

Alters- und Geschlechtsdifferenzen in der Raumkognition - auch eine
Frage der Lösungsstrategien

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

genehmigt durch die Fakultät für Naturwissenschaften
der Otto-von-Guericke-Universität Magdeburg

von Dipl.-Informatiker Robby Schönfeld
(akademischer Grad, Vorname, Name/Geburtsname)

geb. am 03.07.1973 in Hohenmölsen

Gutachter: PD Dr. Wolfgang Lehmann
(akademischer Grad, Vorname, Name)

Prof. Dr. Claudia Quaiser-Pohl
(akademischer Grad, Vorname, Name)

eingereicht am: 25.08.2008

verteidigt am: 27.11.2008

Dank

Ich bedanke mich für die mir jederzeit zur Verfügung gestellte Unterstützung und Beratung, für kritische und nützliche Hinweise meiner Betreuer PD Dr. Wolfgang Lehmann und Prof. Dr. Bernd Lepow. Prof. Dr. Claudia Quaiser-Pohl danke ich für die Übernahme eines der Gutachten. Vielen Dank an Dr. med. Franz-Josef Müller für die initiale Idee zu den Musteranalysen.

Mein Dank gilt außerdem den Dipl.-Psychologinnen Jana Ungewiß, Nadine Mönich, Eileen Strich, Anke Woskowsky und Angelika Drauschke, die mich tatkräftig bei der Datenerhebung unterstützt haben. Sehr dankbar bin ich auch dem Leiter und den Lehrern des Georg-Cantor-Gymnasiums in Halle, der Leiterin der Sektion Rehasport im Universitäts-sportverein Halle über die einige Testungen organisiert und durchgeführt wurden.

Weiterhin danke ich Frau Dipl.-Psych. Yvonne Paelecke-Habermann und Frau cand. Psych. Anja Beinert für das abschließende Lektorat der Dissertationsschrift.

Mein besonderer Dank gilt meiner Ehefrau, Dipl.-Psychologin Sabine Schmidt, ohne die diese Arbeit niemals möglich gewesen wäre. Ich danke ihr für die fachliche, finanzielle und emotionale Unterstützung und das unermüdliche Korrekturlesen während der Entstehung der schriftlichen Arbeit.

Halle (Saale), im März 2009

Inhaltsverzeichnis

Einleitung	7
Theorieteil	9
1 Raumkognition und Lösungsstrategien	11
1.1 Einordnung und Begriff Raumkognition	11
1.2 Informationsverarbeitungsansatz	11
1.3 Strategiemodell	13
2 Räumliche Fähigkeiten und Präferenzen	15
2.1 Räumliche Fähigkeiten	15
2.1.1 Mentale Rotation	16
2.1.2 Place-Learning	17
2.1.3 Pfadintegration	19
2.2 Räumliche Präferenzen	21
2.2.1 Analytisch vs. holistisch	22
2.2.2 Egozentrisch vs. allozentrisch	23
2.2.3 Unimodal vs. bimodal	25
3 Alters- und Geschlechtsdifferenzen in der Raumkognition	26
3.1 Altersdifferenzen	26
3.2 Geschlechtsdifferenzen	28
4 Raumkognition in virtuellen Umgebungen	32
4.1 Räumliche Vorstellung virtueller Umgebungen	32
4.2 Virtuelle Umgebungen	33
4.3 Faktoren virtueller Präsenz	34
5 Zusammenfassung und Fragestellung	36
5.1 Zusammenfassung	36
5.2 Ableitung der Fragestellungen	37
5.2.1 Hypothesen zu den Fragestellungen	38
5.2.2 Fragen der Softwareentwicklung	41

Methodenteil	43
6 Softwareentwicklung	45
6.1 Computer Mental Rotations Test	45
6.1.1 Material und Aufbau	45
6.1.2 Protokoll	47
6.1.3 Punktvergabesysteme	48
6.2 Kompass-Zeigeaufgabe	48
6.2.1 Virtuelle Umgebung	48
6.2.2 Protokoll	50
6.2.3 Messwerte	50
6.3 Water-Maze-Aufgabe	51
6.3.1 Virtuelle Umgebungen	51
6.3.2 Protokoll	53
6.3.3 Merkmalsaggregation	55
7 Datenerhebung	57
7.1 Experiment 1 - Mentale Rotation	57
7.1.1 Stichprobe	57
7.1.2 Messinstrumente	58
7.1.2.1 Raumkognitive Maße	58
7.1.2.2 Fragebögen	59
7.1.2.3 Allgemeine kognitive Maße	60
7.1.3 Technische Apparatur	61
7.1.4 Durchführung	61
7.2 Experiment 2 - Pfadintegration	62
7.2.1 Stichprobe	62
7.2.2 Messinstrumente	63
7.2.2.1 Raumkognitive Maße	63
7.2.2.2 Allgemeine kognitive Maße	64
7.2.3 Technische Apparatur	65
7.2.4 Durchführung	65
7.3 Experiment 3 - Place-Learning	65
7.3.1 Stichprobe	65
7.3.2 Messinstrumente	66
7.3.2.1 Raumkognitive Maße	66
7.3.2.2 Fragebögen	68
7.3.2.3 Allgemeine kognitive Maße	68
7.3.3 Technische Apparatur	68
7.3.4 Durchführung	69
8 Datenanalyse	70

