInnen beruhte. Der größere Teil, d.h. die übrigen zehn TeilnehmerInnen der Stichprobe, produzierten keine falschen Erinnerungen an die Handlungsausführung. Auch wenn demnach nur ein kleiner Teil der untersuchten Personen Probleme bei der Quellendiskrimination aufwies, zeigt das Ergebnismuster ganz eindeutig, dass diese genau an der vermuteten Stelle auftraten: Die Quellendiskrimination schien nur bei einer Itemart schwer zu fallen, nämlich eben bei den fünfmalig beobachteten Handlungen, die nicht ausgeführt wurden. Alle anderen Itemarten wurden ihren Quellen nahezu perfekt zugeordnet. Im Grunde ist es daher erstaunlich, dass überhaupt einige TeilnehmerInnen nach derart kurzem Intervall falsche Erinnerungen an die Handlungsausführung, die ja auch im Alter einen Erinnerungsvorteil aufweist, produzierten und hierbei sehr hohe Sicherheiten aufwiesen. Dies gilt v.a. vor dem Hintergrund, dass die jüngeren TeilnehmerInnen in der Studie von Thomas und Bulevich (2006), die den Gedächtnistest nach zwei Tagen durchliefen, keinerlei Schwierigkeiten zeigten.

Entgegen der Vorhersage konnte nicht nachgewiesen werden, dass die älteren im Vergleich zu den jüngeren TeilnehmerInnen (bei gleichem Intervall) einem stärkeren observation-inflation-Effekt unterliegen. Da das Ergebnismuster deskriptivstatistisch in die erwartete Richtung deutete, könnte dieser Befund jedoch auf die zu kleine Stichprobengröße zurückzuführen sein. Ein Vergleich der Ergebnisse von Thomas und Bulevich (2006), die eine deutlich größere Stichprobe rekrutierten ($n=22$ bzw. 27 TeilnehmerInnen pro Gruppe), mit den in Experiment IV gefundenen soll nur mit Vorsicht gezogen werden, da beide Untersuchungen in mehreren Hinsichten divergieren. Der zentrale Unterschied besteht hierbei darin, dass Thomas und Bulevich mit Vorstellung arbeiteten und das „klassische“ Paradigma nutzten, das die Einführung neuer Items in Phase 2 beinhaltet. Ein (vorsichtiger) Vergleich zeigt jedoch, dass das Ausmaß falscher Erinnerungen an die Handlungsausführung für die älteren TeilnehmerInnen vergleichbar ist, während die jüngeren TeilnehmerInnen in der vorliegenden Untersuchung ein stärkeres Ausmaß an falschen Erinnerungen aufweisen als in der Untersuchung von Thomas und Bulevich. Demnach könnte in der Tat zur Absicherung eines möglicherweise kleineren Alterseffekts hinsichtlich des Phänomens der observation inflation verglichen mit dem Phänomen der imagination inflation eine größere Stichprobe notwendig sein, als die kalkulierte (, die zudem in der vorliegenden Untersuchung noch unterschritten wurde). Allerdings fanden McDaniel et al. (2006) ebenfalls keinen signifikanten Altersunterschied bzgl. der Quellengedächtnisleistung; es existierte sogar nur eine sehr schwache Tendenz in diese Richtung ($n=16$ bzw. 18 TeilnehmerInnen pro Gruppe). Es ist vermutlich auszuschließen, dass im Hinblick auf das diskutierte Ergebnismuster der Altersunterschied zwischen den (älteren) BeobachterInnen und dem (jüngeren) Akteur eine Rolle spielte, da diese Differenz von den BeobachterInnen kaum registriert wurde. Dies könnte jedoch in weiteren Studien näher exploriert werden. Im Übrigen unterlagen vergleichbar viele TeilnehmerInnen beider Altersgruppen dem Effekt; die älteren taten dies jedoch in stärkerem Ausmaß.

Aus dem theoretisch-empirischen Hintergrund ließ sich ableiten, dass falschen Erinnerungen an die Handlungsausführung bei jüngeren und älteren Personen unterschiedliche Mechanismen unterliegen könnten. Aus Kapazitätsgründen konnte jedoch nicht die Anzahl an experimentellen Manipulationen realisiert werden, die notwendig gewesen wäre, um diese für die Gruppe älterer TeilnehmerInnen einzugrenzen; im Mittelpunkt stand daher der reine Nachweis des Phänomens unter (aus v.a. praktischen Gründen) variiertem Intervall zwischen Enkodierung und Abruf. Dennoch lassen die Ergebnisse Vermutungen über die folgenden drei möglicherweise zugrundeliegenden Mechanismen zu: 1) verminderte Qualität spezifischer Erinnerungsspuren, 2) reduziertes Spektrum berücksichtigter spezifischer Erinnerungsspuren und 3)

übermäßige Nutzung globaler Erinnerungscharakteristika, besonders von Vertrautheit (siehe Abschnitt II.1).

Zunächst ist auffällig, dass diejenigen älteren TeilnehmerInnen, die das Paradigma mit langem Intervall durchliefen, zwar gleich viele korrekte *ausgeführt*-Antworten auf Handlungsanweisungen produzierten, die sie in Phase 2 fünf- und einmalig beobachteten, aber deutlich weniger korrekte *ausgeführt*-Antworten auf solche Handlungsanweisungen gaben, die in Phase 2 nicht noch einmal präsentiert wurden. Diese wurden sogar weniger häufig als ausgeführt erinnert als die nicht ausgeführten, aber fünfmalig beobachteten Handlungen! Der extreme Abfall der Quote korrekter *ausgeführt*-Antworten von 76 bzw. 74% nach fünf- bzw. einmaliger auf 34% nach nullmaliger Präsentation (der observation-inflation-Effekt für korrekte *ausgeführt*-Antworten) zeigte sich zwar auch, aber wesentlich schwächer für die Rekognition (94 bzw. 89% vs. 69%) und ist deutlich stärker ausgeprägt als bei den jüngeren TeilnehmerInnen. Diese Befunde zeigen relativ klar, dass die TeilnehmerInnen nicht die bloße Vertrautheit als indikativ für eine *ausgeführt*-Antwort angesehen haben können, denn in diesem Falle hätte sich ein Unterschied zwischen fünf- und einmalig präsentierten Handlungsanweisungen zeigen müssen. Die sehr große Differenz von korrekten *ausgeführt*-Antworten auf solche Handlungsanweisungen, die in Phase 2 erneut präsentiert wurden und solchen, die in Phase 2 nicht mehr präsentiert wurden, lässt den Schluss zu, dass die älteren TeilnehmerInnen ihre Erinnerungsurteile stark auf den durch das Beobachten gestärkten spezifischen Erinnerungsspuren basierten – also wahrscheinlich auf lebhaften visuellen und akustischen, möglicherweise auch auf motorischen Repräsentationen. Hierdurch kam es dann auch zu falschen Erinnerungen, die entweder auf der starken Ähnlichkeit sensorischer (und möglicherweise motorischer) Repräsentationen von Ausführung und Beobachtung beruhen könnten und/ oder darauf, dass die Repräsentationen nicht in voller Bandbreite berücksichtigt und evaluiert wurden.

Hierfür spricht auch das Ergebnismuster, das anhand von zwei älteren TeilnehmerInnen gewonnen wurde, die die in Experiment II beschriebene Instruktion zur Quellenprüfung erhielten. Diese wiesen im Vergleich zur Gruppe ohne Abrufmanipulation mehr korrekte *ausgeführt*-Antworten, besonders nach fünfmaliger, kaum nach nullmaliger Videopräsentation, aber auch deutlich mehr falsche *ausgeführt*-Antworten und zwar sowohl nach fünf- als auch einmaliger, jedoch nicht nach nullmaliger Videopräsentation, auf (die Quote falscher *ausgeführt*-Antworten lag nach fünf- wie nach einmaliger Präsentation bei 60%). Da diese Ergebnisse eben nur auf den Daten zweier TeilnehmerInnen basieren, sind sie mit äußerster Vorsicht zu interpretieren. Sie könnten aber darauf hinweisen, dass die TeilnehmerInnen – wenn sie expli-

zeit hierzu aufgefordert werden – noch intensiver spezifische Erinnerungsspuren in ihre Urteile einbeziehen. Allerdings scheinen sie dann nach wie vor eher visuelle und akustische, möglicherweise auch motorische, weniger taktile oder propriozeptive Repräsentationen zu berücksichtigen. So treffen sie zwar mehr korrekte, aber auch mehr falsche *ausgeführt*-Zuordnungen von Handlungen, die in Phase 2 beobachtet wurden. Demnach könnte für den observation-inflation-Effekt bei älteren TeilnehmerInnen eher die Ähnlichkeit der Erinnerungsspuren als die Evaluation dieser entscheidend sein.

Auch die Ergebnisse von Thomas und Bulevich (2006), die zeigen konnten, dass der imagination-inflation-Effekt trotz Instruktion zur Quellenprüfung persistiert, spricht gegen eine kritische Rolle von Vertrautheit für derartige Erinnerungsfehler. Im Gegensatz zu den soeben berichteten explorativen Befunden, reduzierte sich in der genannten Untersuchung der Effekt jedoch nach Lesen der Instruktion zur Quellenprüfung. Möglicherweise könnten Beobachtung und Ausführung für ältere TeilnehmerInnen schwieriger zu trennen sein als Vorstellung und Ausführung. Dies könnte wiederum auf die besondere Bedeutung lebhafter visueller Gedächtnisspuren für die *ausgeführt*-Urteile älterer TeilnehmerInnen hinweisen.

In dieselbe Richtung weist auch der Vergleich der Ergebnismuster der beiden Stichproben älterer TeilnehmerInnen, die den Gedächtnistest nach unterschiedlich langen Intervallen durchliefen. Die TeilnehmerInnen, die das kurze Intervall durchliefen, hatten – unabhängig von der Frequenz der Beobachtung in Phase 2 – keinerlei Probleme, die ausgeführten Handlungsanweisungen (korrekt) als ausgeführt zu erinnern. Demnach wurde Vertrautheit offensichtlich nicht als cue genutzt. Des Weiteren wurden augenscheinlich auch solche Erinnerungsqualitäten als diagnostisch für die Handlungsausführung angesehen, die nicht durch die Beobachtung gestärkt worden waren, wie z.B. taktile Eindrücke. Dennoch weist ein – wenn auch kleiner Teil – der Untersuchten einen observation-inflation-Effekt auf, d.h. falsche Erinnerungen an die Handlungsausführung nach fünfmaliger Beobachtung. Demnach könnten die Erinnerungsqualitäten beider Quellen so ähnlich sein, dass sie partiell nicht unterscheidbar sind, oder zu ungenau evaluiert werden, als dass bspw. Differenzen in der Perspektive auffallen könnten.

Die geschilderten Überlegungen legen nahe, dass Vertrautheit nicht den kritischen Faktor für falsche Erinnerungen an die Handlungsausführung bei älteren Personen darstellen kann. Demnach erscheint es gerechtfertigt – trotz des Fehlens einer Kontrollgruppe, die die Handlungsanweisungen in Phase 2 nur lesen sollte – tatsächlich von einem *observation*-inflation-

Effekt, also einer Gedächtnistäuschung durch Beobachtung, zu sprechen. Unabhängig davon, welche Qualität der Beobachtung nun die kritische Rolle für die falschen Erinnerungen spielt, scheint die Gedächtnistäuschung doch auf diese und nicht auf das wiederholte Lesen der Handlungsanweisungen zurückzugehen.

Im Folgenden sollen die Zusammenhänge zwischen dem Zuwachs an sowohl korrekten als auch falschen *ausgeführt*-Antworten nach fünf- verglichen mit nullmaliger Präsentation für die beiden Abrufintervalle separat betrachtet werden. Der observation-inflation-Effekt für *korrekte* *ausgeführt*-Urteile hing in der Gruppe mit langem Intervall positiv und signifikant mit der Leistung in fast allen Verfahren zusammen, die zur Erfassung sog. frontaler Funktionen eingesetzt wurden, und daher auch mit dem kombinierten „frontalen Faktor“. Eine differenziertere Betrachtung ergab, dass die TeilnehmerInnen mit höherer frontaler Funktionalität mehr korrekte *ausgeführt*-Antworten nach fünfmaliger, jedoch weniger korrekte *ausgeführt*-Antworten nach nullmaliger Präsentation abgaben. Gleichzeitig fanden sich keine Zusammenhänge mit den erhobenen Maßen für sog. medial-temporale Gehirnfunktionen. Dies könnte darauf hinweisen, dass bzgl. der korrekten *ausgeführt*-Antworten weniger die Qualität der Erinnerungsspuren an sich als die Evaluation dieser die entscheidende Rolle für den gefundenen starken Einfluss der Präsentationshäufigkeit spielt. Möglicherweise urteilten gerade die TeilnehmerInnen mit guten frontalen Funktionen auf der Basis von spezifischen Erinnerungsspuren, die durch die Video-präsentation gestärkt wurden – und erreichten hierdurch zwar eine höhere Quote korrekter *ausgeführt*-Antworten nach fünfmaliger, jedoch gleichzeitig eine geringere Quote korrekter *ausgeführt*-Antworten nach nullmaliger Präsentation. Bzgl. der falschen *ausgeführt*-Antworten, also bzgl. des eigentlichen observation-inflation-Effekts, in derselben Gruppe älterer TeilnehmerInnen mit langem Intervall, zeigte sich nur eine vergleichsweise hohe negative Korrelation zwischen dem Zahlennachsprechen rückwärts als Maß für frontale Funktionen, das besonders den kontrollierten und strategischen Umgang mit bzw. das mentale Manipulieren von Informationen widerspiegelt (dieser Zusammenhang war noch ausgeprägter, wenn nur falsche *ausgeführt*-Antworten nach fünfmaliger Präsentation berücksichtigt wurden; $r(14)=-.608$, $p<.05$, $r^2=.370$). Zusätzlich fanden sich vergleichsweise hohe positive Korrelationen mit dem direkten (und auch leicht mit dem verzögerten) Abruf in einem Untertest der Wechsler Memory Scale, der mit medial-temporalen Gehirnaktivitäten assoziiert sein soll. Demnach könnten einerseits klarere und damit leichter zu verwechselnde Erinnerungsspuren zu der Gedächtnistäuschung beitragen, andererseits könnte die Evaluation dieser eine Rolle spielen: Je geringer die kognitive Kontrolle, desto größer war der Effekt. Allerdings sind die Korrelationen mit den Tests zur Erfassung frontaler Funktionen nicht konsistent negativ und auch durchgängig nicht signifikant, so dass die Interpretationsmöglichkeit nicht überbewertet werden sollte. Im Übrigen

ergab sich in beiden Gruppen kein Zusammenhang zwischen dem Ausmaß des observation-inflation-Effekts für korrekte und demjenigen für falsche *ausgeführt*-Antworten ($r(14)=-.051$, $p>.05$). Allerdings hingen die korrekten *ausgeführt*-Antworten (sowohl nach fünf- als auch nullmaliger Präsentation) jeweils positiv mit den falschen *ausgeführt*-Antworten nach fünfmaliger und nicht mit den falschen *ausgeführt*-Antworten nach nullmaliger Präsentation zusammen (korrekte *ausgeführt*-Antworten nach fünf- bzw. nullmaliger Präsentation und falsche *ausgeführt*-Antworten nach fünfmaliger Präsentation: $r(14)=.302$, $p>.05$; $r(14)=.468$, $p<.05$, $r^2=.219$; korrekte *ausgeführt*-Antworten nach fünf- bzw. nullmaliger Präsentation und falsche *ausgeführt*-Antworten nach nullmaliger Präsentation: $r(14)=-.005$, $p>.05$; $r(14)=.052$, $p>.05$). Dies weist darauf hin, dass korrekten und falschen Quellendiskriminationen ähnliche Mechanismen unterliegen könnten – auch wenn dies nicht in den berichteten Korrelationen zwischen Erinnerungs- und Testleistung zum Ausdruck kommt.

Die Korrelationen für die Gruppe älterer TeilnehmerInnen mit kurzem Intervall sind nur vorsichtig zu interpretieren, da nur eine geringe Variabilität sowohl der korrekten als auch der falschen *ausgeführt*-Antworten gegeben war und zudem nicht alle Testverfahren von allen TeilnehmerInnen durchlaufen wurden. Die Zusammenhänge mit den *korrekten* Erinnerungsleistungen sollen nicht weiter analysiert werden, da hier tatsächlich nahezu keine Varianz bestand. Bzgl. der *falschen* Erinnerungen, des observation-inflation-Effekts, zeigte sich ein ähnliches Muster wie in der Gruppe mit dem langem Intervall: Keine konsistenten Korrelationen mit den frontal assoziierten Maßen, aber eine hohe negative Korrelation mit dem Zahlennachsprechen rückwärts. Daneben ergaben sich – im Gegensatz zu der Gruppe mit dem langen Intervall – durchweg negative Korrelationen mit dem Untertest logical memory der Wechsler Memory Scale. Demnach könnten sowohl eine schlechtere Erinnerungsleistung, aber auch die Evaluation dieser Erinnerungen, eine Rolle für den observation-inflation-Effekt nach kurzem Intervall spielen.

Die berichteten Korrelationen entsprechen nicht vollständig den Befunden von McDaniel et al. (2006). Einerseits ergab sich für die Gruppe mit langem Intervall zumindest eine relativ hohe Korrelation zwischen der Leistung im Zahlennachsprechen rückwärts und dem Ausmaß des observation-inflation-Effekts, wenn auch für den frontalen Gesamtfaktor kein Zusammenhang nachweisbar war. Andererseits zeigte sich, dass die korrekten Erinnerungsleistungen relativ eng mit dem Abschneiden in den Verfahren zur Erfassung frontaler Funktionen zusammenhängen. Die kognitive Kontrolle während des Abrufs scheint also eine wichtige Rolle zu spielen. Warum sich dies nicht konsistent bzgl. falscher Erinnerungen zeigt, obwohl das Ausmaß fal-

scher und korrekter Erinnerungen zusammenhängt, ist eine offene Frage.

Es konnte zwar statistisch nicht nachgewiesen werden, dass die älteren TeilnehmerInnen in stärkerem Maße anfällig für die untersuchte Gedächtnistäuschung waren; das Ergebnismuster weist jedoch sehr klar in diese Richtung. Demnach könnte gerade diese Population im Alltag vermehrt Schwierigkeiten aufweisen, zu unterscheiden, ob einfache Handlungen ausgeführt wurden oder nicht. Die Ergebnisse deuten darauf hin, dass insb. ähnliche Erinnerungsspuren, also z.B. solche, die durch Beobachtung, Vorstellung oder Ausführung zu einem anderen Zeitpunkt enkodiert wurden, aber möglicherweise auch eine nur oberflächliche Evaluation dieser, ursächlich für den Effekt sein könnten. Um praktische Konsequenzen ableiten zu können, müssten die Mechanismen jedoch näher eingegrenzt werden.

II.4.4.3 Résumé und Ausblick

Die Ergebnisse zeigen insgesamt, dass sich falsche Erinnerungen an die Handlungsausführung durch Beobachtung auch bei älteren Personen hervorrufen lassen, und zwar nicht nur nach langem (zweiwöchigem), sondern auch nach kurzem (zwanzigminütigem) Intervall (hier allerdings in geringerem Ausmaß). Der Effekt unterschied sich nach langem Intervall statistisch nicht von demjenigen einer Gruppe jüngerer TeilnehmerInnen; tendenziell wiesen die älteren TeilnehmerInnen jedoch einen stärkeren Effekt auf, der möglicherweise bei größeren Stichprobenumfängen auch abzusichern gewesen wäre. Zukünftige Studien könnten auch die große Zwischengruppe an TeilnehmerInnen, die nicht untersucht wurde (30-60-Jährige) einbeziehen, um die Populationsvalidität auszuweiten und die Veränderungen im Altersverlauf aufzuzeigen.

Die älteren TeilnehmerInnen schienen nach zweiwöchigem Intervall – anders als die jüngeren TeilnehmerInnen – ihre *ausgeführt*-Urteile v.a. auf durch die Videopräsentation gestärkte, spezifische Erinnerungsqualitäten zu basieren, z.B. auf visuelle und akustische, evtl. motorische Repräsentationen. Demnach könnte eher die noch verfügbare Erinnerungsqualität als die Evaluation der Erinnerungsspuren eine Rolle für die Gedächtnistäuschung spielen. Hierauf weisen auch die Ergebnisse derjenigen älteren TeilnehmerInnen hin, die bereits nach 20 Minuten die Abfrage durchliefen. Korrelationen mit Leistungen in neuropsychologischen Testverfahren ließen demgegenüber vermuten, dass auch sog. frontale Funktionen, wie z.B. der strategische Abruf und die Kontrolle bzw. Überwachung mentaler Aktivitäten, einen Einfluss auf das Zustandekommen von *ausgeführt*-Urteilen haben könnten. Ausgeschlossen allerdings scheint

die Möglichkeit zu sein, dass falsche Erinnerungen an die Handlungsausführung bei älteren TeilnehmerInnen allein auf die Vertrautheit der Handlungsanweisungen zurückzuführen sind. Die Ergebnisse stimmen daher größtenteils mit den Befunden von Thomas und Bulevich (2006) sowie McDaniel et al. (2008) überein. Ob sich – wie von McDaniel et al. (2006) vorgeschlagen – Quellen tatsächlich in periphere und zentrale einteilen lassen, wobei nur beim Abruf ersterer frontale Funktionen eine prominente Rolle spielen, bedarf weiterer Überprüfung.

Um die Mechanismen, die dem observation-inflation-Effekt bei älteren TeilnehmerInnen zugrunde liegen, näher eingrenzen zu können, sind weitere Untersuchungsserien notwendig. Die vorliegenden Befunde weisen darauf hin, dass das Paradigma bei der Zielpopulation gut und fruchtbar eingesetzt werden kann. Wären die Mechanismen eingegrenzt, könnten Schlussfolgerungen hinsichtlich der potentiellen Vermeidbarkeit bzw. Reduzierbarkeit derartiger Lapsi gezogen werden, die möglicherweise in den Alltag übertragen werden könnten.

III SYNOPSIS DER BEFUNDE UND AUSBLICK: (WIE) KÖNNEN FREMDE HANDLUNGEN ZU EIGENEN WERDEN?

Falsche Erinnerungen an die Ausführung einfacher Handlungen sind experimentell bis dato ausschließlich durch Vorstellungen hervorgerufen worden. Unter Einnahme einer rein kognitionspsychologischen Perspektive wurde der sog. *imagination-inflation-Effekt* (Garry, Manning, Loftus & Sherman, 1996; Goff & Roediger, 1998) vorrangig als Resultat einer Fehlattribution von durch die Vorstellung generierten, qualitativen Erinnerungsspuren bzw. sensorischen Eindrücken interpretiert. Eine Fehlattribution von Spurstärke als globalem Erinnerungscharakteristikum wurde fast durchgängig ausgeschlossen (z.B. Thomas, Bulevich & Loftus, 2003; Lampinen, Odegard & Bullington, 2003). Bei den zugrundeliegenden Operationalisierungen wurde allerdings nicht die Dimension des Informationsursprungs (internal vs. external) bzw. des Status' der Person (aktiv vs. passiv) beachtet, die jedoch bekanntermaßen über die Verarbeitungstiefe (Craik & Lockhart, 1972) einen Einfluss auf die Spurstärke ausübt. So wurden bspw. Vorstellen und Lesen miteinander verglichen. Beide Prozesse unterscheiden sich aber nicht nur darin, dass beim Vorstellen im Gegensatz zum Lesen sensorische Eindrücke generiert werden, sondern auch darin, dass durch den aktiven Prozess des Generierens von Vorstellungsbildern eine tiefere Verarbeitung erfolgt als durch den passiven Prozess des Lesens. Aber nicht nur müssen die bisherigen Erkenntnisse aufgrund des genannten Mangels mit Vorsicht betrachtet werden, möglicherweise stellt auch die rein kognitionspsychologische eine verkürzte Betrachtungsweise des Phänomens dar. Aufgrund der gewählten Perspektive blieben neurowissenschaftliche Befunde unbeachtet, die – neben einer sensorischen – einhellig auch eine motorische Simulation während des Vorstellens aufzeigen (z.B. Jeannerod, 2001). Demnach existiert also eine Überlappung derjenigen motorischen Areale, die während der tatsächlichen und der bloß vorgestellten Handlungsausführung aktiviert sind (z.B. Grèzes & Decety, 2001). Solche motorischen Repräsentationen können – wie auch für den Fall der tatsächlichen Ausführung gezeigt – beim Abruf reaktiviert werden und damit möglicherweise eine weitere Quelle falscher Erinnerungen darstellen (z.B. Nilsson et al., 2000; Senkfor, Van Petten & Kutas, 2002).

Der zentrale Beitrag der vorliegenden Arbeit besteht darin, neben der Vorstellung erstmals einen weiteren Verarbeitungsprozess aufgezeigt zu haben, der zu falschen Erinnerungen an die eigene Handlungsausführung führen kann: Beobachtung. Diese Form der Verarbeitung war als Untersuchungsgegenstand ausgewählt worden, da ein rapide wachsender Corpus an neurowissenschaftlichen Befunden einhellig zeigt, dass auch die Beobachtung als eine Form der motorischen Simulation begriffen werden kann (z.B. Rizzolatti & Craighero, 2004). Demnach

lösen fremde und eigene Handlungsausführung ähnliche motorische Repräsentationen aus, die wiederum mit der bloß vorgestellten Handlungsausführung überlappen (z.B. Grèzes & Decety, 2001). Zudem existiert eine Überlappung auf sensorischer Ebene; so können z.B. Geräusche sowohl der eigenen als auch der fremden Handlungsausführung entstammen (sowie vorgestellt sein).

Ein zum imagination-inflation- analoger *observation-inflation-Effekt* konnte innerhalb einer Serie von vier Experimenten, an denen insg. über 200 Personen teilnahmen, mehrfach repliziert werden. Hierdurch werden bisherige Erkenntnisse über das Zustandekommen von Gedächtnistäuschungen für die Ausführung einfacher Handlungen auf mehreren Ebenen erheblich erweitert. Die Befunde stehen zudem in Einklang mit Erkenntnissen zum sog. *enactment-Effekt* (z.B. Cohen, 1981; Engelkamp & Krumnacker, 1980), d.h. der besseren Erinnerbarkeit von ausgeführten gegenüber nicht ausgeführten Handlungen.

Durch experimentelle Manipulation der Phasen Enkodierung, weitere Verarbeitung und Abruf konnten die zugrundeliegenden Mechanismen genauer eingegrenzt werden und so ein Ausbau auf theoretischer Ebene erfolgen. Die gewonnenen Erkenntnisse sind nicht mit der Vorstellung vereinbar, dass falsche (und daher vermutlich auch korrekte) Erinnerungen an die Ausführung einfacher Handlungen bei jüngeren Personen ausschließlich durch eine Fehlattribution sensorischer Erinnerungsspuren zustande kommen. Vielmehr scheinen auch (die bisher nicht beachteten) motorischen Repräsentationen eine – oder sogar die – kritische Rolle zu spielen. Bei der Interpretation dieser Befunde ist allerdings zu berücksichtigen, dass die auf behavioraler Ebene erhobenen Daten keine Überprüfung einer gelungenen Manipulation motorischer Repräsentationen erlaubten. Auch existieren keine Vorläuferstudien, die die gewählte Operationalisierung bereits empirisch überprüft hätten. Die Annahme konvergiert jedoch mit Befunden zum enactment-Effekt, die auf eine bedeutsame Rolle motorischer Repräsentationen hinweisen (z.B. Nilsson et al., 2000; Senkfor, Van Petten & Kutas, 2002). Ein Einfluss des Informationsursprungs und der Verarbeitungsleichtigkeit bzw. Vertrautheit auf die Gedächtnistäuschung konnte anhand verschiedener Indikatoren mit großer Sicherheit ausgeschlossen werden. Darüber hinaus konnte eindeutig gezeigt werden, dass es sich bei dem gefundenen Effekt nicht um ein methodologisches Artefakt oder das Resultat einer Antworttendenz bzw. eines bestimmten Antwortkriteriums handelt.

Im Alltag können gerade falsche Erinnerungen an einfache Handlungen (z.B. das Einnehmen einer Tablette, das Ausstellen des Herdes) gravierende Konsequenzen nach sich ziehen. Die gesammelten Befunde lassen jedoch wenig Spielraum für die Hoffnung, dass Fehlattributionen sensorischer und/ oder motorischer Eindrücke durch metakognitive Strategien a posteriori vermieden werden könnten. Allerdings spielen im Alltag selbstverständlich eine Reihe weiterer Faktoren eine Rolle, die im Rahmen der experimentellen Untersuchungen konstant gehalten bzw. ausgeschaltet wurden. So kommen (u.a.) der Aufmerksamkeitszuwendung bei der Enkodierung, dem Intervall zwischen Enkodierung und Abruf und der Komplexität der Handlung, aber auch dem Alter der erinnernden Person entscheidende Rollen zu. Die geschilderten Befunde gelten dabei nur für jüngere Erwachsene, d.h. Personen zwischen etwa 18 und 30 Jahren. Aufgrund verschiedener Studien zu Gedächtnisleistungen im Alter konnte angenommen werden, dass falschen Erinnerungen an einfache Handlungen bei älteren Erwachsenen, d.h. Personen über 60 Jahren, andere Mechanismen unterliegen bzw. die erwähnten Mechanismen anders gewichtet werden (z.B. Craik, 2000; McDaniel et al., 2008). Der Einbezug älterer TeilnehmerInnen konnte aufgrund der beschränkten Zugänglichkeit dieser Population und daher limitierten Anzahl experimenteller Manipulationen nur begrenzte Einsichten liefern. Konvergierend mit existierenden Untersuchungen scheint aber auch bei dieser Population eine Fehlattribution von Vertrautheit nicht den kritischen Mechanismus darzustellen. Inwiefern der Qualität respektive der Evaluation der Erinnerungsspuren die entscheidende Rolle zukommt und inwieweit motorische Repräsentationen eine Rolle spielen, konnte nicht geklärt werden. Auch wenn sich bzgl. des Ausmaßes falscher Erinnerungen an die Handlungsausführung kein signifikanter Unterschied zwischen den untersuchten Altersgruppen zeigte, erbrachte ein Vergleich korrekter *ausgeführt*-Urteile doch erstaunliche Differenzen. Die älteren TeilnehmerInnen wiesen ein deutlich höheres Ausmaß korrekter Erinnerungen an die Handlungsausführung bzgl. solcher Handlungsanweisungen auf, deren Ausführung sie auch beobachteten, verglichen mit solchen, die sie nicht beobachteten. Ein solcher observation-inflation-Effekt für korrekte Erinnerungen an die Handlungsausführung ließ sich zwar auch der Tendenz nach bei den jüngeren TeilnehmerInnen finden, jedoch nicht in der bei den älteren gefundenen Deutlichkeit. Demnach scheinen älterer Personen beim Abruf besonders solche Erinnerungsqualitäten als diagnostisch für die Handlungsausführung anzusehen, die durch die Beobachtung gestärkt werden, also möglicherweise lebendige visuelle und akustische, evtl. auch motorische Repräsentationen.

Auch auf methodologischer Ebene konnten Fortschritte erzielt werden. Durch die Nutzung des von außen kontrollierbaren Prozesses der Beobachtung im Vergleich zur Vorstellung lassen sich falsche Erinnerungen an die Handlungsausführung auch dann untersuchen, wenn Popula-

tionen mit eingeschränktem Instruktionsverständnis bzw. eingeschränkter compliance einbezogen werden sollen. Zudem ist die interindividuelle Variabilität bei der Beobachtung gegenüber der Vorstellung deutlich reduziert, so dass eine bessere interindividuelle Vergleichbarkeit gewährleistet ist. Darüber hinaus konnte das eingesetzte Paradigma so modifiziert werden, dass es bei eindeutigerem Informationsgehalt ökonomischer in der Durchführung wurde und somit auch bei weniger belastbaren TeilnehmerInnen eingesetzt werden kann.

Wie bereits erwähnt besteht die deutlichste Limitierung der gewonnenen Erkenntnisse an jüngeren TeilnehmerInnen darin, dass die (vermutete) Manipulation der motorischen Repräsentationen nicht überprüft und auch nicht auf bestehenden Operationalisierungen basiert werden konnte. Daher sollte die aufgestellte Simulationshypothese weiteren empirischen Prüfungen unterzogen werden. Da sich sensorische und motorische Repräsentationen auf behavioraler Ebene nur sehr schlecht separat manipulieren lassen, sollten hierbei zukünftig v.a. bildgebende und elektrophysiologische Verfahren zum Einsatz kommen, die direkte Evidenz für (oder gegen) die Aktivität motorischer Areale bei der Enkodierung bzw. beim Abruf erbringen könnten. Varianten der Operationalisierung bereits untersuchter Prozesse (z.B. Beobachtung eines tatsächlich anwesenden Gegenübers anstatt der Beobachtung per Video, andere Formen des aktiven Generierens als das Umstellen von Sätzen) sowie die Übertragung des Labor- in ein Feldsetting können helfen, den Geltungsbereich der Erkenntnisse auf Konstrukt- und Situationsebene auszuweiten. Nachfolgende Untersuchungen sollten auch die Händigkeit der TeilnehmerInnen (im Vergleich zu der beobachteten Person) berücksichtigen, da zu vermuten ist, dass sich diese auch auf die Repräsentation von Bewegungen auswirkt (z.B. De'Sperati & Stucchi, 1997).

Lohnenswert erscheint ebenfalls ein weitergehender Vergleich verschiedener Simulationsprozesse (Vorstellung, Beobachtung, Handlungsvorbereitung usw.) hinsichtlich ihrer Ähnlichkeiten und Differenzen bei der Erzeugung falscher Erinnerungen an die Handlungsausführung.

Des Weiteren sollten die Befunde an älteren TeilnehmerInnen weiter ausgebaut werden. Hierbei ist zunächst einmal an Manipulationen gedacht, die eine explizite Kontrolle von Vertraulichkeit beinhalten und eine Trennung der Qualität und der Evaluation der Erinnerungsspuren erlauben. Denn würde bspw. die Evaluation von Erinnerungsspuren eine kritische Rolle spielen, ließen sich aus den gewonnenen Erkenntnissen möglicherweise metakognitive Strategien ableiten, die den älteren TeilnehmerInnen im Alltag hilfreich sein könnten (z.B. Budson, Dodson, Daffner & Schacter, 2005). Aber auch im anderen Falle ließen sich praktische Implikatio-

nen gewinnen: So könnten sich bspw. Manipulationen der Enkodierung, die sich auf die Qualität der Erinnerungsspuren auswirken, nützlich erweisen (z.B. Thomas & Sommers, 2005).

Darüber hinaus könnten durch den Einschluss verschiedener klinischer Stichproben nicht nur die Mechanismen und zugrundeliegenden neuronalen Korrelate eingegrenzt, sondern auch Implikationen für Diagnostik und Therapie abgeleitet werden. So könnten bspw. Populationen, die über ein beeinträchtigtes Spiegelneuronensystem und/ oder beeinträchtigte motorische Systeme verfügen (wie möglicherweise Autisten oder Schizophrene; Brüne, 2005; Buccino & Amore, 2008; Daprati, Nico, Franck & Sirigu, 2003; Dapretto et al., 2006; Oberman, Hubbard, McCleery, Altschuler, Ramachandran & Pineda, 2005), untersucht werden. Des Weiteren ist an den Einbezug von PatientInnen mit Gedächtnisbeeinträchtigungen und/ oder Beeinträchtigungen exekutiver Funktionen (z.B. dementer TeilnehmerInnen bzw. TeilnehmerInnen mit fokalen frontalen/ medialen Schädigungen) zu denken.

Eine Reihe möglicher weiterer Fragestellungen ergibt sich aus kognitions- und sozialpsychologischen Perspektiven. So könnten innerhalb des modifizierten Paradigmas verschiedene klassische gedächtnispsychologische Manipulationen vorgenommen werden (z.B. Manipulation des Behaltensintervalls, Manipulation der Aufmerksamkeit bei der Enkodierung), um das Phänomen weiter einzugrenzen. Auch die Effekte der beiden untersuchten Beobachtungsperspektiven könnten unter Einbezug von sozial-kognitiven Theorien (z.B. *theory of mind*; z.B. Frith & Frith, 2003; siehe auch Apperly, 2008) weiter elaboriert werden.

Der besondere Beitrag der vorliegenden Arbeit besteht darin, gezeigt zu haben, dass fremde Handlungen in der Erinnerung zu eigenen werden können. Die diesem Phänomen unterliegenden Mechanismen konnten eingegrenzt, jedoch nicht abschließend beurteilt werden. Damit konnten die durchgeführten Untersuchungen die Erkenntnisse bzgl. falscher Erinnerungen an die Ausführung einfacher Handlungen hinsichtlich mehrerer Dimensionen erweitern. Aus den gewonnenen Befunden zum observation-inflation-Effekt konnten darüber hinaus zahlreiche weiterführende Fragestellungen und Anknüpfungspunkte für verschiedene psychologische Teildisziplinen abgeleitet werden. Für derartige Untersuchungen kann das im Rahmen der vorliegenden Arbeit modifizierte, gut handhabbare und ökonomische Paradigma genutzt werden, anhand dessen der Effekt bereits mehrfach zuverlässig repliziert werden konnte.

IV KURZZUSAMMENFASSUNG

Falsche Erinnerungen an die Ausführung einfacher Handlungen sind experimentell bis dato ausschließlich durch Vorstellungen hervorgerufen worden. Unter Einnahme einer rein kognitionspsychologischen Perspektive wurde der sog. *imagination-inflation-Effekt* (Garry, Manning, Loftus & Sherman, 1996; Goff & Roediger, 1998) vorrangig als Resultat einer Fehlattribution von durch die Vorstellung generierten, sensorischen Eindrücken interpretiert (Thomas, Bulevich & Loftus, 2003). Neurowissenschaftliche Erkenntnisse zeigen jedoch, dass das Vorstellen nicht nur als eine Form der perzeptuellen, sondern auch der motorischen Simulation der Handlungsausführung begriffen werden kann. Derartigen Studien zufolge werden auch beobachtete Handlungen motorisch simuliert, so dass eine Überschneidung derjenigen motorischen Areale existiert, die während Ausführung, Vorstellung und Beobachtung aktiviert sind (Grèzes & Decety, 2001). Zusätzlich können sensorische Eindrücke sowohl der eigenen als auch der beobachteten Ausführung entstammen (sowie vorgestellt sein). Vor diesem Hintergrund wurde der Frage nachgegangen, ob auch die Beobachtung – ebenso wie die Vorstellung – von Handlungen zu falschen Erinnerungen an die Ausführung dieser Handlungen führen kann. Innerhalb von drei Studien an jungen Erwachsenen konnte ein zum *imagination-inflation-Effekt* analoger *observation-inflation-Effekt* nach Beobachtung eines Gegenübers nachgewiesen werden, während Verarbeitungsformen, die keine sensorischen bzw. motorischen Repräsentationen hervorrufen sollten (Lesen, Generieren), nicht zu einem Anstieg falscher Erinnerungen führten. Der Effekt persistierte, wenn vor dem Abruf eine zusätzliche Instruktion zur Quellenprüfung bzw. eine Warnung ausgesprochen wurde. Er war sogar dann nachweisbar, wenn eine Überlappung sensorischer Repräsentationen von Ausführung und Beobachtung minimiert wurde. Wurde allerdings – bei maximierter sensorischer Überlappung von Ausführung und Beobachtung – der soziale Charakter der Beobachtungssituation reduziert (Beobachtung aus der 1.-Person-Perspektive), verringerte sich das Ausmaß falscher Erinnerungen. Demnach scheint dem Effekt eine Fehlattribution von sensorischen und/ oder durch Simulation hervorgerufenen motorischen Repräsentationen zugrunde zu liegen, wobei letztere möglicherweise den kritischeren Faktor darstellen könnten. Zukünftige Studien, die neben behavioralen auch neuronale Daten erfassen, sind notwendig, um den Beitrag beider Mechanismen näher zu beleuchten. Die reine Vertrautheit mit den Handlungsanweisungen und der Informationsursprung scheinen insgesamt keine Rolle bei der Erzeugung der falschen Erinnerungen zu spielen. Diese Befunde werden durch Daten ergänzt, die an älteren TeilnehmerInnen gewonnen wurden und vermuten lassen, dass diese ihre Erinnerungsurteile bzgl. der Handlungsausführung ebenfalls nicht auf Vertrautheit basieren. Inwiefern bei dieser Population – im Gegensatz zu den jungen Erwachsenen – neben der Qualität auch die Evaluation der Erinnerungsspuren zu der Gedächtnistäuschung beiträgt, und welche Rolle sensorische respektive motorische Repräsentationen spielen, ist in weiteren Studien zu ergründen.