8.1	Deskriptive Verfahren und Inferenzstatistik	70
8.2	Exploratorische Verfahren der Musteranalysen	71
8.2.1	Nicht-negative Matrix-Faktorisierung	72
8.2.2	Consensus-Clustering	74
	Ergebnisteil	77
9	Experiment 1 - Mentale Rotation	79
9.1	Allgemeine kognitive Maße	79
9.2	Vorerfahrung	80
9.3	Raumkognitive Maße	81
9.3.1	Mental Rotations Test A	81
9.3.2	Würfelaufgaben	82
9.3.3	Computer Mental Rotations Test	83
9.4	Musteranalysen	87
9.4.1	Mental Rotations Test A	87
9.4.2	Würfelaufgaben	95
9.4.3	Zusammenhänge zwischen den Klasseneinteilungen	97
9.5	Zusammenhänge zu Strategiefragebögen	98
9.6	Zusammenfassung	101
10	Experiment 2 - Pfadintegration	104
10.1	Allgemeine kognitive Maße	104
10.2	Raumkognitive Maße	105
10.2.1	Raumvorstellungstests	105
10.2.2	Kompass-Zeigeaufgabe	106
10.2.3	Analyse zum Desorientierungseffekt	112
10.2.4	Zusammenhänge zwischen den Maßen	115
10.3	Zusammenfassung	116
11	Experiment 3 - Place-Learning	118
11.1	Allgemeine kognitive Maße	118
11.2	Vorerfahrung	119
11.3	Raumkognitive Maße	120
11.3.1	Raumvorstellungstests	120
11.3.2	Kompass-Zeigeaufgabe	122
11.3.3	Water-Maze-Aufgabe	124
11.3.4	Zusammenhänge zwischen den Maßen	132
11.4	Muster- und Klassenanalysen	134
11.4.1	Desorientierungseffekt der Kompass-Zeigeaufgabe	134
11.4.2	Lösungsstrategien des Mental Rotations Test	136

11.4.3	Suchstrategien der Water-Maze-Aufgabe	139
11.4.4	Zusammenhänge zwischen den Klasseneinteilungen	149
11.4.5	Zusammenhänge zu Strategiefragebögen	152
11.5	Zusammenfassung	156
 Diskussionsteil		159
12	Einzeldiskussionen	161
12.1	Experiment 1 - Mentale Rotation	161
12.2	Experiment 2 - Pfadintegration	168
12.3	Experiment 3 - Place-Learning	172
13	Gesamtdiskussion	181
13.1	Diskussion inhaltlicher Ziele	181
13.2	Diskussion methodischer Ziele	184
 Literaturverzeichnis		187
14	Tests und Fragebögen	198
14.1	Mental Rotations Test (revidierte Testform B)	199
14.2	Test- und Aufgabeninstruktionen	205
14.2.1	Mental Rotations Test	205
14.2.2	Kompass-Zeigeaufgabe	206
14.2.3	Virtuelles Training	207
14.2.4	Water-Maze-Aufgabe	208
14.3	Eigene Fragebögen	210
14.3.1	Demographische Daten	210
14.3.1.1	Experiment 1	210
14.3.1.2	Experiment 2 und 3	211
14.3.2	Computererfahrung und Vorerfahrung Raumvorstellungstest	212
14.3.2.1	Experiment 1	212
14.3.2.2	Experiment 3	213
14.3.3	Strategiefragebögen	215
14.3.3.1	Mental Rotations Test	215
14.3.3.2	Kompass-Zeigeaufgabe	217
14.3.3.3	Water-Maze-Aufgabe	218
15	Technische Details - Softwareentwicklung	220
15.1	Hardwarevoraussetzungen und Installation	220
15.2	Computer Mental Rotations Test	221
15.2.1	Konfiguration	221
15.2.2	Datenformat	224

15.3	Kompass-Zeigeaufgabe	225
15.3.1	Konfiguration	225
15.3.2	Datenformat	227
15.4	Water-Maze-Aufgabe	228
15.4.1	Konfiguration	228
15.4.2	Datenformat	230
15.4.3	Dataviewer	231
16	Technische Details - Datenanalyse	235
16.1	Matlab-Skripte	235
16.1.1	Nicht-negative Matrix-Faktorisierung	235
16.1.2	Consensus-Clustering	236
16.1.3	Vorverarbeitung der Spurdaten der Water-Maze-Aufgabe	237

Einleitung

Wir alle wenden zur Lösung der an uns gestellten Aufgaben in irgendeiner Form Strategien an. Das gilt auch und insbesondere für Aufgaben mit Anforderungen im räumlichen Kontext. Bei der Vorstellung von Objekten und Bewegungen laufen kontinuierlich räumliche Informationsverarbeitungsprozesse ab, die situativ und individuell verschieden sein können. An solchen alltäglichen Handlungen, wie das Auffinden von Wegen, hat jeder schon beobachtet, dass verschiedene Strategien existieren und sich Personen in deren Anwendung ganz offensichtlich unterscheiden.

In wissenschaftlichen Untersuchungen konnte der Einsatz individueller Lösungsstrategien während der Bearbeitung von Raumvorstellungstests nachgewiesen werden (Barratt, 1953; Putz-Osterloh, 1977; Hosenfeld, Strauss & Köller, 1997; Geiser, Lehmann & Eid, 2006). Für die psychometrischen Tests stellt deren Gebrauch aber ein methodisches Problem dar. Auch Raumvorstellungstests operationalisieren in der Mehrzahl eindimensionale Fähigkeitskonstrukte, über die die Leistung auf einer Absolutskala (i. d. R. Summenscores) gemessen wird. Der heterogene Einsatz von Strategien bedeutet nun, dass man durch die „Messung“ auf keinen Fall ein homogenes Verhalten beobachtet, das nur in der Leistungsfähigkeit variiert. Das reduziert die Konstruktvalidität des Messinstruments und somit auch die interne Validität der damit durchgeführten Untersuchung.

Eine Lösung des Problems besteht, im Ansatz der klassischen Testtheorie, in der Analyse und Verbesserung der psychometrischen Messinstrumente. Eine andere Möglichkeit ist die Ausnutzung der strukturellen Mängel in den Tests für eine qualitative Analyse mit einem klassifizierenden Ansatz, wie man ihn beispielsweise in der Item-Response-Theorie (IRT) verfolgt. Ein Vorteil der letzteren Methode ist, trotz des exploratorischen Charakters, die Identifikation weiterer, ansonsten unentdeckter, Störvariablen und deren Verwendung als Faktoren im varianzanalytischen Sinne.

Ein häufig untersuchter Faktor in Bezug auf räumliche Fähigkeiten ist die Geschlechtszugehörigkeit mit dem Ergebnis, dass männliche Versuchspersonen im Mittel bessere Testleistungen erzielen als weibliche Probanden (Kimura, 1999; Linn & Petersen, 1985). In Arbeiten, die einen klassifizierenden Ansatz zur Auswertung der Raumvorstellungstests wählten, konnten eindeutige Zusammenhänge zwischen Testleistung, Geschlecht und Effizienz der verwendeten Strategie aufgezeigt werden (Geiser et al., 2006). Darüber hinaus konnte die Identifikation latenter Subgruppen, deren Personen sich durch die Anwendung unterschiedlicher Lösungsstrategien auszeichnen, auch zur Beurteilung der Validität des eingesetzten Testverfahrens beitragen.

Ein weiteres Problem besteht aber dennoch in der Generalisierbarkeit der Testergebnisse auf das Verhalten in tatsächlichen räumlichen Situationen. Hier sind Raumvorstellungstests häufig der Kritik einer geringen oder gar fehlenden ökologischen Validität ausgesetzt. Der Implementierung ökologisch valider Testsituationen zur Operationalisierung entsprechender Fähigkeitskonstrukte wid-

meten deshalb zahlreiche Wissenschaftler ein besonderes Augenmerk. Seit zirka zehn bis 15 Jahren werden dazu intensiv virtuelle Umgebungen eingesetzt. Diese nutzen technische Möglichkeiten hochentwickelter Computergrafiksysteme zur Simulation einer real nicht-existierenden Umwelt, durch die die Versuchspersonen navigieren.

Die entwickelten virtuelle Umgebungen variieren zwischen hoch-realistischen Nachbauten komplexer Umgebungen (vgl. Bühlhoff & van Veen, 1999) und Adaptionen artifizierlicher Versuchsanordnungen aus Tierexperimenten (z. B. Astur, Ortiz & Sutherland, 1998). Im Vergleich zu den „klassischen“ Raumvorstellungstests gibt es hier bisher keine Standards. Jede Arbeitsgruppe verwendet üblicherweise ihre eigenen selbstentwickelten Programme, die auf ein bestimmtes Untersuchungsdesign zugeschnitten sind. Das schränkt die Vergleichbarkeit der Ergebnisse ein. Zusätzlich bestehen nicht genau kontrollierbare Interaktionen zwischen Versuchsperson und der verwendeten Computertechnik, die sich auf die zu untersuchenden differentiellen Effekte auswirken können. Theoretisch stellt sich darüber hinaus die Frage, welche räumlichen Fähigkeiten mit virtuellen Versuchsanordnungen operationalisiert werden, wenn der zu erfahrende Raum real gar nicht existiert.

Hier besteht nun, aus eigener Sicht, ein ganz neuartiges Problem, bei dem es um die Implementierung und Überprüfung ergometrischer Software geht, die Kriterien valider psychometrischer Verfahren genügt. Diese nicht triviale und komplexe Problematik liegt genau im Schnittpunkt zwischen Informatik und Psychologie und wird in den meisten Untersuchungen nicht explizit angesprochen.

Das Ziel der Arbeit ist deshalb zunächst, rein methodisch, die Implementierung der virtuellen Umgebungen als flexible, für verschiedene Untersuchungsdesigns anwendbare, Versuchsanordnungen. Die zu operationalisierenden räumlichen Fähigkeiten werden auf mentale Rotation, Pfadintegration und Place-Learning begrenzt. Die Auswahl erfolgte auf Grund der weiteren Zielsetzung der Arbeit: der Untersuchung differentieller Effekte in der Raumkognition. Zu allen drei Teilbereichen sind in der Literatur häufig quantitative Geschlechts- bzw. Altersunterschiede berichtet worden. Diese konnten aber bereits in Zusammenhang gebracht werden mit der Anwendung gruppenspezifischer Lösungsstrategien, so dass der Kern der Arbeit in der Untersuchung dieser qualitativen Unterschiede liegt.

Dazu werden im Rahmen der Arbeit zu den Modellen der IRT alternative Datenanalysetechniken für die Identifikation von Lösungsstrategien vorgestellt. Diese sind generische klassifizierende Verfahren aus dem Forschungsbereich der Bioinformatik, basierend auf der automatischen Erkennung von Mustern (Lee & Seung, 1999; Monti, Tamayo, Mesirov & Golub, 2003; Gao & Church, 2005). Diese Entscheidung begründet sich in der Art der hier zur Strategieanalyse anfallenden Daten. Zum einen handelt es sich um kleinere, für die Modelle der IRT unzureichende Datensätze dichotomer Itemantworten aus Raumvorstellungstests, und zum anderen um Spurdaten aus Navigationsexperimenten.

Abschließend sollen dann die bereits berichteten Zusammenhänge zwischen Art der verwendeten Strategie und der Zugehörigkeit zu Geschlechts- bzw. Altersgruppen untersucht werden. Darüber hinaus wird der Frage nachgegangen, wie sich Strategien in verschiedenen Testsituationen mit räumlichen Anforderungen etablieren, d. h. ob diese Strategien eher situations- oder eher personenspezifisch eingesetzt werden.

Theorieteil

1 Raumkognition und Lösungsstrategien

1.1 Einordnung und Begriff Raumkognition

Raumkognition umfasst multiple Fähigkeiten, den umgebenden Raum, Lage und Position darin enthaltener Objekte und der eigenen Person wahrzunehmen, mental zu repräsentieren und zu nutzen. Für eine Betrachtung aus der Sicht der Psychologie ordnet sie sich als Forschungsrichtung in die allgemeineren Themen Wahrnehmung, Lernen und Gedächtnis ein.

Auf der Seite der Wahrnehmung versteht man darunter alle mentalen Prozesse, die externe und interne Informationen über den umgebenden Raum, Lage und Position von Objekte und des eigenen Körpers verarbeiten. Räumliche Wahrnehmung ist in diesem Sinne eine aus der Aufnahme und Verarbeitung anderer Sinnesmodalitäten abgeleitete Wahrnehmung. Demnach sind räumliche Stimuli multimodale Kombinationen aus visuellen, auditiven, vestibulären, propriozeptiven und haptisch-taktilen Reizen.