V LITERATURVERZEICHNIS

- Anderson, N.D. & Craik, F.I.M. (2000). Memory in the aging brain. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford Handbook of Memory* (411-425). Oxford: University Press.
- Anderson, N.D., Craik, F.I.M. & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults. I: Evidence from divided attention costs. *Psychology and Aging, 13*, 405-423.
- Apperly, I.A. (2008). Beyond simulation-theory and theory-theory: Why social cognitive neuroscience should use its own concepts to study "theory of mind". *Cognition, 107*, 266-283.
- Arbib, M.A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences, 28*, 105-124.
- Ayers, M.S. & Reder, L.M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based model. *Psychonomic Bulletin and Review, 5*, 1-21.
- Baddeley, A.D. (1986). *Working Memory*. London: Oxford University Press.
- Bäckman, L. & Nilsson, L.-G. (1984). Aging effects in free recall: An exception to the rule. *Human Learning, 3*, 53-69.
- Bäckman, L. & Nilsson, L.-G. (1985). Prerequisites for lack of age-differences in memory performance. *Experimental Aging Research, 11*, 67-73.
- Balota, D.A., Cortese, M., Duchek, J.M., Adams, D., Roediger, H.L., McDermott, K.B. & Yerys, B. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer type. *Cognitive Neuropsychology, 16*, 361-384.
- Balota, D.A., Dolan, P.O. & Duchek, J.M. (2000). Memory changes in healthy older adults. In E. Tulving & F.I.M. Craik (Eds.). *The Oxford Handbook of Memory* (395-409). Oxford: University Press.
- Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.
- Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge: University Press.
- Belopolsky, A.V., Olivers, C.N.L. & Theeuwes, J. (2008). To point a finger: Attentional and motor consequences of observing pointing movements. *Acta Psychologica, 128*, 56-62.
- Benton, A.L. & Hamsher, K. de S. (1976). *Multilingual Aphasia Examination Manual*. Iowa City: University of Iowa.
- Bernstein, D.M., Godfrey, R.D., Davison, A. & Loftus, E.F. (2004). Conditions affecting the revelation effect for autobiographical memory. *Memory and Cognition, 32*, 455-462.
- Bernstein, E.M. & Putnam, F.W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease, 174*, 727-735.
- Bernstein, D.M., Whittlesea, B.W.A. & Loftus, E.F. (2002). Increasing confidence in remote autobiographical memory and general knowledge: Extensions of the revelation effect. *Memory and Cognition, 30*, 432-438.
- Blakemore, S.J. & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2*, 561-567.
- Blakemore, S.J. & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology, 13*, 219-224.
- Blank, H. (1998). Memory states and memory tasks: An integrative framework for eyewitness memory and suggestibility. *Memory, 6*, 481-529.

- Bonda, E., Petrides, M., Ostry, D. & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, *16*, 3737-3744.
- Borg, E. (2007). If mirror neurons are the answer, what was the question? *Journal of Consciousness Studies*, *14*, 5-19.
- Borst, G. & Kosslyn, S.M. (2008). Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory and Cognition*, *36*, 849-862.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Bower, G.H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, *9*, 529-533.
- Brand, M., Kalbe, E., Fujiwara, E., Huber, M. & Markowitsch, H.-J. (2003). Cognitive estimation in patients with probable Alzheimer's disease and alcoholic Korsakoff patients. *Neuropsychologia*, *41*, 575-584.
- Bransford, J.D. & Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *2*, 331-350.
- Brass, M., Bekkering, H., Wohlschläger A. & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*, 124-143.
- Brass, M., Derrfuss, J., Matthes-von Cramon, G. & von Cramon, D.Y. (2003). Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology*, *17*, 265-271.
- Brass, M. & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*, 489-495.
- Brewer, W.F. (1977). Memory for the pragmatic implications of sentences. *Memory and Cognition*, *5*, 673-678.
- Brüne, M. (2005). Theory of mind in schizophrenia: A review of the literature. *Schizophrenia Bulletin*, *31*, 21-42.
- Buccino, G. & Amore, M. (2008). Mirror neurons and the understanding of behavioral symptoms in psychiatric disorders. *Current Opinion in Psychiatry*, *21*, 281-285.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G. & Freund, H.J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400-404.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C.A. & Rizzolatti, G. (2004a). Neural circuits involved in the recognition of actions performed by non con-specifics: An fMRI study. *Journal of Cognitive Neuroscience*, *16*, 114-126.
- Buccino, G., Vogt, S., Ritzi, A., Fink, G.R., Zilles, K., Freund, H.-J. & Rizzolatti, G. (2004b). Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron*, *42*, 323-334.
- Buchner, A. & Brandt, M. (2002). Gedächtniskonzeptionen und Wissensrepräsentationen. In J. Müsseler & W. Prinz (Hrsg.). *Allgemeine Psychologie* (493-543). Heidelberg: Spektrum.
- Buckner, R.L. & Logan, J.M. (2002). Frontal contributions to episodic memory encoding in the young and elderly. In A. Parker, E.L. Wilding & T.J. Bussey (Eds.), *The Cognitive Neuroscience of Memory. Encoding and Retrieval* (59-81). Hove: Psychology Press.
- Budson, A.E., Dodson, C.S., Daffner, K.R. & Schacter, D.L. (2005). Metacognition and false recognition in Alzheimer's Disease: Further exploration of the distinctiveness heuristic. *Neuropsychology*, *19*, 253-258.

- Budson, A.E., Michalska, K.J., Sullivan, A.L., Rentz, D.M., Daffner, K.R. & Schacter, D.L. (2003). False Recognition in Alzheimer Disease: Evidence from Categorized Pictures. *Cognitive and Behavioral Neurology*, *16*, 16-27.
- Budson, A.E., Sullivan, A.L., Mayer, E., Daffner, K.R., Black, P.M. & Schacter, D.L. (2002). Suppression of false recognition in Alzheimer's disease and in patients with frontal lobe lesions. *Brain*, *125*, 2750-2765.
- Cabeza, R. (2006). Prefrontal and medial temporal lobe contributions to relational memory in young and older adults. In H.D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience* (595-626). Oxford: University Press.
- Calvo-Merino, B., Glaser, D.E., Grèzes, J., Passingham, R.E. & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, *15*, 1243-1249.
- Calvo-Merino, B., Grèzes, J., Glaser, D.E., Passingham, R.E. & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, *16*, 1905-1910.
- Chambers, K.L. & Zaragoza, M.S. (2001). Intended and unintended effects of explicit warnings on eyewitness suggestibility: Evidence from source identification tests. *Memory and Cognition*, *29*, 1120-1129.
- Clark, S., Tremblay, F. & Ste-Marie, C. (2003). Differential modulation of corticospinal excitability during observation, mental imagery and imitation of hand actions. *Neuropsychologia*, *42*, 105-112.
- Cochin, S., Barthelemy, C., Lejeune, B., Roux, S. & Martineau, J. (1998). Perception of motion and EEG activity in human adults. *Electroencephalography and Clinical Neurophysiology*, *107*, 287-295.
- Cochin, S., Barthelemy, C., Roux, S. & Martineau, J. (1999). Observation and execution of movement: Similarities demonstrated by quantified electroencephalography. *The European Journal of Neuroscience*, *11*, 1839-1842.
- Cohen, R.L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology*, *22*, 267-281.
- Cohen, G. & Faulkner, D. (1989). Age differences in source forgetting: Effects on reality monitoring and eyewitness testimony. *Psychology and Aging*, *4*, 10-17.
- Cohen, R.L., Sandler, S.P. & Schroeder, K. (1987). Aging and memory for words and action events: Effects of item repetition and list length. *Psychology and Aging*, *2*, 280-285.
- Cohen-Seat, G., Gastaut, H., Faure, J. & Heuyer, G. (1954). Études expérimentales de l'activité nerveuse pendant la projection cinématographique. *Revue Internationale de Filmologie*, *5*, 7-64.
- Craik, F.I.M. (2006). Remembering items and their contexts: Effects of ageing and divided attention. In H.D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience* (571-594). Oxford: University Press.
- Craik, F.I.M. & Byrd, M. (1982). Aging and cognitive deficits. In F.I.M. Craik & S. Trehub (Eds.). *Aging and Cognitive Processes* (191-211). New York: Plenum.
- Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.
- Daprati, E., Nico, D., Franck, N. & Sirigu, A. (2003). Being the agent: Memory for action events. *Consciousness and Cognition*, *12*, 670-683.
- Daprati, E., Nico, D., Saimpont, A., Franck, N. & Sirigu, A. (2005). Memory and action: An experimental study on normal subjects and schizophrenic patients. *Neuropsychologia*, *43*, 281-293.

- Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y. & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*, 28-30.
- David, N., Cohen, M.X., Newen, A., Bewernick, B.H., Shah, N.J., Fink, G.R. & Vogeley, K. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *NeuroImage*, *36*, 1004-1014.
- David, N., Newen, A. & Vogeley, K. (2008). The 'sense of agency' and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, *17*, 523-534.
- Decety, J. (2002). Is there such a thing as functional equivalence between imagined, observed, and executed action? In A.N. Meltzoff & W. Prinz (Eds.), *The Imitative Mind. Development, Evolution, and Brain Bases* (291-310). Cambridge: University Press.
- Decety, J. & Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Conscious Cognition*, *12*, 577-596.
- Decety, J. & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4-14.
- Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*, 1763-1777.
- Delis, D.C., Kramer, J., Kaplan, E. & Ober, B.A. (1987). *The California Verbal Learning Test*. San Antonio: Psychological Corp.
- Dennis, N.A, Kim, H. & Cabeza, R. (2008). Age-related differences in brain activity during true and false memory retrieval. *Journal of Cognitive Neuroscience*, *20*, 1390-1402.
- De'Sperati, C. & Stucchi, N. (1997). Recognizing the motion of graspable object is guided by handedness. *NeuroReport*, *8*, 2761-2765.
- Dick, M.B., Kean, M.L. & Sands, D. (1989). Memory for action events in Alzheimer type dementia: Further evidence for an encoding failure. *Brain and Cognition*, *9*, 71-87.
- Diehl, J.M. & Arbinger, R. (1992). *Einführung in die Inferenzstatistik*. Eschborn: Klotz.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176-180.
- Dodson, C.S. & Johnson, M.K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology*, *106*, 541-557.
- Drivdahl, S.B. & Zaragoza, M.S. (2001). The role of perceptual elaboration and individual differences in the creation of false memories for suggested events. *Applied Cognitive Psychology*, *15*, 265-281.
- Echterhoff, G., Groll, S. & Hirst, W. (2007). Tainted truth: Overcorrection for misinformation influence on eyewitness memory. *Social Cognition*, *25*, 367-409.
- Echterhoff, G., Hirst, W. & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory and Cognition*, *33*, 770-782.
- Engelkamp, J. (1997). *Das Erinnern eigener Handlungen*. Göttingen: Hogrefe.
- Engelkamp, J. & Jahn, P. (2003). Lexical, conceptual and motor information in memory for action phrases: A multi-system account. *Acta Psychologica*, *113*, 147-165.
- Engelkamp, J. & Krumnacker, H. (1980). Imaginale und motorische Prozesse beim Behalten verbalen Materials. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *27*, 511-533.
- Engelkamp, J. & Zimmer, H.D. (1983). Zum Einfluß von Wahrnehmen und Tun auf das Behalten von Verb-Objekt-Phrasen. *Sprache und Kognition*, *2*, 117-127.
- Engelkamp, J. & Zimmer, H.D. (1984). *The Human Memory*. Seattle, WA: Hogrefe & Huber.

- Engelkamp, J. & Zimmer, H.D. (1994). Motor similarity in subject-performed tasks. *Psychological Research, 57*, 47-53.
- Engelkamp, J., Zimmer, H.D. & Biegelmann, U.E. (1983). Bizarreness effects in verbal tasks and subject-performed tasks. *European Journal of Cognitive Psychology, 5*, 393-415.
- Engelkamp, J., Zimmer, H.D., Mohr, G. & Sellen, O. (1994). Memory of self-performed tasks: Self-performing during recognition. *Memory & Cognition, 22*, 34-39.
- Epstein, S., Pacini, R., Denes-Ray, V. & Heider, H. (1996). Individual differences in intuitive-experiential and analytic-rational thinking styles. *Journal of Personality and Social Psychology, 71*, 390-405.
- Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*, 2608-2611.
- Farah, M.J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal of Experimental Psychology: General, 114*, 91-103.
- Farrer, C., Franck, N., Paillard, J. & Jeannerod, M. (2003). The role of proprioception in action recognition. *Consciousness and Cognition, 12*, 609-619.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.
- Ferguson, S.A., Hashtroudi, S. & Johnson, M.K. (1992). Age differences in using source-relevant cues. *Psychology and Aging, 7*, 443-452.
- Finke, R.A. (1980). Levels of equivalence in imagery and perception. *Psychological Review, 87*, 113-132.
- Fletcher, P.C., Frith, C.D., Grasby, P.M., Shallice, T., Frackowiak, R.S.J. & Dolan, R.J. (1995). Brain systems for encoding and retrieval of auditory-verbal memory: An in vivo study in humans. *Brain, 118*, 401-416.
- Folstein, M.F., Folstein, S.F. & McHugh, P.R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.
- Frith, C.D. & Frith, U. (2003). Development and neurophysiology of mentalizing. *Neuron, 50*, 531-534.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences, 4*, 14-21.
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593-609.
- Gallese, V. & Goldman, A.I. (1998). Mirror neurons and the simulation theory. *Trends in Cognitive Sciences, 2*, 493-501.
- Gallo, D., Roediger, H.L. & McDermott, K.B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin and Review, 8*, 579-586.
- Gangitano, M., Mottaghy, F.M. & Pascual-Leone, A. (2001). Phase specific modulation of cortical motor output during movement observation. *NeuroReport, 12*, 1489-1492.
- Ganis, G., Thompson, W.L. & Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research, 20*, 226-241.
- Garry, M., Frame, S. & Loftus, E.F. (1999). Lie down and let me tell you about your childhood. In S. Della (Ed.). *Mind Myths: Exploring Popular Assumptions about the Mind and the Brain* (113-124). New York: Wiley.
- Garry, M. & Hayes, J.H. (1999). *Imagination Inflation Depends on when the Imagined Event Occurred*. Unpublished Manuscript, Victoria University of Wellington.

- Garry, M., Manning, C.G., Loftus, E.F. & Sherman, S.J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin and Review*, *3*, 208-214.
- Garry, M. & Polaschek, D.L.L. (2000). Imagination and Memory. *Current Directions in Psychological Science*, *9*, 6-10.
- Gastaut, H.J. & Bert, J. (1954). EEG changes during cinematographic presentation. *Electroencephalography and Clinical Neurophysiology*, *6*, 433-444.
- Glaser, W. (2004). Imagery. In H.O. Häcker & K.-H. Stapf (Hrsg.). *Psychologisches Wörterbuch* (433). Bern: Hans Huber.
- Glisky, E.L., Polster, M.R. & Routhieaux, B.C. (1995). Double dissociation between item and source memory. *Neuropsychology*, *9*, 229-235.
- Glisky, E.L., Rubin, S.R. & Davidson, P.S.R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1131-1146.
- Goff, L.M. & Roediger, H.L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory and Cognition*, *26*, 20-33.
- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M. & Prinz, W. (2007). Predicting point-light actions in real-time. *NeuroImage*, *36*, T22-T32.
- Graf, M., Schütz-Bosbach, S. & Prinz, W. (in press). Motor involvement in action and object perception: Similarity and complementarity. In G. Semin & G. Echterhoff (Eds.), *Grounding Sociality: Neurons, Minds, and Culture*. New York: Psychology Press.
- Grafton, S.T., Arbib, M.A., Fadiga, L. & Rizzolatti, G. (1996). Localization of grasp representations in humans by Positron Emission Tomography. *Experimental Brain Research*, *112*, 103-111.
- Greene, E., Flynn, M.S. & Loftus, E.F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior*, *21*, 207-219.
- Greenwald, A.G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, *77*, 73-99.
- Greenwald, A.G. (1972). On doing two things at once: Time sharing as a function of ideomotor compatibility. *Journal of Experimental Psychology*, *94*, 52-57.
- Gregory, R.L. & Gombrich, E.H. (1973). *Illusion in Nature and Art*. London: Duckworth.
- Grèzes, J., Costes, N. & Decety, J. (1998). Top down effect of the strategy to imitate on the brain areas engaged in perception of biological motion: A PET investigation. *Cognitive Neuropsychology*, *15*, 553-582.
- Grèzes, J. & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*, 1-19.
- Guttentag, R.E. & Hunt, R.R. (1988). Adult age differences in memory for imagined and performed actions. *Journal of Gerontology: Psychological Sciences*, *43*, 107-108.
- Hager, W. (1987). Grundlagen einer Versuchsplanung zur Prüfung empirischer Hypothesen in der Psychologie. In G. Lüer (Hrsg.), *Allgemeine Experimentelle Psychologie. Eine Einführung in die methodischen Grundlagen mit praktischen Übungen für das Experimentelle Praktikum* (43-264). Stuttgart: Gustav Fischer.
- Hari, R., Foss, N., Avikainen, S., Kirveskari, S., Salenius, S. & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 15061-15065.
- Harris, P.L. (2000). *The Work of the Imagination*. Oxford: Blackwell Publishers.

- Hart, R.P., Kwentus, J.A., Wade, J.B. & Taylor, J.R. (1988). Modified Wisconsin Card Sorting Test in elderly normal, depressed and demented patients. *Clinical Neuropsychologist*, *2*, 49-56.
- Hashtroudi, S., Johnson, M.K. & Chrosniak, L.D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging*, *5*, 119-126.
- Heaps, C. & Nash, M. (1999). Individual differences in imagination inflation. *Psychonomic Bulletin and Review*, *6*, 313-318.
- Heil, M., Rolke, B., Engelkamp, J., Rosler, F., Ozcan, M. & Hennighausen, E. (1999). Event-related brain potentials during recognition of ordinary and bizarre action phrases following verbal and subject-performed encoding conditions. *European Journal of Cognitive Psychology*, *2*, 261-280.
- Helmholtz, H. von (1903). *Popular Lectures on Scientific Subjects*. New York: Longmans, Green. [Original publiziert 1881]
- Henkel, L.A., Franklin, N. & Johnson, M.K. (2000). Cross-modal source monitoring confusions between perceived and imagined events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 321-335.
- Hintzman, D.L. (1974). Theoretical implications of the spacing effect. In R.L. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium* (77-99). Potomac: Lawrence Erlbaum.
- Hoffman, H.G. (1997). Role of memory strength in reality monitoring decisions: Evidence from source attribution biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 371-383.
- Hornstein, S.L. & Mulligan, N.W. (2001). Memory for action events: The role of objects in memory of self- and other-performed tasks. *American Journal of Psychology*, *114*, 199-217.
- Hornstein, S.L. & Mulligan, N.W. (2004). Memory for actions. Enactment and source memory. *Psychonomic Bulletin and Review*, *11*, 367-372.
- Hostetter, A.B. & Alibali, M.W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin and Review*, *15*, 495-514.
- Hussy, W. & Jain, A. (2002). *Experimentelle Hypothesenprüfung in der Psychologie*. Göttingen: Hogrefe.
- Hyman jr., I.E. & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language*, *35*, 101-117.
- Jacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G. & Mazziotta, J.C. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*, 0529-0535.
- Jacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C. & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526-2528.
- Jacoby, L.L. & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306-340.
- Jacoby, L.L. & Hay, J.F. (1998). Age-related deficits in memory: Theory and application. In M.A. Conway, S.E. Gathercole & C. Cornoldi (Eds.), *Theories of Memory* (111-134). Hove: Psychology Press.
- Jacoby, L.L., Kelley, C., Brown, J. & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*, 326-338.
- Jacoby, L.L., Woloshyn, V. & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, *118*, 115-125.
- James, W. (1890). *The Principles of Psychology*. New York: MaxMillan.

- Janowsky, J.S., Shimamura, A.P. & Squire, L.R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, *27*, 1043-1056.
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage*, *14*, 103-109.
- Jeannerod, M. & Pacherie, E. (2004). Agency, simulation and self-identification. *Mind and Language*, *19*, 113-146.
- Jennings, J.M. & Jacoby, L.L. (1997). An opposition procedure for detecting age-related deficits in recollecting: Telling effects of repetition. *Psychology and Aging*, *12*, 352-361.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201-211.
- Johnson, M.K. (2006). Memory and reality. *American Psychologist*, *61*, 760-771.
- Johnson, M.K., Foley, M.A. & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory and Cognition*, *16*, 337-342.
- Johnson, M.K., Foley, M.A., Suengas, A.G. & Raye, C.L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371-376.
- Johnson, M.K., Hashtroudi, S. & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.
- Johnson, M.K., Kounios, J. & Nolde, S.F. (1997a). Electrophysiological brain activity and memory source monitoring. *NeuroReport*, *8*, 1317-1320.
- Johnson, M.K., O'Connor, M. & Cantor, J. (1997b). Confabulation, memory deficits and frontal dysfunction. *Brain and Cognition*, *34*, 189-206.
- Johnson, M.K. & Raye, C.L. (1981). Reality monitoring. *Psychological Review*, *88*, 67-85.
- Johnson, M.K., Raye, C.L., Wang, A.Y. & Taylor, T.H. (1979). Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 229-240.
- Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (201-208). New York: Cambridge University Press.
- Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A.P., Brand, M. & Bullock, R. (2004). DemTect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal of Geriatric Psychiatry*, *19*, 136-143.
- Kausler, D.H. & Puckett, J.M. (1981). Adult age differences in memory for sex of voice. *Journal of Gerontology*, *36*, 44-50.
- Keller, J., Bohner, G. & Erb, H.-P. (2000). Intuitive und heuristische Urteilsbildung – verschiedene Prozesse? Präsentation einer deutschen Fassung des „Rational-Experiential-Inventory“ sowie neuer Selbstberichtskaalen zur Heuristiknutzung. *Zeitschrift für Sozialpsychologie*, *31*, 87-101.
- Kensinger, E.A. & Schacter, D.L. (2006). Neural processes underlying memory attribution on a reality-monitoring task. *Cerebral Cortex*, *16*, 1126-1133.
- Kilner, J.M., Paulignan, Y. & Blakemore, S.J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*, 522-525.
- Knuf, L., Aschersleben, G. & Prinz, W. (2001). An analysis of ideomotor action. *Journal of Experimental Psychology: General*, *113*, 779-798.
- Korda, M. (1997, 10/06/97). Prompting the president. *New Yorker*, *73*, 88-95.