In der mnestischen Komponente umfasst Raumkognition den Erwerb und die Organisation von Informationen über räumliche Umgebungen und Strukturen. Die Primitiven räumlicher Informationen sind Positionen, Richtungen und Distanzen. Wie diese Informationen im Gehirn gespeichert werden, wird anhand unterschiedlicher Modelle diskutiert. Zur Nutzung dieser mentalen Repräsentation in einem natürlichen Kontext räumlichen Verhaltens zählen Wegesuchen und Navigieren. Spezielle Aspekte sind dabei z. B. Explorations- und Heimkehrverhalten. Beim Menschen wird dazu noch allgemeiner das räumliche Denken gezählt, zu dem räumliche Vorstellungen sowie der Umgang mit räumlichen Wissen gehören (Freksa, Bauer, Habel & Wender, 2000).

1.2 Informationsverarbeitungsansatz

Die Repräsentation räumlicher Information im Gehirn sollte nun nicht nur als unmittelbare Speicherung der Primitiven verstanden werden, obwohl auch dafür direkte neuronale Abbildungen existieren (Samsonovich & McNaughton, 1997). Grundsätzlich kann aber davon ausgegangen werden, dass nicht alle räumlichen Informationen dauerhaft erhalten bleiben. Sie unterliegen wie andere Informationen auch zeitlichen Zerfallsprozessen, die sich in das Prozessebenenmodell nach Atkinson und Shiffrin (1968) einordnen lassen.

Innerhalb der drei größeren Verarbeitungseinheiten sensorischer Puffer, Kurz- und Langzeitgedächtnis müssen auch die räumlichen Informationen gehalten werden. Der sensorische Puffer ist für

räumliche Informationen relevant. Hier finden Merkmalsanalysen und Modalitätstrennungen statt. Die komplexen Signalmuster werden also zunächst separiert bevor sie weiterverarbeitet werden. Im Kurzzeitgedächtnis geschieht dann ein Teil dieser Weiterverarbeitung. Da dieser Speicher neue und alte Informationen mit dem Ziel der Verhaltenssteuerung bewusst bearbeitet, wird er oft analog als Arbeitsgedächtnis bezeichnet.

Im Arbeitsgedächtnismodell von Baddeley (2000) wird, spätestens durch die Ergänzung des episodischen Puffers, von einer Dualität der Speicherung räumlicher Informationen ausgegangen. Mit der Existenz eines Speichers für visuell-räumliche Skizzen wird dem statischen Charakter räumlicher Informationen Rechnung getragen. Hier handelt es sich tatsächlich um die Encodierung von Lagebeziehungen in Form von Positionen und Richtungen im Raum. Außerdem wird die enge Beziehung zwischen räumlichen Information und dem visuellen System betont. Mit dem Postulat eines episodischen Puffers wird dagegen Wert gelegt auf die Speicherung von Ereignissen, die in einem zeitlichen und räumlichen Kontext stattfanden. Hier werden räumliche Informationen als dynamisch und kontextbereitend aufgefasst.

Die Debatte, ob räumliche Information als Kontextinformation oder in ihren Primitiven repräsentiert sind, wird besonders heftig in der Frage nach ihrer langfristigen Speicherung diskutiert. Wenigstens fünf allgemeinere Langzeitgedächtnis-Theorien werten aktuell dazu vertreten (Bird & Burgess, 2008). Sie alle akzeptieren Tulvings Einteilung in episodische und semantische Gedächtnisinhalte (Tulving, 1972). Räumliche Informationen werden als kontextuelle Informationen im episodischen Gedächtnis gespeichert. Unterschiedlich sind die Auffassungen der Theorien über die neuronalen Mechanismen beim Erinnern räumlich-kontextueller Information (s. Tab. 1.1).

Tabelle 1.1: Übersicht aktueller Langzeitgedächtnis-Theorien und deren Auffassung zur Speicherung räumlicher Informationen.

Theorie des deklarativen Gedächtnisses (Squire, 1986)	Bewusste, abrufbare Speicherprozesse verwalten räumliche Informationen. Nach der Konsolidierung sind diese ausschließlich neokortikal gespeichert. Es gibt keine besonderen Unterschiede zwischen räumlichen und anderen Gedächtnisspuren.
Multiple-Trace-Theorie (Moscovitch et al., 2006)	Die Speichermechanismen nutzen räumliche Informationen als Kontextinformation für episodische Inhalte. Diese unterscheiden sich dadurch von semantischen Inhalten. Kontextinformationen bleiben zeitlebens in einem Netzwerk aus medialem Temporallappen, Hippocampus und anderen kortikalen Arealen repräsentiert.
Dual-Prozess-Theorie (Yonelinas, 2002)	Räumliche Informationen sind kontextuelle Details. Sie sind in einem hippocampalen Netzwerk dauerhaft an episodische Inhalte geknüpft und können bewusste Erinnerungen hervorrufen oder durch diese hervorgerufen werden, d.h. sie haben assoziativen Charakter beim Erinnern von Ereignissen.
Theorie des Relationalen Lernens (Eichenbaum, 2000)	Die Speicherung räumlicher Informationen stellt einen Spezialfall relationaler Verknüpfungen dar. Inhalte einzelner räumlicher Szenen werden modalitätsabhängig in neokortikalen Arealen encodiert. Der Hippocampus verknüpft die verschiedenen Inhalte relational miteinander zu einer Erinnerung.

Wie auch immer räumliche Informationen neuronal gespeichert werden, ob als statische Primitiven oder dynamische Kontextinformationen, auf dem Weg der Informationsverarbeitung von der sensorischen Aufnahme bis zur langfristigen Speicherung erfolgt, ausgehend von einer Separation der Modalitäten eines räumlichen Signals, eine wieder zunehmende Integration dieser. Aus der Beschaffenheit der mnestischen Systeme, die in diesem Verarbeitungsprozess beteiligt sind, ergeben sich einige ernsthafte Probleme. Die Begrenzung des Arbeitsgedächtnisses auf wenige Items (Miller, 1956) stellt einen Engpass bei der bewussten Bearbeitung von Informationen dar. So sind dadurch möglicherweise die räumliche Vorstellung von Objekten und deren Transformation limitiert. Eine Steigerung der Effizienz ist nur durch Erhöhung des Informationsgehalts pro Item (Chunkbildung) und den Einsatz von Strategien möglich. Durch eine Automatisierung dieser Strategien besteht im besten Fall praktisch kein Speicherplatzbedarf.

Beim Langzeitgedächtnis besteht diese Limitierung nicht. Durch den assoziativen Charakter binden neue Informationen besser an vorhandenes Wissen, d. h. je mehr Inhalte vorhanden sind, um so leichter können neue daran geknüpft werden. Dadurch ist die Kapazität praktisch unbegrenzt, woraus sich die Notwendigkeit einer Organisation des Wissens ergibt. Beim Aufbau dieser Wissensstrukturen spielen ebenfalls Strategien eine entscheidende Rolle. Diese können beispielsweise Informationen selektieren, gruppieren, in Kategorien einteilen bzw. in Hierarchien einordnen.

1.3 Strategiemodell

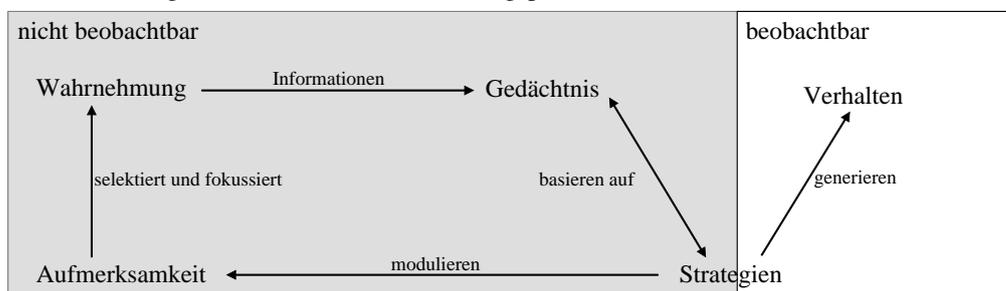
Allgemein werden Strategien aufgabenbezogen und bewusst eingesetzt (Naus & Ornstein, 1983). Sie dienen der Steigerung der Effizienz mnestischer Systeme, der Entscheidungsfindung, der Verhaltenssteuerung und der Automatisierung. Erfasst werden können Strategien durch Verhaltensbeobachtungen über viele Situationen hinweg und der Entdeckung möglicher Gemeinsamkeiten in der Performanz verschiedener Individuen. Diese Verhaltensinvarianten können als Ausdruck einer gemeinsamen Strategie interpretiert werden.

Räumliche Strategien werden speziell angewendet während der zielgerichteten Antwort auf eine Umgebung. Dabei werden in den Verhaltensantworten beobachtbare und unterscheidbare Muster erzeugt. Eine räumliche Strategie ist eine Komposition aus diesen Verhaltensmustern. Auf der Basis dieser Muster können wiederum die Strategien voneinander unterschieden werden. Entgegen der allgemeinen Definition von Strategien führt bei einer räumlichen Strategie auch der nicht-bewusste Einsatz meist zu systematischen Verhaltensmustern. Deshalb zählen diese zu den räumlichen Strategien dazu.

Räumliches Verhalten in natürlichen Umgebungen beinhaltet eine besondere Art der räumlichen Strategie, die Suchstrategie (Kallaia, Makany, Karadia & Jacobs, 2005). Diese besteht aus wiederholten Bewegungsmustern, die unterscheidbare Teile der Gesamtbewegung bilden. So sind Bewegungsmuster qualitative Daten, die in der Trajektorie der Bewegung unterschiedlich stark repräsentiert sind. Die Suchstrategien bilden somit die Hauptkomponenten der Trajektorie.

Im Informationsverarbeitungsprozess stehen Strategien zwischen dem nicht beobachtbaren internen Prozessen, wie Wahrnehmung, Gedächtnis, Aufmerksamkeit etc. und dem beobachtbaren Verhalten. In einem stark vereinfachten Modell wird angenommen, dass Strategien in enger Wechselwirkung mit der im Gedächtnis abgelegten mentalen Repräsentation der Umwelt stehen (s. Abb. 1.1). Über die Wahrnehmung gelangen Informationen in mnestiche Strukturen, die aufgrund von Kapazitätslimits und der Notwendigkeit der Strukturierung den Einsatz von Strategien erfordern. Die Strategien sind in Form von Metawissen selbst Bestandteil des Gedächtnisses und wären ohne dieses nicht existent. Auf das Verhalten üben sie eine generierende Wirkung aus, indem wiederholt planvoll Handlungsmuster aufgerufen werden. Intern können Strategien eine rückkoppelnde Wirkung auf die Wahrnehmung haben, indem die Aufmerksamkeit selektiv und fokussiert auf handlungsrelevante Details der Umgebung gelenkt wird.

Abbildung 1.1: Modell der aufmerksamkeitsmodulierenden und verhaltensgenerierenden Bedeutung von Strategien im Informationsverarbeitungsprozess.



Für die individuelle Auswahl der Strategien wird im Modell davon ausgegangen, dass diese intern und extern getriggert werden. Letzteres wird durch geeignete Hinweisreize aus der Umwelt hervorgerufen. Interne Triggerung basiert auf vielfältigen mentalen Repräsentation, wie z. B. Erwartungshaltungen und der Verfügbarkeit einer Strategie. Ein Ansatz, diese internen Auswahlprozesse zu verstehen, ist die Annahme, dass die Wahl einer Strategie sowohl von Kompetenzen als auch von Präferenzen abhängt. Kompetenzen sind eindimensionale Fähigkeiten eines Individuums, die quantitativ bewertet werden. Sie können durch psychometrische Verfahren auf einer unipolaren Skala abgebildet, d. h. gemessen werden. Eine bessere Testleistung ist direkt interpretierbar als eine höhere Ausprägung der Fähigkeit.