- Kormi-Nouri, R. (1995). The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology, 7*, 337-363.
- Kosslyn, S.M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review, 88*, 46-66.
- Kosslyn, S.M. (2005). Mental images and the brain. *Cognitive Neuropsychology, 22*, 333-347.
- Kosslyn, S.M., Behrmann, M. & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia, 33*, 1335-1344.
- Kosslyn, S.M., Thompson W.L. & Ganis, G. (2006). *The Case of Mental Imagery*. Oxford: University Press.
- Landau, J.D. & von Glahn, N. (2004). Warnings reduce the magnitude of the imagination inflation effect. *American Journal of Psychology, 117*, 579-593.
- Lampinen, J.M., Odegard, T.N. & Bullington, J.L. (2003). Qualities of memories for performed and imagined actions. *Applied Cognitive Psychology, 17*, 881-893.
- Lane, S.M. & Zaragoza, M.S. (2007). A little elaboration goes a long way: The role of generation on eyewitness suggestibility. *Memory and Cognition, 35*, 1255-1266.
- Léonard, G. & Tremblay, F. (2007). Corticomotor facilitation associated with observation, imagery and imitation of hand actions: A comparative study in young and old adults. *Experimental Brain Research, 177*, 167-175.
- Libby, L.K. (2003). Imagery perspective and source monitoring in imagination inflation. *Memory and Cognition, 31*, 1072-1081.
- Light, L.L., LaVoie, D., Valencia-Laver, D., Owens, S.A.A. & Mead, G. (1992). Direct and indirect measures of memory for modality in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1284-1297.
- Lindsay, D.S. & Johnson, M.K. (1989). The eyewitness suggestibility effect and memory for source. *Memory and Cognition, 17*, 349-358.
- Loftus, E.F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560-572.
- Loftus, E.F., Miller, D.G. & Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19-31.
- Loftus, E.F. & Palmer, J.C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585-589.
- Lotze, H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmannsche Buchhandlung.
- Luo, C-R. (1993). Enhanced feeling of recognition: Effect of identifying and manipulating test items on recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 405-413.
- Maeda, F., Kleiner-Fisman, G. & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: Specificity of the effect and role of observer's orientation. *Journal of Neurophysiology, 87*, 1329-1335.
- Mäntylä, T. (1993). Knowing but not remembering: Adult age differences in recollective experience. *Memory and Cognition, 21*, 379-388.
- Mammarella, N. (2007). The role of spaced repetitions in imagination inflation. *European Journal of Cognitive Psychology, 19*, 882-897.
- Mammarella, N., Fairfield, B. & Cornoldi, C. (2007). Reality monitoring and resistance to forgetting under short delay intervals. *The Quarterly Journal of Experimental Psychology, 60*, 551-570.

- Marks, D.J. (1973). Visual imagery in the recall of pictures. *British Journal of Psychology*, *64*, 17-24.
- Masumoto, K., Yamaguchi, M., Sutani, K., Tsuneto, S., Fujita, A. & Tonoike, M. (2006). Reactivation of physical motor information in the memory of action events. *Brain Research*, *1101*, 102.
- Mayr, U. (2003). Normales kognitives Altern. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (715-726). Heidelberg: Springer.
- Mazzoni, G., Loftus, E.F. & Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: A little plausibility goes a long way. *Journal of Experimental Psychology: Applied*, *7*, 51-59.
- Mazzoni, G. & Memmon, A. (2003). Imagination can create false autobiographical memories. *Psychological Science*, *14*, 186-188.
- McDaniel, M.A., Butler, K.M. & Dornburg, C.C. (2006). Binding of source and content: New directions revealed by neuropsychological and age-related effects. In H.D. Zimmer, A. Mecklinger & U. Lindenberger (Hrsg.), *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience* (595-626). Oxford: University Press.
- McDaniel, M.A. & Einstein, G.O. (1986). Bizarre imagery as effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 54-65
- McDaniel, M.A., Einstein, G.O., DeLosh, E.L., May, C.P. & Brady, P. (1995). The bizarreness effect: It's not surprising, it's complex. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 422-435.
- McDaniel, M.A., Lyle, K.B., Butler, K.M. & Dornburg, C.C. (2008). Age-related deficits in reality monitoring of action memories. *Psychology and Aging*, *23*, 646-656.
- McDermott, K.B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212-230.
- McDermott, K.B. & Roediger, H.L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language*, *39*, 508-520.
- McIntyre, J.S. & Craik, F.I.M. (1987). Age differences in memory for item and source information. *Canadian Journal of Psychology*, *41*, 175-192.
- Meister, I.G., Krings, T., Foltys, H., Müller, M., Töpper, R. & Thron, A. (2004). Playing piano in the mind – an fMRI-study on music imagery and performance in pianists. *Cognitive Brain Research*, *19*, 219-228.
- Mitchell, K.J. & Johnson, M.K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford Handbook of Memory* (179-195). Oxford: University Press.
- Mitchell, K.J., Johnson, M.K. & Mather, M. (2003). Source-monitoring and suggestibility to misinformation: Adult age-related differences. *Applied Cognitive Psychology*, *17*, 107-119.
- Morcom, A.M., Li, J. & Rugg, M.D. (2007). Age effects on the neural correlates of episodic retrieval: Increased cortical recruitment with matched performance. *Cerebral Cortex*, *17*, 2491-2506.
- Mori, M. (1982). *The Buddha in the Robot*. Boston: Tuttle.
- Müsseler, J. (2002). Visuelle Wahrnehmung. In J. Müsseler & W. Prinz (Hrsg.), *Allgemeine Psychologie* (15-65). Heidelberg, Berlin: Spektrum.
- Murnane, K. & Bayen, U.J. (1996). An evaluation of empirical measures of source identification. *Memory and Cognition*, *24*, 417-428.

- Murnane, K. & Bayen, U.J. (1998) Measuring memory for source: Some theoretical assumptions and technical limitations. *Memory and Cognition*, *26*, 674-677.
- Mutz, G. (2002). *Variotest: Eine Entwicklungsumgebung für psychologische Testverfahren* [Computerprogramm]. Universität zu Köln: Unveröffentlichtes Manuskript.
- Nachtigall, C. & Wirtz, M. (1998). *Wahrscheinlichkeitsrechnung und Inferenzstatistik. Statistische Methoden für Psychologen, Teil 2*. Weinheim: Juventa.
- Naveh-Benjamin, M. (2006). Binding of memories: Adult-age differences and the effects of divided attention in young adults on episodic memory. In H.D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience* (627-656). Oxford: University Press.
- Naveh-Benjamin, M. & Craik, F.I.M (1995). Memory for context and its use in item memory: Comparisons of younger and older persons. *Psychology and Aging*, *10*, 284-293.
- Naveh-Benjamin, M., Craik, F.I.M., Gavrilesco, D. & Anderson, N.D. (2000). Asymmetry between encoding and retrieval processes: Evidence from divided attention and a calibration analysis. *Memory and Cognition*, *28*, 965-976.
- Nelson, D.L., Reed, U.S. & Walling, J.R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 523-528.
- Neuschatz, J.S., Benoit, G.E. & Payne, D.G. (2003). Effective warnings in the Deese-Roediger-McDermott false-memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 35-40.
- Nilsson, L.-G. (2000). Remembering actions and words. In E. Tulving & F.I.M. Craik (Eds.). *The Oxford Handbook of Memory* (137-148). Oxford: University Press.
- Nilsson, L.-G. & Craik, F.I.M. (1990). Additive and interactive effects in memory for subject-performed tasks. *European Journal of Cognitive Psychology*, *2*, 305-324.
- Nilsson, L.-G., Nyberg, L., Klingberg, T., Åberg, C., Persson, J. & Roland, P.E. (2000). Activity in motor areas while remembering action events. *NeuroReport*, *11*, 2199-2201.
- Nolde, S.F., Johnson, M.K. & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: An event-related fMRI study. *NeuroReport*, *9*, 3509-3514.
- Norman, K.A. & Schacter, D.L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition*, *25*, 838-848.
- Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U. & Nilsson, L.-G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *Journal of Gerontology: Psychological Science*, *51B*, 234-240.
- Nyberg, L. & Nilsson, L.-G. (1995). The role of enactment in implicit and explicit memory. *Psychological Research*, *57*, 215-219.
- Nyberg, L., Nilsson, L.-G. & Bäckman, L. (1992). Recall of actions, sentences, and nouns: Influences of adult age and passage of time. *Acta Psychologica*, *79*, 245-254.
- Nyberg, L., Petersson, K.M., Nilsson, L.-G., Sandblom, J.A. & Ingvar, C.M. (2001). Reactivation of motor brain areas during explicit memory for actions. *NeuroImage*, *2*, 521-528.
- Nyberg, L., Persson, J. & Nilsson, L.-G. (2002). Individual differences in memory enhancement by encoding enactment: Relationships to adult age and biological factors. *Neuroscience and Biobehavioral Reviews*, *26*, 835-839.
- Oberman, L.M., Hubbard, E.M., McCleery, J.P., Altschuler, E.L., Ramachandran, V.S. & Pineda, J.A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research*, *24*, 190-198.
- O'Connor, K.P. & Aardema, F. (2005). The imagination: Cognitive, pre-cognitive, and meta-cognitive aspects. *Consciousness and Cognition*, *14*, 233-56.

- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*, 241-263.
- Parkin, A.J. & Walter, B.M. (1992). Recollective experience, normal aging, and frontal dysfunction. *Psychology and Aging*, *7*, 290-298.
- Perani, D., Fazio, F., Borghese, A., Tettamanti, M., Ferrari, S., Lomonaco, F., Gilardi, M.C. & Decety, J. (2001). How the brain works for watching at real and virtual (hand) action. *NeuroImage*. [Angabe nach Grèzes & Decety, 2001]
- Pezdek, K., Blandon-Gitlin, I. & Gabbay, P. (2006). Imagination and memory: Does imagining implausible events lead to false autobiographical memories? *Psychonomic Bulletin and Review*, *13*, 764-769.
- Pezdek, K., Blandon-Gitlin, I., Lam, S., Hart, R.E. & Schooler, J.W. (2006). Is knowing believing? The role of event plausibility and background knowledge in planting false beliefs about the personal past. *Memory and Cognition*, *34*, 1628-1635.
- Pezdek, K. & Eddy, R.M. (2001). Imagination inflation: A statistical artefact of regression towards the mean. *Memory and Cognition*, *29*, 707-718.
- Prinz, W. (2002). Experimental approaches to imitation. In A.N. Meltzoff & W. Prinz (Eds.), *The Imitative Mind. Development, Evolution, and Brain Bases* (143-162). Cambridge: University Press.
- Pylyshyn, Z.W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*, 16-45.
- Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385-401.
- Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F.I.M. Craik & T.A. Salthouse (Eds.), *Handbook of Aging and Cognition* (1-90). Mahwah: Erlbaum.
- Rizzolatti, L. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169-192.
- Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1996a). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131-141.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Perani, D. & Fazio, F. (1996b). Localization of cortical areas responsive to the observation of hand grasping movements in humans: A PET study. *Experimental Brain Research*, *111*, 246-252.
- Rizzolatti, G., Fogassi, L. & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*, 661-670.
- Rizzolatti, G. & Sinigaglia, C. (2008). *Mirrors in the Brain – How our Minds Share Actions and Emotions*. Oxford: University Press.
- Roediger, H.L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76-100.
- Roediger, H.L. & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803-814.
- Roediger, H.L. & McDermott, K.B. (2000). Distortions of memory. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford Handbook of Memory* (149-162). Oxford: University Press.
- Rönnlund, M., Nyberg, L., Bäckman, L. & Nilsson, L.-G. (2003). Recall of subject-performed tasks, verbal tasks, and cognitive activities across the adult life span: Parallel age-related deficits. *Aging, Neuropsychology, and Cognition*, *10*, 182-201.
- Ross, B.H., Wang, R.F., Kramer, A.F., Simons, D.J. & Crowell, J.A. (2007). Action information from classification learning. *Psychonomic Bulletin and Review*, *14*, 500-504.

- Salthouse, T.A. (1996a). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403-428.
- Salthouse, T.A. (1996b). General and specific speed mediation of adult age differences in memory. *Journal of Gerontology: Psychological Sciences*, *51B*, P30-P42.
- Schacter, D.L., Harbluk, J. & McLachlan, D. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior*, *23*, 593-611.
- Schacter, D.L., Koutstaal, W., Johnson, M.K., Gross, M.S. & Angell, K.E. (1997). False recollection induced by photographs: A comparison of older and younger adults. *Psychology and Aging*, *12*, 203-215.
- Schacter, D.L., Verfaellie, M. & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, *35*, 319-334.
- Schneider, W. & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1-66.
- Schütz-Bosbach, S., Mancini, B., Aglioti, S.M. & Haggard, P. (2006). Self and other in the human motor system. *Current Biology*, *16*, 1830-1834.
- Seamon, J.G., Philbin, M.M. & Harrison, L.G. (2006). Do you remember proposing marriage to the Pepsi machine? False recollections from a campus walk. *Psychonomic Bulletin and Review*, *13*, 752-756.
- Senkfor, A.J. (2008). Memory for pantomimed actions versus actions with real objects. *Cortex*, *44*, 820-833.
- Senkfor, A.J. & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1005-1025.
- Senkfor, A.J., Van Petten, C. & Kutas, M. (2002). Episodic action memory for real objects: An ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. *Journal of Cognitive Neuroscience*, *14*, 402-419.
- Senkfor, A.J., Van Petten, C. & Kutas, M. (2008). Enactment versus conceptual encoding: Equivalent item memory but different source memory. *Cortex*, *44*, 649-664.
- Sharman, S.J. & Barnier, A.J. (2008). Imagining nice and nasty events in childhood or adulthood: Recent positive events show the most imagination inflation. *Acta Psychologica*, *129*, 228-233.
- Sharman, S.J., Garry, M. & Beuke, C.J. (2004). Imagination or exposure causes imagination inflation. *American Journal of Psychology*, *117*, 157-168.
- Sharman, S.J., Garry, M. & Hunt, M. (2005a). Using source cues and familiarity cues to resist imagination inflation. *Acta Psychologica*, *120*, 227-242.
- Sharman, S.J., Manning, C.G. & Garry, M. (2005b). Explain this: Explaining childhood events inflates confidence for those events. *Applied Cognitive Psychology*, *19*, 67-74.
- Shepard, R.N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701-703.
- Shiffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-190.
- Shimamura, A.P. & Squire, L.R. (1987). A neuropsychological study of fact memory and source amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 464-473.

- Slamecka, N.J. & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, 4, 592-604.
- Spencer, W.D. & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, 10, 527-539.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74 (Whole No. 498).
- Spitzer, C., Stieglitz, R.-D. & Freyberger, H.-J. (2005). *Fragebogen zu Dissoziativen Symptomen (FDS)*. Göttingen: Hogrefe.
- Städtler, T. (2003). *Lexikon der Psychologie*. Stuttgart: Kröner.
- Stark, L.-J. & Perfect, T.J. (2006). Elaboration inflation: How your ideas become mine. *Applied Cognitive Psychology*, 20, 641-648.
- Stefan, K., Classen, J., Celnik, P. & Cohen, L.G. (2008). Concurrent action observation modulates practice-induced motor memory information. *European Journal of Neuroscience*, 27, 730-738.
- Stefan, K., Cohen, L.G., Duque, J., Mazzocchio, R., Celnik, P., Sawaki, L., Ungerleider, L. & Classen, J. (2005). Formation of a motor memory by action observation. *The Journal of Neuroscience*, 25, 9339-9346.
- Stephan, K.M., Fink, G.R., Passingham, R.E., Silbersweig, D., Ceballos-Baumann, A.O., Frith, C.D. & Frackowiak, R.S.J. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology*, 73, 373-386.
- Stern, E.R. & Rotello, C.M. (2000). Memory characteristics of recently imagined events and real events experienced previously. *American Journal of Psychology*, 113, 569-590.
- Thomas, A.K. & Bulevich, J.B. (2006). Effective cue utilization reduces memory errors in older adults. *Psychology and Aging*, 21, 379-389.
- Thomas, A.K., Bulevich, J.B. & Loftus, E.F. (2003). Exploring the role of repetition and sensory elaboration in the imagination inflation effect. *Memory and Cognition*, 31, 630-640.
- Thomas, A.K. & Loftus, E.F. (2002). Creating bizarre false memories through imagination. *Memory and Cognition*, 30, 423-431.
- Thomas, A.K. & Sommers, M.S. (2005). Attention to item-specific processing eliminates age effects in false memories. *Journal of Memory and Language*, 52, 71-86.
- Tye, M. (1991). *The Imagery Debate*. Cambridge, MA: MIT Press.
- Underwood, B.J. & Freund, J.S. (1970). Testing effects in the recognition of words. *Journal of Verbal Learning and Verbal Behavior*, 9, 117-125.
- Vogele, K., May, M., Ritzl, A., Falkai, P., Zilles, K. & Fink, G.R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience*, 16, 817-827.
- Watson, J.M., McDermott, K.B. & Balota, D.A. (2004). Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory and Cognition*, 32, 135-141.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised Manual*. New York: Psychological Corp.
- Wechsler, D. (1987). *Wechsler Memory Scale – Revised Manual*. New York: Psychological Corp.
- West, R. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, 120, 272-292.
- Westerman, D.L. & Greene, R.L. (1996). On the generality of the revelation effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1147-1153.

- Whitten, W.B. & Bjork, R.A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning & Verbal Behavior*, 16, 465-478.
- Whittlesea, B.W.A. & Leboe, J.P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, 49, 62-79.
- Whittlesea, B.W.A. & Price, J.R. (2001). Implicit/ explicit memory versus analytic/ nonanalytic processing: Rethinking the mere exposure effect. *Memory and Cognition*, 29, 234-246.
- Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin and Review*, 8, 44-57.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9, 625-636.
- Wilson, M. & Knoblich, G. (2005). The case of motor involvement in perceiving conspecifics, *Psychological Bulletin*, 131, 460-473.
- Wohlschläger, A. & Prinz, W. (2006). Wahrnehmung. In H. Spada (Hrsg.), *Allgemeine Psychologie* (25-114). Bern: Huber.
- Wolpert, D.M., Ghahramani, Z. & Jordan, M.I. (1995). An internal model of sensorimotor integration. *Science*, 269, 1880-1882.
- Zaragoza, M.S. & Lane S.M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 934-945.
- Zaragoza, M.S. & Lane S.M. (1998). Processing resources and eyewitness suggestibility. *Legal and Criminological Psychology*, 3, 305-320.
- Zaworka, W., Hand, I., Jauernig, G. & Lünenschloß, K. (1983). *Hamburger Zwangsinventar (HZI). Fragebogen zur Erfassung von Zwangsgedanken und Zwangsverhalten*. Göttingen: Hogrefe.
- Zimmer, H.D. & Engelkamp, J. (1985). An attempt to distinguish between kinematic and motor memory components. *Acta Psychologica*, 58, 81-106.

VI ABILDUNGS- UND TABELLENVERZEICHNIS

Abbildungsverzeichnis

Abb.-Nr.	Titel	Seite
Abb. 1	Zeichnung von Saul Steinberg (abgeb. nach Aronson, 1994, S. 348).	8
Abb. 2	Gehirnaktivität beim cued-recall-Test. Ausführung kontrastiert mit verbaler Enkodierung in der Studie von Nyberg et al. (2001). Die Aufgabe war vergleichbar mit derjenigen von Nilsson et al. (2000). Erhöhte Aktivität fand sich im kontralateralen (linken) motorischen/ somatosensorischen, prämotorischen, inferioren und superioren parietalen Cortex (abgeb. nach Nyberg et al., 2002, S. 836).	29
Abb. 3	Antwort eines Spiegelneurons für Greifen auf die Beobachtung des greifenden Experimentators (A) bzw. auf den selbst ausgeführten Greifakt (B) (di Pellegrino et al., 1992; abgeb. nach Rizzolatti & Sinigaglia, 2008, S. 81).	33
Abb. 4	Das menschliche Spiegelneuronensystem. Rot und gelb: Teile des Parietal- bzw. Frontallappens, die während der Ausführung und Beobachtung aktiviert sind. Manchmal wird auch der dorsale Anteil der Area 6 zum Spiegelneuronensystem gezählt, der allerdings möglicherweise eher zum Zwecke der Handlungsvorbereitung aktiviert ist. Blau: Teil des Frontallappens, der in einigen experimentellen Bedingungen bei der Beobachtung aktiviert war; hierbei könnte es sich möglicherweise aber auch um andere als spiegelneuronale Aktivität handeln (abgeb. nach Rizzolatti & Sinigaglia, 2008, S. 120).	34
Abb. 5	Lateralansicht der Aktivationsfoki unter den verschiedenen Verarbeitungsbedingungen in der Metaanalyse von Grèzes und Decety (2001). Gelb: Ausführung. Grün: Vorstellung. Blau: Beobachtung. Rot: mentale Verbalisierung. Die Zahlen entsprechen den in die Metaanalyse einbezogenen Studien (abgeb. nach Grèzes & Decety, 2001, S. 5).	35
Abb. 6	Ausschnitte aus den Videos zum ersten Experiment. Die Handlungsausführung erfolgte durch eine Akteurin (links) bzw. einen Akteur (rechts).	69
Abb. 7	Ablauf der ersten Phase des ersten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.	74
Abb. 8	Ablauf der zweiten Phase des ersten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.	76
Abb. 9	Ablauf der dritten Phase des ersten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung. Rekog.: Rekognitionstest. Quell.: Quellengedächtnistest. Sicherh.: Sicherheitsrating.	77
Abb. 10	Ablauf der ersten Phase des zweiten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.	120
Abb. 11	Ablauf der zweiten Phase des zweiten Experiments. S-g: selbst-gesteuert, d.h. ohne Zeitbegrenzung.	121
Abb. 12	Ausschnitte aus den Videos zum dritten Experiment. Die Handlungsausführung wurde aus der 1.- (links) bzw. 2.-Person-Perspektive (rechts) gezeigt.	152
Abb. 13	Ausmaß des enactment-Effekts in verschiedenen Altersgruppen, erfasst anhand des freien Abrufs (nach Nyberg et al., 1996; abgeb. nach Nyberg et al., 2002, S. 836).	180

Tabellenverzeichnis

Tab.-Nr.	Titel	Seite
Tab. 1	Ablauf der Untersuchung von Garry, Manning, Loftus und Sherman (1996).	16
Tab. 2	Ablauf der Untersuchung von Goff und Roediger (1998).	22
Tab. 3	Ergebnisse des ersten Experiments von Thomas, Bulevich und Loftus (2003). <i>Ausgeführt</i> -Antworten und <i>remember-/ know</i> -Urteile im Quellengedächtnistest (abgeb. nach Thomas et al., 2003, S. 633).	24
Tab. 4	Ergebnisse des zweiten Experiments von Thomas, Bulevich und Loftus (2003). <i>Ausgeführt</i> -Antworten und <i>remember-/ know</i> -Urteile im Quellengedächtnistest (abgeb. nach Thomas et al., 2003, S. 635).	24
Tab. 5	Charakterisierung der Prozesse Ausführen, Vorstellen, Beobachten, Generieren und Lesen hinsichtlich der vier Dimensionen Vertrautheit, Informationsursprung, sensorische und motorische Eindrücke.	46
Tab. 6	Ablauf des ersten Experiments.	54
Tab. 7	Versuchsplan des ersten Experiments.	57
Tab. 8	Materialpool des ersten Experiments.	66
Tab. 9	Materialkombinationen des ersten Experiments.	67
Tab. 10	Regeln zur Erstellung umgestellter Sätze des ersten Experiments.	68
Tab. 11	Darbietung der intraindividuell variierten unabhängigen und der abhängigen Variablen des ersten Experiments (am Beispiel der Materialkombination 1).	70
Tab. 12	Beispiel einer fixierten Zufallsreihenfolge des ersten Experiments (Materialkombination 1).	71
Tab. 13	Relative Häufigkeiten von <i>Phase 1</i> -Antworten im Rekognitionstest des ersten Experiments.	79
Tab. 14	Relative Häufigkeiten von <i>ausgeführt</i> -Antworten im Quellengedächtnistest des ersten Experiments.	80
Tab. 15	Sicherheitsratings (siebenstufige Skala) bzgl. falscher <i>ausgeführt</i> -Antworten im ersten Experiment.	88
Tab. 16	Falsche <i>ausgeführt</i> -Antworten in Abhängigkeit vom Geschlecht der Akteurin/ des Akteurs bzw. der Beobachterin/ des Beobachters im ersten Experiment.	89
Tab. 17	Darbietung der intraindividuell variierten unabhängigen und der abhängigen Variablen der Voruntersuchung zum zweiten Experiment (am Beispiel der Materialkombination 1).	106
Tab. 18	Relative Häufigkeiten von <i>alt-</i> bzw. <i>ausgeführt</i> -Antworten im Rekognitions- bzw. Quellengedächtnistest der Voruntersuchung zum zweiten Experiment.	107
Tab. 19	Einschätzungen verschiedener Items (siebenstufige Skala) in der Voruntersuchung zum zweiten Experiment.	109
Tab. 20	Ablauf des zweiten Experiments.	112
Tab. 21	Versuchsplan des zweiten Experiments.	115
Tab. 22	Relative Häufigkeiten von <i>alt</i> -Antworten im Rekognitionstest des zweiten Experiments.	124

Tab.-Nr.	Titel	Seite
Tab. 23	Relative Häufigkeiten von <i>ausgeführt</i> -Antworten im Quellengedächtnistest des zweiten Experiments.	125
Tab. 24	Sicherheitsratings (siebenstufige Skala) bzgl. falscher <i>ausgeführt</i> -Antworten im zweiten Experiment.	129
Tab. 25	Einschätzung verschiedener Items (siebenstufige Skala) im zweiten Experiment.	131
Tab. 26	Charakterisierung der Untersuchungsbedingungen des dritten Experiments.	146
Tab. 27	Ablauf des dritten Experiments.	147
Tab. 28	Versuchsplan des dritten Experiments.	150
Tab. 29	Relative Häufigkeiten von <i>alt</i> -Antworten im Rekognitionstest des dritten Experiments.	154
Tab. 30	Relative Häufigkeiten von <i>ausgeführt</i> -Antworten im Quellengedächtnistest des dritten Experiments.	155
Tab. 31	Sicherheitsratings (siebenstufige Skala) bzgl. falscher <i>ausgeführt</i> -Antworten im dritten Experiment.	158
Tab. 32	Einschätzung verschiedener Items (siebenstufige Skala) im dritten Experiment.	161/162
Tab. 33	Ergebnisse des ersten Experiments von Thomas und Bulevich (2006). <i>Ausgeführt</i> -Antworten im Quellengedächtnistest (abgeb. nach Thomas & Bulevich, 2006, S. 383).	181
Tab. 34	Ergebnisse des zweiten Experiments von Thomas und Bulevich (2006). <i>Ausgeführt</i> -Antworten im Quellengedächtnistest (abgeb. nach Thomas & Bulevich, 2006, S. 385).	182
Tab. 35	Versuchsplan des vierten Experiments.	193
Tab. 36	Ablauf des vierten Experiments.	199
Tab. 37	Relative Häufigkeiten von <i>alt</i> -Antworten im Rekognitionstest des vierten Experiments.	200
Tab. 38	Relative Häufigkeiten von <i>ausgeführt</i> -Antworten im Quellengedächtnistest des vierten Experiments.	201
Tab. 39	Sicherheitsratings (siebenstufigen Skala) bzgl. falscher <i>ausgeführt</i> -Antworten im vierten Experiment.	206
Tab. 40	Einschätzung verschiedener Items (siebenstufige Skala) im vierten Experiment.	208
Tab. 41	Korrelationen von Leistungen in Testverfahren zur Erfassung sog. frontaler bzw. medial-temporalen Funktionen und dem observation-inflation-Effekt für falsche bzw. korrekte <i>ausgeführt</i> -Antworten im vierten Experiment.	209

VII ANHANG

A Materialien zum ersten Experiment

A.1 Vertrag zwischen Untersuchungsleiterin und UntersuchungsteilnehmerInnen

Liebe Teilnehmerin, lieber Teilnehmer,

Sie haben sich freiwillig dafür entschieden, an dieser Untersuchung teilzunehmen. Ich danke Ihnen sehr für Ihre Bereitschaft und versichere, dass Ihnen in diesem Zusammenhang keine übermäßigen seelischen und körperlichen Belastungen zugemutet werden, denn *Ihr Wohlergehen hat selbstverständlich Vorrang vor reinem wissenschaftlichem Interesse.*

Leider ist es aus methodischen Gründen manchmal nicht möglich, TeilnehmerInnen vor einer Untersuchung vollständig über deren Ziele und genaue Durchführung zu informieren. *Sofort nach Abschluss der Studie werden Sie jedoch alle gewünschten Informationen erhalten.*

Sollten entgegen meiner Erwartungen und bisherigen Erfahrungen für Sie während der Untersuchung Belastungen auftreten, die Sie als zu schwerwiegend erachten, so haben Sie selbstverständlich *jederzeit die Möglichkeit, die Untersuchung abzubrechen. Es werden für Sie daraus keinerlei nachteilige Folgen entstehen.*

Diesen Rechten stehen auch nachvollziehbare Pflichten gegenüber. Die Planung und Durchführung einer Untersuchung sind mit erheblichen zeitlichen und finanziellen Aufwendungen verknüpft. Es ist deshalb wichtig, dass Sie versuchen, die *Aufgabenstellungen der Untersuchung, so wie Sie Ihnen durch die Anweisungen vermittelt werden, so gut wie möglich zu erfüllen.* Die Untersuchung bedarf Ihrer *vollen Konzentration.* Eine uninteressierte und oberflächliche Mitarbeit gefährdet die Erreichung der Untersuchungsziele erheblich und bedeutet damit gleichzeitig eine mögliche Verschwendung von zeitlichen und finanziellen Mitteln. Darüber hinaus möchte ich Sie bitten, vor Abschluss der Studie keine konkreten Inhalte der Untersuchung sowie Informationen über Ziele und Ablauf der Untersuchung weiterzugeben.

Ich möchte Sie abschließend darauf hinweisen, dass Ihre *Untersuchungsergebnisse selbstverständlich streng vertraulich behandelt* werden. Zu diesem Zweck vereinbare ich mit Ihnen ein Codewort, das die Zuordnung der Daten zu Ihrer Person verhindert.

Vielen Dank für Ihre ehrliche und bereitwillige Mitarbeit!!!

Ich habe den Vertrag gelesen und verstanden und erkläre mein Einverständnis.

Ort, Datum: _____ Unterschrift: _____

A.2 Umgestellte Sätze

Würfeln Sie mit den Würfeln!	<p>Würfeln0Sie4den8mit1Würfeln! Würfel_mitSie_ürfelnden! Würf_InSWürfelnJm_tOdenPSie! Si_1W_rfelndWürfeln5d_nOmit! m_t sie w_rf_In w_rf_In d_n!</p>
Schütteln Sie die Flasche!	<p>Sie3Schütteln9die5Flasche! Flas_hedieSieSchüt_eln! d_eMSieWSch_ttelnAFlasche! S_eMSchütt_In7Flasch_Gdie! d__fl_sche sch_tt_In s_e!</p>
Wickeln Sie das Band auf!	<p>Sie7Wickeln2das8Band0auf! Ba_dSieaufWickelnda_! Wick_InXBandPdaskS_eEauf! d_s6BandTW_ckeln9SieK_uf! a_f b_nd das si_ wick_In!</p>
Falten Sie das Blatt Papier!	<p>Falten5Sie0das9Papier6Blatt! da_BlattFa_tenSiePapier! Bl_ttOSieVd_sBFaltenNPapier! Pap_erOBlatt2d_sGSie3Falt_n! d_s blatt f_lt_n s__p_pier!</p>
Betätigen Sie die Fernbedienung!	<p>Betätigen2die1Sie0Fernbedienung! dieFer_bedienungBetätigen_ie! B_tätigenGFernbedienungHdieLS_e! F_rnbedienung2d_eRSie6Bet_tigen! si_b_tät_gen fernb_dien_ng di_!</p>
Mischen Sie die Karten!	<p>Mischen9Sie3Karten8die! dieKar_enMische_Sie! dieOS_eNMischenIK_rten! Si_LM_schen7K_rtenMdie! k_rt_n di_s_e m_sch_n!</p>
Tun Sie das Ei in den Eierbecher!	<p>Tun7Sie5das0in2Ei1den4Eierbecher! Eier_echerSieinEidasde_Tun! T_nYdenSdasWEiUinOSieKEi_rbecher! Ei8Eierb_cherPdas4T_nBin1denWS_e! d_s s__ tun ei e_erb_cher den _n!</p>
Ziehen Sie die Socke auf links!	<p>Sie8Ziehen2die3Socke5auf9links! So_keSielinksZiehenau_die! Zi_henDlinksHdieLS_ckeMaufGSie! a_fKSie5linksLSock_3ZiehenUd_e! l_nks s_ck_ die s_e auf z_eh_n!</p>

Spitzen Sie den Bleistift an!	Spitzen1Sie8den5an3Bleistift! de_anSpitze_BleistiftSie! _nQSiePdenSBl_istiftMSpitzen! Sp_tzen2Ble_stiftBan3S_eOden! _n bl_ist_ft d_n sie sp_tz_n!
Reißen Sie einen Zettel von dem Block ab!	Reißen3Sie6einen0von4Zettel7dem2Block9ab! einenSieRei_enBlockvonde_Zettelab! ReißenDBl_ckGeinenÜZettelLvonOd_mBSieVab! vonOSie6_inenXBlock4R_ißenQdem2Zett_IPab! einen sie r_iß_n z_tt_l ab d_m block v_n!
Legen Sie die Seife in die Dose!	Legen6Sie4Seife1die9in3die7Dose! dieSieLege_DoseindieSei_e! Leg_nXdieLdieESeifeRinOSiePD_se! Seif_9SieWdie7L_genEDose3d_eOin! in s_e d_s_seife l_g_n die di_!
Öffnen Sie die Sicherheitsnadel!	Öffnen3die9Sie6Sicherheitsnadel! Sicher_eitsnadeldie_ieÖffnen! _ffnenBSicherh_itsnadelHdieKSie! Si_NÖffnen4Sicherheitsn_delSd_e! di_sich_rheitsn_del öffn_n s_!
Lochen Sie das Papier!	Lochen5das9Sie0Papier! Pa_ierdasSieLo_hen! P_pierRSieEd_sTLochen! S_e7L_chenHPapi_r1das! d_s p_p_er l_ch_n s_e!
Biegen Sie eine Büroklammer auseinander!	Biegen0Sie3Büroklammer1eine8auseinander! eineausei_anderBiegenBürok_ammerSie! BiegenTaus_inanderZeineUBürokl_mmerISie! B_egenVBür_klammer3auseinanderMS_e5eine! b_rokl_mmer sie a_sein_nder bi_gen e_ne!
Nehmen Sie ein Streichholz aus der Schachtel!	Nehmen9Sie2ein3aus1Streichholz6der7Schachtel! ausSieei_NehmenSchachtelderStreich_olz! Nehm_nDSieKeinOderLausBStreichh_lzMSchachtel! ein5S_eONehmen3Stre_chholzWSchachtel7derDa_s! sch_chtel sie d_r stre_chh_lz aus ein_n_hm_n!
Falten Sie das Handtuch!	Falten4das6Sie3Handtuch! dasHa_dtuchFalten_ie! dasPS_eVFaltenNHandt_ch! SieÜF_lten4Handt_chTd_s! h_ndt_ch d_s s_e f_lt_n!
Dehnen Sie das Gummiband!	Dehnen6das3Sie0Gummiband! da_GummibandDeh_enSie! SiePD_hnenUGumm_bandOdas! d_s1G_mmibandOD_hnen8Sie! g_mmib_nd d_s s_e d_hn_n!

Zerschneiden Sie das Papier!	Sie5Zerschneiden7das2Papier! SieZer_chneidenPapier_as! dasFSi_MZerschneidenIPap_er! Papi_rTdas6S_eRZ_rschneiden! d_s_p_pi_r z_rschne_den si_!
Schalten Sie die Taschenlampe an!	Schalten2Sie8Taschenlampe5die6an! die_aschenlampeSchalten_jean! SchaltenATaschenl_mpeKdieGS_ePan! an2T_schenlampeMd_e3SiePSchalt_n! tasch_nlamp__n die sch_lt_n s_e!
Nehmen Sie den Deckel von der Schachtel!	Nehmen0Sie3den4Deckel6der2von9Schachtel! Ne_menvondenSchachte_SiederDeckel! Nehm_nÄderNdenPDeckelCv_nXSieÜSchachtel! Deck_lÜSie4derÖN_hmen8vonWden7Sch_chtel! sch_cht_l sie d_n der von d_ck_l nehm_n!
Betätigen Sie die Klingel!	Sie9Betätigen4die7Klingel! Kli_geldieSieBe_ätigen! dieOSieRB_tätigenPKling_l! S_e3Betätig_nMKlingel9d_e! d_e kl_ng_l bet_tig_n s_e!
Wickeln Sie ein Bonbon aus!	Wickeln1Sie5ein2aus8Bonbon! Bo_bonSieausWic_elnein! WickelnRS_eTeinUBonb_nEaus! e_nMBonb_n1WickelnNSie6a_s! a_s b_nb_n ein s_e w_ck_ln!
Stellen Sie das Teelicht in das Glas!	Stellen6das0Sie1Teelicht9in4das7Glas! Teelichtdasda_StelleninSieG_as! inUSieGd_sÖTeelichtWStell_nPdasXGlas! das4SieÖSt_llen8Te_lichtSGlas1da_Rin! st_ll_n das d_s glas _n sie t_el_cht!
Nehmen Sie ein Stück Tesaband von der Rolle!	Nehmen7Sie6ein5Stück3Tesaband2der4von1Rolle! Neh_enTesabandeinRolleSievonderStüc_! St_ckFSiePeinGNehmenSTes_bandAvonIderCRolle! v_n8StückZein1SieITesab_nd4NehmenMder6R_lle! n_hm_n sie tes_b_nd rolle e_n von der st_ck!
Schlagen Sie das Buch auf!	Schlagen3Sie8das9auf0Buch! da_BuchS_hlagenSieauf! B_chOSieIdasÄSchlagenR_uf! aufJBuch2d_sHSi_9Schl_gen! d_s __f schl_g_n b_ch sie!
Stoßen Sie das Spielzeugauto an!	Stoßen6Sie5das7an4Spielzeugauto! a_SieSpielzeu_autodasStoßen! dasPS_eFStoßenRSp_elzeugautoMan! Spi_lzeugauto5Si_Van8St_ßenÖdas! st_ßen sp_elze_gauto _n s__ das!

Tun Sie den Teebeutel in die Tasse!	Tun8Sie4den7in1Teebeutel3die2Tasse! inSieTa_seTeebeutelTun_ieden! T_nUSieXdenÖTasseDinBdieLTe_beutel! denBS_e1TunOTeeb_utel9Tasse2d_eCin! te_beut_l sie d_n tun t_ss_ die_n!
Stellen Sie die Uhr eine Stunde vor!	Stellen0Sie8die5eine6Uhr2Stunde9vor! UhrSiedieStel_envorStun_eeine! St_llenASTundeVdieOUhrP_ineRSieÜvor! v_r1SieRdie8St_ndeTeine4UhrMStell_n! st_nd_ ein_ die _hr sie st_ll_n vor!
Nehmen Sie eine Tintenpatrone aus der Schachtel!	Nehmen2Sie7eine4aus9Tintenpatrone0der5Schachtel! Tinte_patroneSieSchachtelNeh_enausdereine! Nehm_nETintenpatroneZein_DSieWausPderXSchachtel! Tintenp_troneGSie3derFN_hmen9ausJeine6Sch_chtel! nehmen a_s e_ne sch_cht_l sie der t_ntenpatr_ne!
Klappen Sie die Sonnenbrille zusammen!	Klappen4Sie3die0zusammen1Sonnenbrille! Son_enbrilleSiezusamme_Klappendie! KlappenOSieBdieÄS_nnenbrilleJz_sammen! die8Sonnenbr_lleGKI_ppen6SieOzusamm_n! zus_mmen s_nnenbrill_ die sie kl_pp_n!
Stempeln Sie auf das Papier!	Stempeln7Sie6auf4Papier9das! au_dasStempelnSie_apier! Stemp_lnRSieMPapierEd_sGauf! dasXP_pier4a_fgSt_mpeln3Sie! a_f_p_pi_r st_mp_ln d_s sie!
Legen Sie die Nudeln in die Tasche!	Sie8Legen6die4Nudeln2in7die0Tasche! i_SiedieNudelnLegendieT_sche! LegenHN_delnAdieRSieVTaschePdi_Lin! die4S_ekLegen3di_9inVNud_ln8Tasche! t_sch_ d_e die n_d_ln in sie l_gen!
Betätigen Sie das Feuerzeug!	Betätigen3das5Sie9Feuerzeug! da_FeuerzeugBe_ätigenSie! B_tätigenGFeuerzeugPd_sÖSie! S_eZBetätig_n2Feuerze_gWdas! fe_erz_ug d_s S_e b_tätig_n!
Öffnen Sie das Mäppchen!	Sie1Öffnen1das7Mäppchen! Mä_pchendas_ieÖffnen! d_sHSieBöffn_nNMäppchen! d_s6Mäppch_nRöffn_n5Sie! s_e_ffn_n m_ppch_n d_s!
Nehmen Sie ein Taschentuch heraus!	Nehmen5Sie1ein6heraus9Taschentuch! Herau_TaschentucheinSieNe_men! T_schentuchXSieVe_nKNehmenPher_us! Nehm_nHTasch_ntuch4h_rausRSie3ein! t_schent_ch her_us ein n_hm_n s_e!
Schrauben Sie den Deckel vom Deo!	Schrauben9Sie5den1vom2Deckel8Deo! Decke_SievomSchraube_denDeo! SchraubenID_ckelYdenÄSieLv_mODeo! Deckel0D_oWden4Schr_ubenXvom7S_e! d_n sie schra_b_n d_o vom d_ck_l!