Präferenzen dagegen stellen Neigungen oder Stile einer Person dar, eben eine „typische Art“ der Informationsverarbeitung, die u. a. in Abhängigkeit von Persönlichkeitsmerkmalen für verschiedene Fähigkeiten zutreffen kann. Sie sind als theoretisches Konstrukt oft als Ausprägung zweier komplementärer Eigenschaften auf einer bipolaren Skala formuliert. Das Ergebnis aus einem entsprechenden Test wird qualitativ wertfrei interpretiert, in der Art, dass bei einer Person eine stärkere Ausprägung der einen gegenüber einer schwächeren Ausprägung der anderen Eigenschaft vorliegt (in Tiedemann, 1988 und Messick, 1984; zit. n. Hergovich, 1999).

2 Räumliche Fähigkeiten und Präferenzen

2.1 Räumliche Fähigkeiten

In der Differenzierung eindimensionaler räumlicher Fähigkeiten (Kompetenzen) werden im wesentlichen zwei Ansätze verfolgt. Der Ansatz mehrfaktorieller Fähigkeitsmodelle, wie sie in der Intelligenzforschung üblich sind (z. B. Thurstone & Thurstone, 1941), oder über Defizitmodelle, die primär im klinischen Anwendungsbereich und z. T. in der Verhaltensforschung zu finden sind.

Ausgehend von einer Beurteilung beobachtbarer Störungen räumlicher Fähigkeiten können räumlich-perzeptive, räumlich-konstruktive, räumlich-kognitive und räumlich-topographische Fähigkeiten unterschieden werden (Kerkhoff, 2002). Räumlich-perzeptive Fähigkeiten umfassen Komponenten der Raumwahrnehmung mit verschiedenen Sinnesmodalitäten. Zu den räumlich-konstruktiven Fähigkeiten zählt man Fähigkeiten, wie Teil-Ganzes-Analysen, mentale Segmentierung und die Fähigkeit zum perspektivischen Zeichnen. Unter räumlich-kognitiven Fähigkeiten fasst man mentales Rotieren und die Fähigkeit zum mentalen räumlichen Perspektivwechsel zusammen. Räumlich-topographische Fähigkeiten basieren auf Wissen über den umgebenden Raum oder räumliche Anordnungen.

Die meisten Mehr-Faktoren-Modelle gehen von einem Raumvorstellungskonstrukt aus, in dem drei bis fünf Fähigkeitsfacetten unterschieden werden können (Thurstone, 1950; Linn & Petersen, 1985)). Maier (1996) geht von einem Fünf-Faktoren-Modell aus mit den Facetten Veranschaulichung, räumliche Wahrnehmung, mentale Rotation, räumliche Relationen und räumliche Orientierung. Die Wahrnehmung von Objekten in Verbindung mit der Vorstellung vom eigenen Körper (Körperschema) bildet die Facette der räumlichen Wahrnehmung. Veranschaulichung beschreibt die Fähigkeit, sich Veränderungen von Objekten (Faltungen und Schnitte) und deren Bewegungen im Raum (Verschiebungen und Drehungen) vorstellen zu können. Mentale Rotation bezieht sich auf die Vorstellung einer mentalen Abbildung von Figuren nach deren Drehung (Rotation). Räumliche Relationen beziehen sich auf das Erkennen von Konfigurationen im Raum und innerhalb zusammengesetzter Objekte. Räumliche Orientierung bedarf dagegen der realen oder mentalen Einordnung der eigenen Person in eine räumliche Situation, d. h. „Sich-Zurechtfinden“ im umgebenden Raum.

Der Bereich räumlich-kognitive Fähigkeit ist auf Grund der Möglichkeit zur Anwendung vielfältiger Strategien im folgenden besonders relevant. Ausgewählt wurden mentale Rotation, die als eigenständige Facette der Raumwahrnehmung auch verstärkt räumlich-perzeptive Fähigkeiten erfordert, Place-Learning und Pfadintegration, die zusätzlich auch räumlich-topographische Fähigkeiten in sich vereinen.

2.1.1 Mentale Rotation

Mentale Rotation beschreibt primär die Fähigkeit, sich schnell und exakt Rotationen zweidimensionaler Figuren und dreidimensionaler Objekte vorzustellen. Es handelt sich also um die mentale Abbildung von Objekten und deren Transformation vor dem „inneren Auge“. Sekundär sind deshalb mindestens auch räumlich-perzeptive Fähigkeiten und visuell-räumliche Gedächtnisleistungen damit verbunden. Es handelt sich also um eine multidimensionale Fähigkeit aus dem Zusammenspiel von räumlicher Wahrnehmung und dem räumlichen Arbeitsgedächtnis. Neuronale Netzwerke, die bei der Verarbeitung räumlicher Information während der mentalen Rotation eine Rolle spielen könnten, wurden konsistent zu dieser Annahme primär im visuellen und parietalen Kortex lokalisiert (Jordan, Heinze, Lutz, Kanowski & Jäncke, 2001). Entlang des dorsalen Pfades (Ungerleider & Mishkin, 1982) wird ein hypothetischer Verarbeitungsweg vom visuellen über den extrastriaten und inferior-parietalen bis zum prämotorischen Kortex angenommen (Jordan, Wüstenberg, Heinze, Peters & Jäncke, 2002).

Dass in diesen Hirnarealen ein physiologisches Korrelat zur mentalen Repräsentation und Rotation räumlicher Stimuli vorliegt, belegen mehrere Beobachtungen. Erstens besteht ein linearer Zusammenhang zwischen Reaktionszeit und Winkel der durchzuführenden Rotation (Shepard & Metzler, 1971). Zweitens konnten Keehner, Guerin, Miller, Turk und Hegarty (2006) einen linearen Trend zwischen Hirnaktivität im rechten superior-parietalen Sulcus (BA 40) und rechten dorsolateralen-präfrontalen Kortex (BA 46) mit dem Winkel der Rotation nachweisen. Drittens rekrutieren komplexere Stimuli ausgedehntere Anteile der entsprechenden Areale (Jordan et al., 2001).