Schneiden Sie ein Stück Garn von der Rolle ab!	Schneiden6Sie1Stück7ein9Garn2von3der5Rolle4ab! SchneidenSieei_abGarnvonderRol_eStück! RolleUS_eZeinXderIGarnPv_nEStückYSchne_denNab! Schneid_nLStück5Roll_CSie9GarnEvon7d_rDein4ab! schne_d_n von ein ab g_rn sie der r_ll_st_ck!
Zerreißen Sie das Papier!	Zerreißen0Sie2Papier8das! Pa_ierdasSieZer_eißen! dasCS_eHZerreißenUP_pier! Sie8Z_rreißenWPapi_r3d_s! d_s p_p_er zerr_iß_n s_e!
Nehmen Sie die CD aus der Hülle!	Nehmen2Sie9CD3die5aus8der6Hülle! Neh_enderHülleCDausSie_ie! NehmenFderXdi_KCDMausVSieEH_ile! CDRS_e4derLN_hmen9ausHdie7Hüll_! h_ll_sie d_e der a_s cd n_hm_n!
Schließen Sie das Schloss auf!	Sie1Schließen4das5Schloss7auf! Schlo_sSieaufSchließe_das! dasJS_eFSchließenCSchl_ssUauf! a_f1Schl_ssHdas5SieLSchli_ßen! schl_ss auf das schli_ß_n s_!
Werfen Sie den Ball von der einen in die andere Hand!	Werfen7Sie3den8Ball0von1der2in9einen4die5andere6Hand! Werfe_SiedeneinenvonderB_llindieHandandere! WerfenMSieÖH_ndIBallRvonWderUein_nAinEdieWanderePden! vonASie4denJB_ll3WerfenPd_r8einenRdie7inWand_re5Hand! h_nd sie den ball in der e_n_n v_n die andere w_rf_n!
Nehmen Sie einen Becher von dem Stapel!	Nehmen4Sie7Becher1einen3von5dem9Stapel! Be_herSieeinenNehmenStapelde_von! N_hmenPStapelAvonLBecherNein_nDdemTSie! von9SieZSt_pel1BecherVnehm_n5d_mOeinen! einen b_ch_r nehm_n sie v_n dem st_p_!
Entfernen Sie die Batterie aus dem Wecker!	Entfernen5Sie9die3aus0Batterie7dem1Wecker! BatterieSiedieE_tfernenWeckerde_aus! ausWS_eÜdieOBatterieIEntfernenUdemJWeck_r! dieÖS_e7EntfernenAW_cker0ausFdem2Batt_rie! entf_rn_n a_s die w_ck_r sie dem b_tterie!
Schrauben Sie den Kugelschreiber auf!	Schrauben3Sie8Kugelschreiber5den1auf! denKu_elschreiberSchraube_Sieauf! Schr_ubenXKugelschreiberMdenZS_eÖauf! Kug_lschreiber7SieDa_f4Schra_benIden! d_n auf schr_ub_n k_gelschr_iber s_e!
Stecken Sie die Karte in den Umschlag!	Stecken8Sie5die4in7Karte1den3Umschlag! Stecke_KartedieSieUm_chlagdenin! St_ckenVUmschlagIdieRK_rteÜinZdenPSie! inGUmschl_g8dieFKarte5Steck_nLden6S_e! k_rt_den d_e stecken_n sie _mschl_g!

A.3 Instruktionen

Phase 1

Liebe Untersuchungsteilnehmerin, lieber Untersuchungsteilnehmer, diese Untersuchung besteht aus mehreren Phasen, von denen Sie heute die ersten drei durchlaufen werden. Im Rahmen der ersten Untersuchungsphase, mit der Sie nun beginnen werden, werden Ihnen mehrere einfache Handlungsanweisungen präsentiert, wie z.B. 'Zeichnen Sie einen Kreis auf das Papier!'. Diese Anweisungen sollen Sie entweder lesen oder tatsächlich ausführen. Wie genau Sie in einem konkreten Fall vorgehen sollen, wird Ihnen jeweils zusammen mit der Handlungsanweisung mitgeteilt. Der Ablauf wird konkret so aussehen, dass Ihnen zuerst mitgeteilt wird, welche Utensilien im Zusammenhang mit einer folgenden Handlungsanweisung gebraucht werden könnten. Die erste Aufforderung könnte also z.B. lauten: 'Bitte holen: Stift und Blatt Papier'. Sie würden sich dann beide Utensilien heraussuchen und vor sich auf den Tisch legen. Sobald Sie dies getan haben, drücken Sie den Start-Button. Es folgt die Handlungsanweisung zusammen mit dem Hinweis, ob Sie diese nun ausführen oder nur lesen sollen. Ganz wichtig für die Durchführung ist, dass Sie sich bei beiden Vorgehensweisen (Lesen bzw. Ausführen) gleich lang mit der Handlungsanweisung beschäftigen. Um dies sicherzustellen, ist für jede Handlungsanweisung die gleiche Zeit vorgesehen, die nach Erscheinen der Handlungsanweisung abläuft. Ihnen steht so viel Zeit zu Verfügung, dass Sie die Handlungsanweisungen mehrmals lesen bzw. ausführen können. Bitte wiederholen Sie das Lesen bzw. Ausführen einfach so lange, bis ein Endsignal ertönt und die Handlungsanweisung ausgeblendet wird. Auf der nächsten Seite folgt noch ein bisschen Instruktion...

Noch einmal ganz konkret: Hätten Sie sich z.B. einen Stift und ein Blatt Papier holen sollen und läge Ihnen nun die Handlungsanweisung 'Zeichnen Sie einen Kreis auf das Papier!' zusammen mit der Aufforderung, diese zu lesen, vor, würden Sie Stift und Papier einfach nur vor sich liegen lassen, d.h. die Handlung NICHT ausführen, sondern die Anweisung so lange lesen, bis das abschließende Signal ertönt und die Handlungsanweisung ausgeblendet wird. Läge Ihnen dieselbe Handlungsanweisung mit der Bitte vor, diese auszuführen, würden Sie so lange Kreise auf das Papier zeichnen, bis das Endsignal ertönt ist. Bitte legen Sie die Utensilien jeweils wieder an ihren Platz zurück! So geht es dann immer weiter, bis die nächste Untersuchungsphase startet. Sie sollen die Handlungsanweisungen nun allerdings NICHT möglichst schnell lesen oder ausführen, sondern Sie sollen in einem angemessenen Tempo vorgehen. Bei den meisten Handlungsanweisungen ist es im Übrigen möglich, sie auf verschiedene Arten und Weisen auszuführen bzw. zu wiederholen (z.B. jeweils ein neues Blatt Papier nehmen, jeweils einen neuen Kreis auf das alte Blatt Papier zeichnen oder denselben Kreis immer wieder nachzeichnen). Die genaue Umsetzung liegt aber ganz bei Ihnen. Ich möchte Sie nur bitten, nicht allzu verschwenderisch mit den Materialien umzugehen. Damit ich sicher gehen kann, welche Anweisungen Sie gelesen und welche Sie ausgeführt haben, werden Sie jedes Mal nach Ablauf der Zeit gefragt, ob Sie die jeweilige Anweisung gelesen oder ausgeführt haben. Es folgt die Frage, wie oft sie dies getan haben. Geben Sie dann bitte EHRlich an, ob bzw. wie häufig Sie die Handlungsanweisung VOLLSTÄNDIG gelesen bzw. KOMPLETT ausgeführt haben. Für diese Untersuchung ist es wirklich ganz WICHTIG, dass Sie die Handlungen - je nach Instruktion - tatsächlich ausführen oder eben NICHT ausführen, d.h. lesen! BEVOR Sie nun den Weiter-Button anklicken, möchte ich Sie noch bitten, sich einen kurzen Überblick über die Materialien zu verschaffen. Lassen Sie einfach kurz Ihren Blick über die Utensilien schweifen, so dass Sie eine gewisse Übersicht haben, und starten Sie erst dann das Programm. Haben Sie noch Fragen? Dann stellen Sie diese bitte jetzt! Wenn Sie keine Fragen mehr haben, starten Sie das Programm bitte über den Weiter-Button.

Phase 2

Vorstellen

Im Folgenden werden Ihnen wiederum Handlungsanweisungen präsentiert. In dieser Untersuchungsphase haben Sie allerdings immer nur eine einzige Aufgabe: Sich die jeweilige Handlungsausführung - so häufig wie es Ihnen in der vorgegebenen Zeit möglich ist - vorzustellen. Dabei kommt es allerdings wiederum NICHT auf Schnelligkeit an, sondern darauf, dass Sie sich die Handlung MÖGLICHST LEBHAFT vorstellen! Konkret wird Ihnen gleich - genau wie in der ersten Untersuchungsphase - jeweils eine Handlungsanweisung präsentiert. Sobald diese auf dem Bildschirm erschienen ist und von Ihnen einmal gelesen wurde, schließen Sie die Augen und stellen sich deren Ausführung vor, bis das Endsignal ertönt bzw. der Bildschirm gelöscht wird. Die Zeit ist wieder vorgegeben und zählt nun vom Erscheinen der Handlungsanweisung an. Nach dem Ausblenden der Handlungsanweisung werden Sie gefragt, wie oft Sie sich die Handlung vorgestellt haben. Bitte geben Sie hier an, wie oft Sie sich die Handlungsausführung VOLLSTÄNDIG vorgestellt haben. Wie bereits erwähnt, geht es aber nicht darum, sich die Ausführung der Handlungsanweisung (z.B. 'Zeichnen Sie einen Kreis auf das Papier!') möglichst oft vorzustellen, sondern Sie sollen sich die Ausführung möglichst LEBHAFT vorstellen. Stellen Sie sich also nicht nur ganz deutlich vor, wie es AUSSEHEN würde, wenn Sie die Handlung ausführen würden (z.B. was für eine Farbe hat der Stift in Ihrer Hand?), sondern auch, welche GERÄUSCHE (und ggf. GERÜCHE) entstehen würden (z.B. welches Geräusch macht der Stift auf dem Papier?) und wie sich die Handlungsausführung ANFÜHLEN würde (z.B. wie fühlt sich der Stift beim Zeichnen in Ihrer Hand an?!). Es ist ganz wichtig, dass Sie sich die Handlung möglichst lebhaft, d.h. angereichert mit möglichst vielen sensorischen Eindrücken, vorstellen! Daher werden Sie anschließend gebeten, die Lebhaftigkeit Ihrer Vorstellung auf einer siebenstufigen Skala einzuschätzen. Ich möchte Sie darauf hinweisen, dass die Handlungsanweisungen sich z.T. wiederholen. Die Aufgabe erfordert also relativ viel Anstrengung und Ausdauer. Ich bitte Sie jedoch eindringlich, sich zu bemühen und wirklich gut zu konzentrieren! Haben Sie noch Fragen? Dann stellen Sie diese bitte jetzt! Wenn Sie keine Fragen mehr haben, starten Sie das Programm bitte über den Weiter-Button.

Beobachten

Im Folgenden werden Ihnen wiederum Handlungsanweisungen präsentiert, diesmal allerdings zusammen mit Videos der Handlungsausführungen. In dieser Untersuchungsphase haben Sie daher immer nur eine einzige Aufgabe: Die jeweilige Ausführung der Handlungsanweisung genau zu verfolgen - d.h. genau zuzuschauen und zuzuhören. Es kommt also darauf an, dass Sie sich (visuell und akustisch) vollkommen auf die Videos konzentrieren! Konkret wird Ihnen gleich - genau wie in der ersten Untersuchungsphase - jeweils eine Handlungsanweisung präsentiert; nun allerdings zusammen mit einem Video. Sobald diese auf dem Bildschirm erschienen ist, lesen Sie zuerst einmal die Anweisung und schauen dann das Video so lange, bis das Endsignal ertönt bzw. der Bildschirm gelöscht wird. Die Zeit ist wieder vorgegeben und zählt nun vom Erscheinen der Handlungsanweisung an. Nach dem Ausblenden werden Sie gefragt, wie oft Sie die Handlungsausführung gesehen haben. Bitte geben Sie hier an, wie oft Sie die Handlungsausführung VOLLSTÄNDIG gesehen haben. Anschließend werden Sie gebeten, auf einer siebenstufigen Skala einzuschätzen, wie konzentriert Sie die Handlung visuell und akustisch verfolgt haben. Ich möchte Sie darauf hinweisen, dass die Handlungsanweisungen sich z.T. wiederholen. Die Aufgabe erfordert also relativ viel Anstrengung und Ausdauer. Ich bitte Sie jedoch eindringlich, sich zu bemühen und wirklich gut zu konzentrieren! Haben Sie noch Fragen? Dann stellen Sie diese bitte jetzt! Wenn Sie keine Fragen mehr haben, starten Sie das Programm bitte über den Weiter-Button.

Generieren

Im Folgenden werden Ihnen wiederum Handlungsanweisungen präsentiert. Allerdings liegen diese diesmal in Anagrammform vor, d.h. dass die Wortstellungen durcheinander geraten sind. In dieser Untersuchungsphase haben Sie daher immer nur eine einzige Aufgabe: Das jeweilige Satzanagramm zu lösen, die Lösung einmal zu schreiben und dann - so häufig wie es Ihnen in der vorgegebenen Zeit möglich ist - diese Lösung zu lesen. In dieser Phase kommt es also schon auf Schnelligkeit an! Konkret wird Ihnen gleich - genau wie in der ersten Untersuchungsphase - jeweils eine Handlungsanweisung präsentiert, nun allerdings in Anagrammform und zusammen mit einem grauen Eingabefeld. Sobald Sie das Anagramm gelöst haben, schreiben Sie die Lösung bitte einmal in das Feld. Sie können einfach drauflos schreiben (Groß- und Kleinschreibung bitte beachten) und müssen nicht zuerst in das Feld klicken. Danach lesen Sie Ihre Lösung einfach immer wieder, bis das Endsignal ertönt bzw. der Bildschirm gelöscht wird. Die Zeit ist wieder vorgegeben und zählt nun vom Erscheinen der Handlungsanweisung an. Damit ich nachher weiß, ob Sie das Anagramm richtig gelöst haben, wird Ihnen im Folgenden die richtige Lösung präsentiert und Sie werden gefragt, wie oft Sie diese Lösung gelesen haben. Falls Sie diese Lösung zwar herausbekommen, jedoch in der kurzen Zeit nicht vollständig aufschreiben bzw. nicht mehr lesen konnten, tragen Sie hier bitte eine 0 ein. Falls Sie eine andere Lösung herausbekommen haben, tragen Sie hier bitte eine 99 ein und bei richtiger Lösung die Anzahl an VOLLSTÄNDIG gelesenen Anweisungen. Anschließend werden Sie gebeten, die Schwierigkeit des Anagramms auf einer siebenstufigen Skala einzuschätzen. Ich möchte Sie darauf hinweisen, dass die Handlungsanweisungen sich z.T. wiederholen und die entsprechenden Anagramme immer schwieriger werden. Neben zunehmenden Wortvertauschungen fehlen auch zunehmend einzelne Buchstaben, es werden zusätzlich Zahlen und/ oder Buchstaben eingefügt bzw. die Groß- und Kleinschreibung geht verloren. Die Aufgabe ist also nicht ganz leicht und erfordert zudem relativ viel Anstrengung und Ausdauer. Ich bitte Sie jedoch eindringlich, sich zu bemühen und wirklich gut zu konzentrieren! Haben Sie noch Fragen? Dann stellen Sie diese bitte jetzt! Wenn Sie keine Fragen mehr haben, starten Sie das Programm bitte über den Weiter-Button.

Lesen

Im Folgenden werden Ihnen wiederum Handlungsanweisungen präsentiert. In dieser Untersuchungsphase haben Sie allerdings immer nur eine einzige Aufgabe: Die jeweilige Handlungsanweisung - so häufig wie es Ihnen in der vorgegebenen Zeit möglich ist - zu lesen. In dieser Phase kommt es also schon auf Schnelligkeit an! Konkret wird Ihnen gleich - genau wie in der ersten Untersuchungsphase - jeweils eine Handlungsanweisung präsentiert. Sobald diese auf dem Bildschirm erschienen ist, lesen Sie sie einfach so lange, bis das Endsignal ertönt bzw. der Bildschirm gelöscht wird. Die Zeit ist wieder vorgegeben und zählt nun vom Erscheinen der Handlungsanweisung an. Nach dem Ausblenden der Handlungsanweisung werden Sie gefragt, wie oft Sie die Anweisung gelesen haben. Bitte geben Sie hier an, wie oft Sie die Anweisung VOLLSTÄNDIG gelesen haben. Anschließend werden Sie gebeten, anzugeben, wie viele Konsonanten die Handlungsanweisung enthielt. Zählen Sie daher bitte möglichst direkt beim ersten Durchlesen, wie viele Konsonanten vorhanden sind, merken Sie sich diese Zahl und tragen Sie sie später in das leere Feld ein. Ich möchte Sie darauf hinweisen, dass die Handlungsanweisungen sich z.T. wiederholen. Die Aufgabe erfordert also relativ viel Anstrengung und Ausdauer. Ich bitte Sie jedoch eindringlich, sich zu bemühen und wirklich gut zu konzentrieren! Haben Sie noch Fragen? Dann stellen Sie diese bitte jetzt! Wenn Sie keine Fragen mehr haben, starten Sie das Programm bitte über den Weiter-Button.

Phase 3

Liebe Untersuchungsteilnehmerin, lieber Untersuchungsteilnehmer, im Rahmen des ersten Untersuchungstermins haben Sie mehrere Untersuchungsphasen durchlaufen. Bitte denken Sie jetzt ausschließlich an die ERSTE Untersuchungsphase zurück, in der Sie Handlungsanweisungen entweder gelesen oder ausgeführt haben. Alles, was Sie innerhalb der zweiten und dritten Untersuchungsphase gemacht haben, ist für die folgende Aufgabe irrelevant. Ihnen werden im Folgenden noch einmal Handlungsanweisungen präsentiert. Sie sollen bitte jeweils angeben, ob diese Handlungsanweisungen aus der ERSTEN Untersuchungsphase stammen, ob Sie diese Handlungsanweisungen in dieser ersten Untersuchungsphase gelesen oder ausgeführt haben und wie sicher Sie sich bei diesen Einschätzungen sind. Beziehen Sie sich bitte nur auf die ALLERERSTE UNTERSUCHUNGSPHASE! Haben Sie noch Fragen? Dann stellen Sie diese bitte jetzt! Wenn Sie keine Fragen mehr haben, starten Sie das Programm bitte über den Weiter-Button.

B Materialien des zweiten Experiments (Vor- und Hauptuntersuchung)

B.1 Einverständniserklärung



University
of
Alberta

Faculties of Arts, Science, and Law REB (#1632 DK-11-16-07-070)

INFORMED CONSENT FORM: Undergraduate Research Pool Members

Project: Reality monitoring in young and older people

Principal Investigator: Patrick Davidson, PhD; Department of Psychology, University of Alberta; (780) 492-2237. Additional investigators: Isabel Lindner and Leslie McGhan; (780) 492-1277.

This study takes place in two sessions with an interval of a few days between these sessions. During both you will be asked to examine different kinds of materials, such as action statements, words or pictures. Some of these tasks will involve remembering the materials, but others will not. You will be given specific instructions before each test, and we will explain the purpose of each test at the end of the second session. Our goal is to assess how memory and other cognitive processes are affected by normal aging.

The two testing sessions should each last between half an hour and an hour. At the end of each session, you will be assigned your credits for the undergraduate research participation pool. Your decision to participate in this study is entirely voluntary. You have the right to withdraw from the study at any time and for any reason without penalty or loss of rights. If you choose not to participate or withdraw after you have begun, but would like your credit for participation, you may complete an alternative educational activity. The time it takes to complete this activity will take no longer than the time it takes to participate in this study. Little physical or psychological risk is involved. This research is not designed to aid or provide insight into your case specifically, but rather to help us learn about potential similarities and differences between relatively large groups. The results from your testing session will be put together with those from other people, in order for us to make large-scale comparisons among different groups (for example, young people versus older people). The information obtained from this study may be presented at scientific conferences or in scientific journals, but your name will never appear in any public document. In order to ensure confidentiality, your name and other identifying information will be filed separately from the experimental data in our laboratory, and all data are stored in locked filing cabinets and on password-protected computers in our department.

Contact Information. If you have any questions or comments on the study, or if you wish a clarification of rights as a research participant, you can contact Dr. Patrick Davidson or the Human Research Ethics Committee at the number and address below.

Investigator

Dr. Patrick Davidson
Department of Psychology
University of Alberta
Edmonton, AB T6G 2E9
(780) 492-2237

Ethics Committee Member

Dr. Don Heth
Department of Psychology
University of Alberta
Edmonton, AB T6G 2E9
(780) 492-2662

Signature Section. This research project has been explained to me in detail, and I have been provided with a copy of the Informed Consent Form. I have been given an opportunity to ask questions concerning these procedures, and any questions that I have asked have been adequately answered. I understand that all information will be kept confidential and I will not be identified in any way in any publication of the results. I understand that I do not have to take part in this study, and my refusal to participate will involve no penalty or loss of rights to which I am entitled. I may withdraw from this study at any time and for any reason without penalty or loss of rights. I hereby consent to participate.

Name: _____

Signature: _____ Date: _____

Name of Person Obtaining Consent: _____

Signature: _____ Date: _____

B.2 Aufklärungsformular



University
of
Alberta

DEBRIEFING FORM—Undergraduate Pool Members

Project: Reality monitoring in young and older people

Principal Investigator: Patrick Davidson, PhD; Department of Psychology, University of Alberta; (780) 492-2237. Additional investigators: Isabel Lindner and Leslie McGhan (780) 492-1277.

Thank you very much for taking part in our research. You completed an experiment that we designed in order to study the effects of healthy aging on memory. To understand how memory is different between young and older people, we will compare the results of young and older adults on these tasks.

Memory often declines with advancing age, but not all aspects of memory are affected equally. Previous research has shown that although healthy older adults are good at recognizing studied items, they are less accurate when recognizing the *context* associated with these items. For instance, one might recognize a face as being highly familiar (i.e., item memory), but be unable to recall the person's name or the situation in which they encountered this person (i.e., context memory). In our special paradigm, we have examined memory for the modality in which an action is encoded. For instance, in real life one might wonder if one has actually taken one's pills or merely thought about taking them. To study this kind of memory phenomenon (known as reality monitoring) in the laboratory, we asked you to perform some action statements but not others in the first session, and to indicate the encoding modality – whether you have performed it or not – in the second session. We also asked you to rate how confident you were in your source decisions. This allowed us to have a measure of your subjective experience. Our main goal is to find out whether older people consistently show problems with memory for context, or whether it depends on the exact way that we test memory.

Remember, the results that we have collected are confidential. Data will be coded using either a participant number or your initials, and stored in a secured room. Only Dr. Davidson and those he authorizes will have access to the log linking participants' numbers or initials to names. Your name, and other identifying information, will NOT be used when we report our results in presentations and/or scientific publications.

The results from your testing session will be put together with those from other people, in order for us to make large-scale comparisons among different groups (for example, young people versus older people).

Thanks again for taking part. Feel free to contact our lab (492-1277) or our primary investigator, Professor Patrick Davidson (492-2237), should you have any questions or comments.

If you have general questions, contact Dr. Tom Spalding (Director, Research Participation) at spalding@ualberta.ca or 492-7778, or Sharon Randon (Research Participation Coordinator) at rescred@ualberta.ca or 492-5689. Please don't tell other people about what we had you do here because other students in the class may participate in this study.

B.3 Handlungsanweisungen

Pool	Nr.	Handlungsanweisung	Darst. (Video)	Objekt
Pool 1				
A	1	Roll the dice!	m	
	2	Shake the bottle!	m	
	3	Wrap up the ribbon around your fingers!*	m	
	4	Fold the paper!	w	
	5	Operate the remote control!*	w	
B	6	Shuffle the cards!	w	
	7	Put the egg in the eggcup!	w	
	8	Turn the sock inside out!	m	
	9	Sharpen the pencil!*	w	
	10	Tear off a piece of paper from the pad!*	m	
C	11	Put the soap in the case!	m	

	12	Open the safety pin!	w	
C	13	Hole punch the paper!	m	
	14	Unbend a paper clip!***	w	
	15	Take a match out of the match box!***	m	
	Pool 2			
	16	Fold the towel!	w	
	17	Stretch the rubber band!	w	
A	18	Draw a circle on the paper!**	m	
	19	Shine the flashlight on the table!*	w	
	20	Take the lid off of the box!	m	
B	21	Ring the bell!	m	
	22	Shake the tube of sprinkles!**	w	

	23	Put the candle in the glass!	m	
B	24	Take a piece of the tape!*	m	
	25	Open the book!	m	
	26	Push the toy car!	w	
	27	Put the teabag in the cup!	m	
C	28	Put the watch one hour forward!	w	
	29	Squeeze the sponge!**	m	
	30	Close the sunglasses!***	m	
Pool 3				
	31	Stamp the paper!	m	
A	32	Put the pack of noodles in the bag!	m	
	33	Put the toothbrush in the cup!**	w	

A	34	Open the pencil case!*	w	
	35	Get out a tissue!	m	
B	36	Screw off the lid of the deodorant!*	w	
	37	Cut off a small piece of thread from the spool!	w	
	38	Tear the paper!	w	
	39	Take the CD out of the case!***	m	
	40	Unlock the lock!	m	
C	41	Throw the ball from one hand to the other!*	w	
	42	Take a cup from the stack!	w	
	43	Put the glove on your hand!**	m	
	44	Unscrew the pen!	w	
	45	Put the card in the envelope!	m	

Pool 4 (Distraktoren)		
A	46	Unwrap a candy!
	47	Operate the lighter!
	48	Take the battery out of the alarm clock!
	49	Use the airpump!
	50	Flip the coin!
B	51	Squeeze the rubber duck!
	52	Sign the paper!
	53	Use the calculator!
	54	Take the ring off of the key!
	55	Make a knot in the cord!
B	56	Staple the papers!
	57	Open the penknife!
	58	Pull the plug out of the outlet!
	59	Take the cap off of the highlighter!
	60	Crack a nut!

Ausgetauschte Videos in der Voruntersuchung zu Experiment II (Anzahl gesamt, Anzahl pro Akteur/ Akteurin):

* = altes Item, neues Objekt; $n=9$; 3 m, 6 w

** = neues Item, neues Objekt; $n=5$; 3 m, 2 w

*** = altes Item, altes Objekt; $n=4$; 3 m, 1 w

→ $N = 18$ ausgetauschte Videos; davon 9 des Akteurs und 9 der Akteurin; alle anderen Videos waren alt, d.h. stammten aus Experiment I

In der Hauptuntersuchung wurden nur neue Videos genutzt und daher alle Handlungen nur noch durch einen männlichen Akteur ausgeführt.

B.4 Instruktionen (Voruntersuchung)

Phase 1

Dear participant,
This study comprises several phases and today you will pass through the first three of these phases. Within the first phase, several simple action statements like 'cut the paper' will be presented. Your task is to either read these action statements or to read and then actually perform the actions. The instruction how to handle each action statement - if you should just read it or read and perform it - will be provided together with each action statement. The procedure will be as follows: First you will be told which object(s) the following action statement will involve. Thus, an example of the first request could be: 'Please get: scissors and paper'. Then, you would get these objects from the assortment beside and lay them on the table in front of you. When you have done this, click on the continue-button. Now, the action statement will appear together with an instruction how to handle it (i.e. if you should just read or read and perform it). It's very important, that you take the same amount of time for each action statement - regardless of the exact handling (reading or performing). To ensure this, the same amount of time is allotted for each presentation of action statements. The action statements will just disappear when the allotted time has expired. There will be sufficient time for reading or performing, and in most cases you will even have more time than you would need to read the action statement or to perform the action once. Please, repeat reading or performing until you'll hear a signal and the action statement disappears. The procedure is explained in more detail on the following page. Please click the continue-button for further instructions.

For example, if you had been told to get the scissors and a sheet of paper, you would have gotten these objects from the assortment and lay them in front of you. You would then press the continue-button and the action statement would appear together with the instruction how to handle it. If the action statement was 'cut the paper' and you were instructed to read it, you would just ignore the objects lying in front of you, i.e. you would NOT perform the action with the materials and NOT even touch them, but read and reread the action statement until you hear the signal and the action statement disappears. In case you were given the same action statement together with the instruction to perform it, you would perform and re-perform the action, i.e. cutting the paper, until you hear the signal and the action statement disappears. Please replace every object after being done with each action statement. Please note: You don't need to hurry when reading or performing an action statement. It's not important that you do it as many times as possible, but rather that you do it with an appropriate tempo. In many cases there are different possibilities how to perform and re-perform an action statement (e.g. you could first cut the paper in half, then cut these halves in half again etc. or you could cut the paper in strips). It's left up to you how exactly you perform the action. Please don't work too lavishly. I won't observe you during the course of the study, but I want to ensure that each action statement has been handled as requested. Thus, after each presentation you're asked to indicate whether you read or performed the last action statement. In addition, you are asked how often you read or performed it. Please indicate HONESTLY whether and how many times you completely read or performed an action statement. For this study it's really important that you handle the action statements as instructed! You will work autonomously for the rest of this phase, but first I'll show you two examples and give you time to ask questions. Please let me know when you have finished reading this instruction.

Phase 2

In this third phase, action statements will be presented again, but now together with videos showing action performance. You should proceed in exactly the same way with each presentation: First, you should read the action statement once and then, you should faithfully observe action performance, i.e. concentrate visually and acoustically on the videos. The most important thing is to concentrate on the shown action - to watch and listen carefully to these videos. One action statement after the other will be presented together with a video showing action performance. Please, read the action statement once and watch the video until you hear a signal and the screen will be cleared. Time is again allotted for each presentation. After an action statement and the corresponding video have disappeared, you're asked how many times you saw action performance. Please indicate how many times you saw COMPLETE action performance. Then, you're asked to rate how much you concentrated visually and acoustically on action performance on a seven-point scale. Please note: Some action statements are presented repeatedly during the course of this phase, but it's really important that you concentrate on each video as much as possible - even if you already know it! This task thus requires a little bit of endurance. Please, try hard to concentrate as much as possible on each video! You will work autonomously for the rest of this phase, but first I'll show you an example and give you time to ask questions. Please let me know when you have finished reading this instruction.

Phase 3

Dear participant,
Please think back to the first session of this study now. Just to remind you, in the first phase of this first session, several action statements were presented and you were asked to just read or to read and actually perform them. Then, you completed a short questionnaire and afterwards, you observed some videos. Now, action statements will be presented, again. Your first task today will be to indicate for each action statement if it was presented in the FIRST phase of the first session or if it was not presented then. If you indicate that an action statement was presented in the first phase, you will furthermore be asked if you performed this action or if you didn't perform it, that is merely read the action statement. In addition, you will be asked to rate confidence concerning your memory judgment. Please don't hesitate to ask me if you have any questions! You can start the program by clicking on the continue-button.

B.5 Instruktionen (Hauptuntersuchung)

Phase 1

Dear participant,
This study comprises several phases and will involve computer-based as well as paper-pencil tasks. Within the first computer-based phase, several simple action statements like 'cut the paper' will be presented. Your task is to read these action statements. Also, you will be asked to actually perform some, but not all of them. In addition, we would like to know how familiar the objects which are involved in the actions are for you. Thus, for every object you will be asked on how many days you use it in a usual week. In many cases there are different possibilities how to perform an action (e.g. cutting the paper in two halves or cutting an edge off of the paper). It's left up to you how exactly you perform the action. Also, you don't need to hurry - there's enough time to perform every action. Now, we'll go through the details of the procedure. I'll show you two examples and give you time to ask questions.

Phase 2

Lesen

In this second computer-based phase, some of the action statements will be presented again, but now together with some shapes. You should proceed in exactly the same way with each presentation: First, you should read the action statement once and then, you should faithfully observe the different shapes which will appear on the screen. Both tasks are equally important. One action statement after the other will be presented together with different shapes. Please, read the action statement once and then observe the shapes until you hear a signal and the screen will be cleared. Time is again allotted for each presentation. After an action statement and the shapes have disappeared, you're asked how many circles you saw. Please indicate then how many circles appeared during the preceding presentation. Please note: Some action statements are presented repeatedly during the course of this phase, but it's really important that you concentrate on each presentation as much as possible - even if you already know it! This task thus requires a little bit of endurance. Please, try hard to concentrate as much as possible on each presentation! First, I'll show you an example and give you time to ask questions.

Beobachten

In this second computer-based phase, some of the action statements will be presented again, but now together with videos showing action performance. You should proceed in exactly the same way with each presentation: First, you should read the action statement once and then, you should faithfully observe the action performance, i.e. concentrate visually and acoustically on the videos. The most important thing is to concentrate on the shown action - to watch and listen carefully to these videos. One action statement after the other will be presented together with a video showing action performance. Please, read the action statement once and watch the video until you will hear a signal and the screen will be cleared. Time is again allotted for each presentation. After an action statement and the corresponding video have disappeared, you're asked how many times you saw the action performance. Please indicate then, how many times you saw COMPLETE action performance. Please note: Some action statements are presented repeatedly during the course of this phase, but it's really important that you concentrate on each presentation as much as possible - even if you already know it! This task thus requires a little bit of endurance. Please, try hard to concentrate as much as possible on each presentation! Now, I'll show you an example and give you time to ask questions.

Phase 3

Lesen ohne Abrufhilfe

Dear participant,
Please think back to the first session of this study now. Just to remind you: In the first phase of the first session, several action statements were presented and you were asked to just read or to read and actually perform them. Later, within the shape-observation task, some of these action statements were re-presented. Now, action statements will be presented again. Your task is to indicate for each action statement if it was presented within the first session, i.e. if it is an old or a new action statement. If you indicate that an action statement was presented, you will furthermore be asked if you performed the action or if you didn't perform it, that is merely read the action statement. In addition, you will be asked to rate confidence concerning your memory judgments. Please don't hesitate to ask me if you have any questions!

Beobachten ohne Abrufhilfe

Dear participant,
Please think back to the first session of this study now. Just to remind you: In this first phase of the first session, several action statements were presented and you were asked to just read or to read and actually perform them. Later, within the video-observation task, some of these action statements were re-presented. Now, action statements will be presented again. Your task is to indicate for each action statement if it was presented within the first session, i.e. if it is an old or a new action statement. If you indicate that an action statement was presented, you will furthermore be asked if you performed the action or if you didn't perform it, that is merely read the action statement. In addition, you will be asked to rate confidence concerning your memory judgments. Please don't hesitate to ask me if you have any questions!

Beobachten mit Instruktion zur Quellenprüfung

Dear participant,
Please think back to the first session of this study now. Just to remind you: In the first phase of the first session, several action statements were presented and you were asked to just read or to read and actually perform them. Later, within the video-observation task, some of these action statements were represented. Now, action statements will be presented again. Your task is to indicate for each action statement if it was presented within the first session, i.e. if it is an old or a new action statement. If you indicate that an action statement was presented, you will furthermore be asked if you performed the action or if you didn't perform it, that is merely read the action statement. In addition, you will be asked to rate confidence concerning your memory judgments.

Research has determined that memory for perceptual and contextual information is a useful indicator of whether an action statement was actually performed, merely read, or not presented at all. We will explain this with the following example action statement 'Cut the paper!'. Please click the continue-button for further explanation!

If an action statement is OLD, you might have a vivid memory for the objects. You might remember, for example, that I brought out black scissors and a white sheet of paper. If an action statement is NEW, you probably will NOT have a vivid memory for the objects, because they were never shown to you within this study.

BE CAREFUL! You should make sure that – if you have a vivid memory – it comes from this study, and not from somewhere else: For example, you might have used scissors at home yesterday. We only want you to say 'old' to action statements that were presented within this study.

If an action statement was PERFORMED by you, you might in addition have a vivid memory for performing the action. You might clearly remember the way you performed the action, and/or how it felt, sounded, or looked when you were performing it. For example, you might remember how it felt when you were cutting the paper and that you cut it in half. Also, you might remember what you felt or thought while performing the action. For example, you might remember that you were not sure whether you should cut the paper the whole way through or not. Or you might, for example, remember that you thought that the scissors were similar to the ones you use at home.

You should make sure that – if you have a vivid memory for performing an action – it actually comes from your OWN performance. Remember, you also saw videos showing another person performing some of these actions, and so you have to make sure not to mix up performance of the person in the video with your own performance. For example, try to remember if you actually held the scissors in your hand, and whether you viewed cutting the paper from your own point of view.

If you DIDN'T PERFORM an action statement, you might also NOT remember the way you performed the action, and/ or how it felt, sounded, or looked when you were performing it, because you never touched the objects. You might nevertheless remember what you felt or thought in conjunction with that action statement – for example you might have been disappointed or glad that you were not supposed to perform a certain action.

So, try to ask yourself the following questions for each action statement and base your memory judgments on your answers to these questions:

- Do I have a vivid memory for the object?
- Do I have a memory for the way I performed the action?
- Do I have a memory for how action felt, looked, or sounded?
- Do I remember what I was feeling or thinking when I saw this action statement?

MAKE SURE you refer to what you experienced within THIS study and not other contexts. And make sure you refer to your OWN action performance.

Please don't hesitate to ask me if you have any questions.

Beobachten mit Warnung

Dear participant,
Please think back to the first session of this study now. Just to remind you: In the first phase of the first session, several action statements were presented and you were asked to just read or to read and actually perform them. Later, within the video-observation task, some of these action statements were re-presented. Now, action statements will be presented again. Your task is to indicate for each action statement if it was presented within the first session, i.e. if it is an old or a new action statement. If you indicate that an action statement was presented, you will furthermore be asked if you performed the action or if you didn't perform it, that is merely read the action statement. In addition, you will be asked to rate confidence concerning your memory judgments. Please click the continue-button for further instructions!

In the first session you watched some videos showing another person performing some actions that you DID actually perform yourself. But also, you watched some videos showing another person performing some actions that you DID NOT actually perform yourself.
BE CAREFUL concerning these vidoes about actions that you DID NOT perform yourself!
Research has determined that observing another person performing actions, especially repeatedly, can lead to false memories for action performance. That is, people might sometimes think that they actually performed an action when in fact they only observed another person performing it. So, you should be careful when performing this memory test because you watched some videos showing another person performing some actions that you didn't actually perform yourself.
Try not to let these videos influence your judgments about whether you performed an action or not.

C Materialien zum dritten Experiment

C.1 Handlungsanweisungen

Pool	Nr.	Handlungsanweisung
Pool 1		
A	1	Stoßen Sie das Spielzeugauto an!
	2	Schütteln Sie die Flasche!
	3	Wickeln Sie das Band auf!
	4	Falten Sie das Papier!
	5	Nehmen Sie ein Taschentuch heraus!
B	6	Mischen Sie die Karten!
	7	Tun Sie das Ei in den Eierbecher!
	8	Ziehen Sie die Socke auf links!
	9	Spitzen Sie den Bleistift an!
C	10	Stellen Sie den Wecker eine Stunde vor!
	11	Legen Sie die Seife in die Dose!
	12	Nehmen Sie die CD aus der Hülle!
	13	Lochen Sie das Papier!
	14	Biegen Sie eine Büroklammer auseinander!
15	Nehmen Sie ein Streichholz aus der Schachtel!	
Pool 2		
A	16	Falten Sie das Handtuch!
	17	Blättern Sie in dem Buch!
	18	Nehmen Sie ein kleines Stück Tesaband von der Rolle!
	19	Leuchten Sie mit der Taschenlampe!
B	20	Nehmen Sie den Deckel von der Schachtel!
	21	Betätigen Sie die Klingel!
	22	Schütteln Sie die Perlen!
	23	Stellen Sie das Teelicht in das Glas!
	24	Zeichnen Sie einen Kreis auf das Papier!
C	25	Dehnen Sie das Gummiband!
	26	Würfeln Sie mit den Würfeln!
	27	Tun Sie den Teebeutel in die Tasse!
	28	Reißen Sie einen Zettel vom Block ab!
	29	Wringen Sie den Schwamm!
	30	Klappen Sie die Sonnenbrille zusammen!
Pool 3		
A	31	Stempeln Sie auf das Papier!
	32	Legen Sie die Nudelpackung in die Tasche!
	33	Tun Sie die Zahnbürste in den Becher!
	34	Öffnen Sie das Mäppchen!
	35	Ziehen Sie den Handschuh an!
B	36	Schrauben Sie den Deckel vom Deo!
	37	Öffnen Sie die Sicherheitsnadel!
	38	Zerreißen Sie das Papier!
	39	Stecken Sie den Stecker in die Dose!
C	40	Werfen Sie den Ball von der einen in die andere Hand!
	41	Schließen Sie das Schloss auf!
	42	Nehmen Sie einen Becher vom Stapel!
	43	Betätigen Sie die Fernbedienung!
	44	Schrauben Sie den Kugelschreiber auseinander!
45	Stecken Sie die Karte in den Umschlag!	