Unter der Annahme einer begrenzten Kapazität des räumlichen Arbeitsgedächtnisses wird letzteres relevant für den Gebrauch effizienter Strategien zur mentalen Rotation. So genügt es bei einfachen zweidimensionalen Abbildungen, wie sie der *Card Rotation Test* (Thurstone, 1950) oder die Aufgabe *Spiegelbilder* im Wilde-Intelligenztest (Jäger & Althoff, 1994) verwenden, wenige Merkmale zu extrahieren und zu encodieren, um die mentale Rotation unter geringer Speicherbelastung ausführen zu können.

Komplexe Stimuli, wie die dreidimensionalen Abbildungen der Shepard-Metzler-Figuren im *Mental Rotations Test* (Peters et al., 1995), führen öfter dazu, dass Versuchspersonen sich in der Art der verwendeten Strategie unterscheiden (Geiser et al., 2006). Das liegt zum einen daran, dass es mehrere Möglichkeiten gibt, das Speicherkapazitätsproblem zu lösen, und zum anderen daran, dass die komplexeren Muster der Stimuli unterschiedliche Interpretationen der Aufgabe zulassen.

Eine Interpretationsmöglichkeit der mentalen Rotation ist die Vorstellung eines Perspektivwechsels (Shelton & McNamara, 2004; Hegarty & Waller, 2004). Obwohl bei diesem gleiche kortikale Areale aktiviert werden, fanden Keehner et al. (2006) ein zur mentalen Rotation inverses Aktivitätsmuster im superior-parietalen Sulcus. Außerdem berichteten Probanden heterogene Strategien in der Bearbeitung der Aufgaben zum Perspektivwechsel aber nicht in der Rotationsaufgabe. Es stellte sich heraus, dass nur einige Versuchspersonen mentale Rotation als Basis für den Perspektivwechsel nutzen. Zusätzlich ist nämlich für die Bewältigung der Aufgabe eine räumliche Vorstellung von

der eigenen Position und Blickrichtung in Bezug auf die Position des zu rotierenden Objekts notwendig. Diese Vorstellung hängt eng mit der Fähigkeit zum Wechsel zwischen Referenzsystemen zusammen (Hegarty & Waller, 2004).

2.1.2 Place-Learning

Mit dem Begriff Place-Learning wird allgemein eine Fähigkeit zur Orientierung im Raum bezeichnet, durch wiederholtes Aufsuchen einer Position diese mental zu repräsentieren. Die mentale Repräsentation erfolgt dabei in Bezug auf Hinweisreize aus der Umgebung. Diese können weithin sichtbare distale Landmarken, proximale Wegweiser aus der unmittelbaren Nähe oder aus der Geometrie des Raumes sein. Während des Place-Learnings werden, wie in jedem Lernprozess, Wissenstrukturen aufgebaut. Dieser Aufbau folgt dem Schema der Aufnahme neuen Wissens durch Exploration einer Umwelt, der Konsolidierung durch wiederholtes Aufsuchen und des Abrufs durch Wiedererkennen unter veränderten Bedingungen.

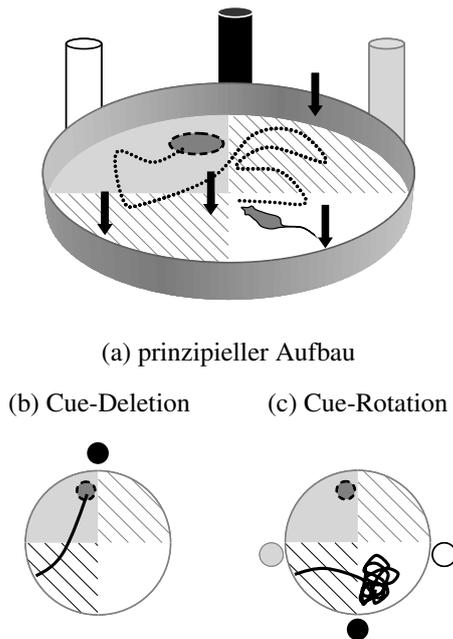
Wie stabil dieses Wissen bezüglich Veränderungen ist hängt primär von den verwendeten Umgebungshinweisen ab. Die Cognitive-Map-Theorie unterscheidet dahingehend zwei Arten von Raumwissen, Überblickswissen und Routenwissen (O'Keefe & Nadel, 1978). Diese Theorie ordnet sich als Spezialfall und Erweiterung für räumliche Informationen in die Theorie des Relationalen Lernens ein (Eichenbaum, 2000). Nach ihrem, für die Raumkognition grundlegenden Modell, wird Überblickswissen in einer kognitiven Landkarte repräsentiert. In diese werden Orte und mehrere Landmarken in einem globalen Referenzrahmen in Relation zueinander gesetzt. Ist eine bestimmte Landmarke nicht mehr sichtbar, können problemlos andere zur Orientierung genutzt werden. Bei Veränderungen der Umgebung werden neue Informationen in das existierende Referenzsystem integriert. O'Keefe und Nadel (1978) bezeichnen das dafür notwendige Orientierungssystem als Place-Learning-Mapping-System und sehen den Hippocampus als dessen neuronales Substrat. Spezielle Ortszellen im Hippocampus werden dabei mit rezeptiven Feldern im Raum verknüpft und bilden so die Umgebung in einer Art Koordinatensystem ab.

Der Begriff Place-Learning wird meist synonym mit der Fähigkeit zum Erwerb einer kognitiven Karte verwendet. Das global Referenzsystem basiert hier auf der Verwendung einer allozentrischen, d. h. von der eigenen Position unabhängigen mentalen Repräsentation des umgebenden Raumes. Sind die mentalen Repräsentationen aber abhängig von der eigenen Position und Lage, handelt es sich um egozentrische Referenzsysteme. Diese Arten von Wissen fassen O'Keefe und Nadel (1978) unter dem Begriff Routenwissen zusammen. Das entsprechende neuro-biologische System bezeichnen sie als Route-Learning-System, dass durch extrahippocampale Hirnstrukturen, hauptsächlich durch Striatum und Thalamus, gebildet wird. Zu diesem Routensystem zählen im Wesentlichen zwei räumliche Verhaltensweisen.