Pool 4 (Distraktoren)

- | | | |
|----------|----|--|
| A | 46 | Wickeln Sie ein Bonbon aus! |
| | 47 | Legen Sie die Gabel auf den Teller! |
| | 48 | Legen Sie die Uhr an! |
| | 49 | Betätigen Sie die Luftpumpe! |
| | 50 | Werfen Sie die Münze! |
| B | 51 | Drücken Sie das Quietscheentchen! |
| | 52 | Ziehen Sie mit dem Lineal einen Strich auf das Papier! |
| | 53 | Betätigen Sie den Taschenrechner! |
| | 54 | Klappen Sie den Zollstock auseinander! |
| | 55 | Machen Sie einen Knoten in die Kordel! |
| C | 56 | Stanzten Sie einen Stern aus dem Papier aus! |
| | 57 | Öffnen Sie die Gürtelschnalle! |
| | 58 | Stechen Sie eine Nadel in das Nadelkissen! |
| | 59 | Nehmen Sie die Kappe vom Textmarker! |
| | 60 | Wählen Sie eine Nummer auf dem Handy! |
-

C.2 Instruktionen

Phase 1

Durchführung mit geöffneten Augen

Liebe Teilnehmerin, lieber Teilnehmer, vielen herzlichen Dank für Ihre Bereitschaft, zu dieser Untersuchung beizutragen! Diese Untersuchung besteht aus mehreren Phasen, von denen Sie heute die ersten drei durchlaufen werden. Im Rahmen der ersten Untersuchungsphase, mit der Sie nun beginnen werden, werden Ihnen mehrere einfache Handlungsanweisungen präsentiert, wie z.B. 'Zerschneiden Sie das Papier!'. Diese Anweisungen sollen Sie entweder nur lesen oder lesen und mit Gegenständen, die ich Ihnen anreichen werde, tatsächlich ausführen. Zusätzlich würden wir gerne wissen, wie vertraut Sie mit den jeweiligen Gegenständen sind. Daher werden Sie für jeden Gegenstand gefragt, an wievielen Tagen in einer gewöhnlichen Woche Sie ihn benutzen. In vielen Fällen gibt es mehrere Möglichkeiten, eine Handlungsanweisung umzusetzen. Z.B. könnten Sie das Papier längs oder quer zerschneiden usw. Es ist vollkommen Ihnen überlassen, wie Sie eine Handlungsanweisung umsetzen. Außerdem ist es Ihnen überlassen, in welcher Geschwindigkeit Sie eine Handlung ausführen - es steht Ihnen für jede Handlungsanweisung genug Zeit zur Verfügung und Sie brauchen sich nicht zu beeilen. Nun werden wir anhand von zwei Beispielen den konkreten Ablauf durchgehen. Dabei werden Sie auch Gelegenheit haben, Fragen zu stellen.

Durchführung mit geschlossenen Augen

Liebe Teilnehmerin, lieber Teilnehmer, vielen herzlichen Dank für Ihre Bereitschaft, zu dieser Untersuchung beizutragen! Diese Untersuchung besteht aus mehreren Phasen, von denen Sie heute die ersten drei durchlaufen werden. Im Rahmen der ersten Untersuchungsphase, mit der Sie nun beginnen werden, werden Ihnen mehrere einfache Handlungsanweisungen präsentiert, wie z.B. 'Zerschneiden Sie das Papier!'. Diese Anweisungen sollen Sie entweder nur lesen oder lesen und mit Gegenständen, die ich Ihnen anreichen werde, tatsächlich ausführen. Dabei sind Sie gebeten, während der Handlungsausführung die Augen geschlossen zu halten. Zusätzlich würden wir gerne wissen, wie vertraut Sie mit den jeweiligen Gegenständen sind. Daher werden Sie für jeden Gegenstand gefragt, an wievielen Tagen in einer gewöhnlichen Woche Sie ihn benutzen. In vielen Fällen gibt es mehrere Möglichkeiten, eine Handlungsanweisung umzusetzen. Z.B. könnten Sie das Papier längs oder quer zerschneiden usw. Es ist vollkommen Ihnen überlassen, wie Sie eine Handlungsanweisung umsetzen. Außerdem ist es Ihnen überlassen, in welcher Geschwindigkeit Sie eine Handlung ausführen - es steht Ihnen für jede Handlungsanweisung genug Zeit zur Verfügung und Sie brauchen sich nicht zu beeilen. Nun werden wir anhand von zwei Beispielen den konkreten Ablauf durchgehen. Dabei werden Sie auch Gelegenheit haben, Fragen zu stellen.

Phase 2

Durchführung mit Ton

Im Folgenden werden Ihnen wiederum Handlungsanweisungen präsentiert, diesmal allerdings zusammen mit Videos, die die Handlungsausführung zeigen. In dieser Untersuchungsphase haben Sie für jede Handlungsanweisung die gleiche Aufgabe: Sie sollen zunächst die jeweilige Handlungsanweisung lesen und dann das entsprechende Video aufmerksam verfolgen - d.h. mit voller Konzentration zuschauen und zuhören. Ihre wichtigste Aufgabe ist es, sich (visuell und akustisch) vollkommen auf die Videos zu konzentrieren! Konkret wird Ihnen gleich jeweils eine Handlungsanweisung zusammen mit einem Video präsentiert. Sobald eine Handlungsanweisung auf dem Bildschirm erschienen ist, lesen Sie diese bitte einmal durch und beobachten dann das Video so lange, bis das Endsignal ertönt bzw. der Bildschirm gelöscht wird. Die Präsentationsdauer ist jeweils vorgegeben. Nach dem Ausblenden werden Sie gefragt, wie oft Sie die Handlungsausführung gesehen haben. Bitte geben Sie hier an, wie oft Sie die VOLLSTÄNDIGE Handlungsausführung gesehen haben. Dabei ist es Ihnen überlassen, was Sie unter 'vollständiger Handlungsausführung' verstehen. Bezogen auf die Handlungsanweisung 'Zerschneiden Sie das Papier!' könnten Sie z.B. jeden Schnitt zählen, den die Person getätigt hat. Sie könnten auch zählen, wie häufig die Person das Papier insg. zerschnitten hat. Es gibt kein richtig oder falsch! Zählen Sie so, wie es Ihrer Meinung nach am meisten Sinn macht. Ich möchte Sie darauf hinweisen, dass die Handlungsanweisungen sich z.T. wiederholen. Die Aufgabe erfordert also relativ viel Anstrengung und Ausdauer. Ich bitte Sie jedoch eindringlich, sich zu bemühen und sich jedesmal erneut gut zu konzentrieren! Bitte starten Sie diese Phase über den Weiter-Button. Bei Fragen wenden Sie sich gerne an mich.

Durchführung ohne Ton

Im Folgenden werden Ihnen wiederum Handlungsanweisungen präsentiert, diesmal allerdings zusammen mit Videos, die die Handlungsausführung zeigen. In dieser Untersuchungsphase haben Sie für jede Handlungsanweisung die gleiche Aufgabe: Sie sollen zunächst die jeweilige Handlungsanweisung lesen und dann das entsprechende Video aufmerksam verfolgen - d.h. mit voller Konzentration zuschauen. Ihre wichtigste Aufgabe ist es, sich (visuell) vollkommen auf die Videos zu konzentrieren! Konkret wird Ihnen gleich jeweils eine Handlungsanweisung zusammen mit einem Video präsentiert. Sobald eine Handlungsanweisung auf dem Bildschirm erschienen ist, lesen Sie diese bitte einmal durch und beobachten dann das Video so lange, bis das Endsignal ertönt bzw. der Bildschirm gelöscht wird. Die Präsentationsdauer ist jeweils vorgegeben. Nach dem Ausblenden werden Sie gefragt, wie oft Sie die Handlungsausführung gesehen haben. Bitte geben Sie hier an, wie oft Sie die VOLLSTÄNDIGE Handlungsausführung gesehen haben. Dabei ist es Ihnen überlassen, was Sie unter 'vollständiger Handlungsausführung' verstehen. Bezogen auf die Handlungsanweisung 'Zerschneiden Sie das Papier!' könnten Sie z.B. jeden Schnitt zählen, den die Person getätigt hat. Sie könnten auch zählen, wie häufig die Person das Papier insg. zerschnitten hat. Es gibt kein richtig oder falsch! Zählen Sie so, wie es Ihrer Meinung nach am meisten Sinn macht. Ich möchte Sie darauf hinweisen, dass die Handlungsanweisungen sich z.T. wiederholen. Die Aufgabe erfordert also relativ viel Anstrengung und Ausdauer. Ich bitte Sie jedoch eindringlich, sich zu bemühen und sich jedesmal erneut gut zu konzentrieren! Bitte starten Sie diese Phase über den Weiter-Button. Bei Fragen wenden Sie sich gerne an mich.

Phase 3

Liebe Teilnehmerin, lieber Teilnehmer,
bitte denken Sie nun an die letzte Sitzung zurück. Um Ihnen diese wieder ins Gedächtnis zu rufen: In der ersten Phase der letzten Sitzung wurden Ihnen Handlungsanweisungen präsentiert und Sie wurden gebeten, diese tatsächlich auszuführen oder nur zu lesen. Später, während der Video-Beobachtungsaufgabe, wurden Ihnen einige dieser Handlungsanweisungen erneut präsentiert. Nun werden Ihnen wiederum Handlungsanweisungen dargeboten. Ihre Aufgabe ist es, für jede dieser Handlungsanweisungen zu entscheiden, ob diese innerhalb der letzten Sitzung bereits präsentiert wurde oder nicht, d.h. ob es sich um eine ALTE oder heute erstmals NEU präsentierte Handlungsanweisung handelt. Wenn Sie der Meinung sind, eine Handlungsanweisung sei Ihnen in der letzten Sitzung bereits präsentiert worden bzw. sei alt, werden Sie des Weiteren gefragt, ob Sie diese Handlungsanweisung tatsächlich ausgeführt haben oder nicht, d.h. sie nur gelesen haben. Zusätzlich werden Sie gebeten, die Sicherheit Ihres Erinnerungsurteils auf einer siebenstufigen Skala von <<gar nicht sicher>> bis <<absolut sicher>> anzugeben. Bei Fragen wenden Sie sich gerne an mich.

D Materialien zum vierten Experiment

D.1 Einverständniserklärung



University
of
Alberta

Faculties of Arts, Science, and Law REB (#1632 DK-11-16-07-070)

INFORMED CONSENT FORM: Community Members

Project: Reality monitoring in young and older people

Principal Investigator: Patrick Davidson, PhD; Department of Psychology, University of Alberta; (780) 492-2237. Additional investigators: Isabel Lindner and Leslie McGhan; (780) 492-1277.

This study takes place in two sessions with an interval of a few days between these sessions. During both you will be asked to examine different kinds of materials, such as action statements, words or pictures. Some of these tasks will involve remembering the materials, but others will not. You will be given specific instructions before each test, and we will explain the purpose of each test at the end of the second session. Our goal is to assess how memory and other cognitive processes are affected by normal aging.

The two testing sessions should each last between an hour and a half and two hours. We will take at least one short break, halfway through each session. At the end of the session, you will be reimbursed for your expenses (\$10 per session). Your decision to participate in this study is entirely voluntary. Little physical or psychological risk is involved. You have the right to withdraw from the study at any time and for any reason without penalty or loss of rights (and still receive reimbursement for research participation). This research is not designed to aid or provide insight into your case specifically, but rather to help us learn about potential similarities and differences between relatively large groups. The results from your testing session will be put together with those from other people, in order for us to make large-scale comparisons among different groups (for example, young people versus older people). The information obtained from this study may be presented at scientific conferences or in scientific journals, but your name will never appear in any public document. In order to ensure confidentiality, your name and other identifying information will be filed separately from the experimental data in our laboratory, and all data are stored in locked filing cabinets and on password-protected computers in our department.

Contact Information. If you have any questions or comments on the study, or if you wish a clarification of rights as a research participant, you can contact Dr. Patrick Davidson or the Human Research Ethics Committee at the number and address below.

Investigator

Dr. Patrick Davidson
Department of Psychology
University of Alberta
Edmonton, AB T6G 2E9
(780) 492-2237

Ethics Committee Member

Dr. Don Heth
Department of Psychology
University of Alberta
Edmonton, AB T6G 2E9
(780) 492-2662

Signature Section. This research project has been explained to me in detail, and I have been provided with a copy of the Informed Consent Form. I have been given an opportunity to ask questions concerning these procedures, and any questions that I have asked have been adequately answered. I understand that all information will be kept confidential and I will not be identified in any way in any publication of the results. I understand that I do not have to take part in this study, and my refusal to participate will involve no penalty or loss of rights to which I am entitled. I may withdraw from this study at any time and for any reason without penalty or loss of rights. I hereby consent to participate.

Name: _____

Signature: _____ Date: _____

Name of Person Obtaining Consent: _____

Signature: _____ Date: _____

Anmerkung: Die Einverständniserklärung für die Untersuchung mit kurzem Intervall (eine Sitzung) unterschied sich nur darin, dass nicht von zwei, sondern von einer Sitzung gesprochen wurde.

D.2 Aufklärungsformular



University
of
Alberta

Project: Reality monitoring in young and older people

Principal Investigator: Patrick Davidson, PhD; Department of Psychology, University of Alberta; (780) 492-2237. Additional investigators: Isabel Lindner and Leslie McGhan (780) 492-1277.

Thank you very much for taking part in our research. You completed a battery of psychology tests that we put together in order to study the effects of healthy aging on memory. To understand how memory is different between young and older people, we will compare the results of young and healthy older adults on these tasks.

Memory often declines with advancing age, but not all aspects of memory are affected equally. Previous research has shown that although healthy older adults are good at recognizing studied items, they are less accurate when recognizing the *context* associated with these items. For instance, one might recognize a face as being highly familiar (i.e., item memory), but be unable to recall the person's name or the situation in which they encountered this person (i.e., context memory). In our special paradigm, we have examined memory for the modality in which an action is encoded. For instance, in real life one might wonder if one has actually taken one's pills or merely thought about taking them. To study this kind of memory phenomenon (known as reality monitoring) in the laboratory, we asked you to perform some action statements but not others in the first session, and to indicate the encoding modality – whether you have performed it or not – in the second session. We also asked you to rate how confident you were in your source decisions. This allowed us to have a measure of your subjective experience. Our main goal is to find out whether older people consistently show problems with memory for context, or whether it depends on the exact way that we test memory.

You also completed a few tests of memory, problem-solving, vocabulary, mood, and so on within the two sessions, to help us find out whether any group differences that we find in memory are related to these other factors.

Remember, the results that we have collected are confidential. Data will be coded using either a participant number or your initials, and stored in a secured room. Only Dr. Davidson and those he authorizes will have access to the log linking participants' numbers or initials to names. Your name, and other identifying information, will NOT be used when we report our results in presentations and/or scientific publications.

The results from your testing session will be put together with those from other people, in order for us to make large-scale comparisons among different groups (for example, young people versus older people). We are unable to provide insight into a specific person's case, but if you are worried about your memory we suggest you contact your family doctor.

Thanks again for taking part. Feel free to contact our lab (492-1277) or our primary investigator, Professor Patrick Davidson (492-2237), should you have any questions or comments.

Anmerkung: Das Aufklärungsformular für die Untersuchung mit kurzem Intervall (eine Sitzung) unterschied sich nur darin, dass nicht von zwei, sondern von einer Sitzung gesprochen wurde.

D.3 Instruktionen

Beobachten – Abruf nach 20 Minuten (Experiment IV)

Dear participant, Please think back to the beginning of this session now. Just to remind you: In the first phase of this session, several action statements were presented and you were asked to just read or to read and actually perform them. Later, within the video-observation task, some of these action statements were re-presented. Now, action statements will be presented again. Your task is to indicate for each action statement if it was already presented within this session, i.e. if it is an old or a new action statement. If you indicate that an action statement was presented, you will furthermore be asked if you performed the action or if you didn't perform it, that is merely read the action statement. In addition, you will be asked to rate confidence concerning your memory judgments. Please don't hesitate to ask me if you have any questions!

Danke.

Meine beiden Doktorväter, Walter Hussy und Gerald Echterhoff, haben mich während der gesamten Zeit meiner Promotion fachlich und menschlich großartig betreut und hierdurch ganz maßgeblich zum Entstehen dieser Arbeit beigetragen. Walter Hussy danke ich besonders für die perfekte Balance zwischen dem Freiraum, den er mir ließ, und dem Halt, den er mir vermittelte. Bei Gerald Echterhoff möchte ich mich ganz besonders für seine große Begeisterung und für die zahlreichen, wichtigen Impulse, die er der Arbeit verliehen hat, bedanken. Ohne das absolute Vertrauen, das Sie beide mir entgegengebracht haben, gäbe es diese Arbeit nicht.

Patrick Davidson hatte immer ein offenes Ohr für meine Fragen und nahm sich viel Zeit für Diskussionen verschiedener Sachverhalte. Hilfreich waren aber auch sein steter Optimismus und sein reichlich vorhandener Humor. Esther Fujiwara hat mich auf eine so selbstverständliche Art und Weise in Edmonton aufgenommen und mich bei der Planung meines Forschungsaufenthaltes unterstützt, dass ich hiervon immer noch beeindruckt bin. Matthias Brand gab wichtige Anregungen gerade in der Anfangsphase der Promotion. Leslie McGhan half mir sehr bei der Rekrutierung der älteren TeilnehmerInnen – und bei der Bewältigung von Kauderwelsch.

Meine Kollegin Anita Jain stand mir unermüdlich mit Rat und Tat zur Seite und nahm sich jeglichen (menschlichen, praktischen, methodischen) Problems an – und das in jeder noch so stressigen Situation.

Gerhard Mutz hat mich aus purer Nettigkeit mit unglaublicher Geduld und besonderem Engagement bei der technischen Umsetzung der Untersuchungen unterstützt. Ich bin froh, dass ich ihm im Gegenzug durch die seltsamen Phänomene, die ich immer wieder erzeugte, und die Sonderwünsche, die ich immer wieder hatte, wenigstens bei der Optimierung seiner Software helfen konnte ;).

Insgesamt haben über 200 Personen an den geschilderten Untersuchungen teilgenommen. Mein besonderer Dank gebührt denjenigen TeilnehmerInnen, die mich aus reiner Hilfsbereitschaft unterstützt haben. Hierzu zählen v.a. die älteren TeilnehmerInnen der vierten Untersuchung, deren Einsatzbereitschaft und Interesse außergewöhnlich waren.

Zahlreiche Familienmitglieder bzw. Nahezu-Familienmitglieder haben mir bei der Erstellung der Videoaufnahmen geholfen. Ich weiß es sehr zu schätzen, dass sie dieser nicht sehr attraktiven Aufgabe mit Ausdauer nachkamen, Handtücher falteten und Perlen schüttelten, auch wenn es extrem anstrengend wurde („Und jetzt bitte noch einmal den Schwamm wringen – aber diesmal bitte mit beiden Händen!“). Als AkteurInnen stellten sich namentlich zur Verfügung Michael Mentzner, Alexandra Hill, Edith Dammann und Wilfried Mentzner. Technische Unterstützung beim Überspielen der Videos (und darüber hinaus) leisteten Christian Lindner und Volker Mentzner.

Meinen Eltern, Brigitte und Wilhelm Lindner, danke ich besonders für ihr uneingeschränktes Vertrauen, ihr Mit-Leiden und Mit-Freuen sowie ihre bedingungslose Unterstützung, ohne die u.a. mein Kanada-Aufenthalt nur ein Traum geblieben wäre. Bedanken möchte ich mich auch bei meiner Oma, Edith Dammann, die diese Zeit ebenfalls voll Anteilnahme mit mir durchlebt hat – und dies mit großer Begeisterung, Neugier und Interesse.

Mein Freund, Michael Mentzner, hat die Zeit meiner Promotion zweifellos am intensivsten miterlebt und ausnahmslos alle Höhen und Tiefen mit mir durchgestanden. Durch ihn habe ich die wunderbarste Unterstützung erfahren, die man durch einen anderen Menschen nur erfahren kann – von unendlicher Geduld auch bei der 3. Wiederholung des Videodrehs, weil das Licht immer noch nicht perfekt war, über verständnisvolle und motivierende Worte bis hin zu Lebensmittelanlieferungen kurz vor meiner Disputation ;). Um seinen Beitrag zu dieser Arbeit angemessen zu würdigen, wäre eine eigenständige Abhandlung notwendig. Daher sage ich an dieser Stelle einfach von ganzem Herzen: Danke für alles!!!

Und Danke auch allen anderen – KollegInnen, FreundInnen, (Nahezu-)Familienmitgliedern –, die hier nicht explizit erwähnt wurden, aber auch auf die eine oder andere Art und Weise für mich da waren und mich bei der Umsetzung meines Vorhabens unterstützt haben.