Sequence-Learning ist eine auf erlernten, z. T. automatisierten, Bewegungssequenzen basierende Orientierung. In ihr werden einmal abstrahierte Handlungsabläufe, wie „gerade aus und dann rechts“, in strikter und rigider Weise abgearbeitet. Reiz-Reaktionsverhalten, oder Cue-Response,

setzt dagegen schon die Einbeziehung lokaler Hinweisreize im verwendeten Referenzsystem voraus. Die entsprechenden Handlungen beziehen sich auf die Annäherung oder die Vermeidung bestimmter Orte in der Umgebung, die durch relevante Hinweise gekennzeichnet sind. Diese „Cues“ oder „Beacons“ erlauben das Erlernen von Reiz-Reaktionsketten, wie „an der alten Eiche rechts abbiegen“. Die egozentrischen Referenzsysteme sind rigider und anfälliger für Interferenzen, da sie bei Desorientierung, z. B. durch Richtungswechsel, nicht auf das von der eigenen Position unabhängige Überblickswissen der kognitiven Karte zurückgreifen können.

Abbildung 2.1: Place-Learning im Morris-Water-Maze.



Der Hidden-Plattform-Test im Water-Maze ist in der Verhaltensforschung eine Standardprozedur für Mäuse und Ratten. Innerhalb eines runden Tanks befindet sich eine im trüben Wasser versteckte Plattform.

(a) Während des Trainings lernt das Tier von verschiedenen Startpositionen aus (Pfeile in der Abbildung), diese auf direktem Weg anzuschwimmen. Dabei encodiert es deren Position in Relation zu distalen Landmarken (weiße, schwarze und graue Säulen bzw. Kreise im Bild).

(b) Der globale Referenzrahmen bleibt auch beim Löschen einiger Landmarken erhalten, was sich in einem direkten Anschwimmverhalten zeigt.

(c) Bei Veränderung der räumlichen Relationen zwischen Ziel und Landmarken wird die Plattform an einer falschen Position gesucht. Im dazugehörigen Quadranten beobachtet man ein intensives Suchverhalten.

Zur Differenzierung zwischen egozentrischen oder allozentrischen Raumrepräsentationen werden auf der Ebene der Verhaltensbeobachtung häufig Labyrinthaufgaben eingesetzt. Eines der Standardparadigmen in der Überprüfung der Place-Learning-Fähigkeit ist das Morris-Water-Maze (MWM; Morris, 1984). Darin wird dem Versuchstier ein globaler Referenzrahmen für das Auffinden einer verdeckten Zielposition anhand distaler Landmarken bereitgestellt (s. Abb. 2.1 a). Durch wiederholtes Aufsuchen von verschiedenen Startpositionen erlernen die Versuchstiere, sich in direkter Trajektorie zum Ziel zu bewegen. Das funktioniert, solange der globale Referenzrahmen nicht zu stark gestört wird. Selbst das Löschen einiger Landmarken ergibt bei Vorhandensein von stabilem Überblickswissen keine Interferenzen des Verhaltens mit der veränderten Umwelt (s. Abb. 2.1 b). Erst bei drastischen Eingriffen zeigt auch das globale Referenzsystem Störungen (s. Abb. 2.1 c).

Zahlreiche Belege existieren, dass eine Beeinträchtigung hippocampaler Hirnstrukturen einen vollständigen oder wenigstens partiellen Verlust dieser räumlichen Fähigkeit verursacht (für einen Überblick s. D’Hooge & De Deyn, 2001). Genau darin besteht ein Problem in der Interpretation des Water-Maze-Paradigmas. Das MWM ist im Grunde genommen die Implementierung eines Defizitmodells (Morris, Garrud, Rawlins & O’Keefe, 1982). Zeigen Tiere nach einer Deaktivierung des

Hippocampus nicht mehr die bekannten, für Place-Learning typische Orientierungsmuster, kann davon ausgegangen werden, dass die Hirnstruktur eine wichtige Rolle in der Aufrechterhaltung dieser Fähigkeit spielt. Bei einem nicht beobachteten Defizit sind andererseits keine sicheren Aussagen darüber möglich, welche mentalen Repräsentationen und Informationsverarbeitungsprozesse das Tier zur Verhaltensgenerierung nutzte.

So könnte das unter Abbildung 2.1 b dargestellte direkte Anschwimmen auch als Reiz-Reaktionsverhalten in Bezug auf die verbliebene Landmarke interpretiert werden, und nicht als bewusste Orientierungsreaktion anhand einer kognitiven Karte. Es ist möglich, dass während der wiederholten Lerndurchgänge von verschiedenen Startpositionen aus mehrere potentielle Routen encodiert und bei Bedarf als Handlungssequenzen abgerufen werden. Belege für diese Hypothese liefern schlechtere Testleistungen im MWM, wenn während der Lerndurchgänge bestimmte Bereiche im Maze nicht betreten werden durften (Hamilton, Driscoll & Sutherland, 2002).

Eine zweite Schwierigkeit in der Interpretation von Verhalten im Morris-Water-Maze besteht in der Multidimensionalität der Aufgabe. Neben der Erfassung räumlich-mnestischer Fähigkeiten wird im MWM auch eine Navigationsaufgabe operationalisiert. In dieser sind automatische, motorische Komponenten und Informationsverarbeitungsprozesse über das Routensystem enthalten. Bildgebende Studien zeigen zwar die Beteiligung des Hippocampus und angrenzender parahippocampaler Areale bei Navigationsaufgaben und insbesondere beim Auffinden neuer Wege und Abkürzungen (Maguire et al., 1998), dies muss aber tatsächlich nicht mit dem Abruf einer kognitiven Karte zusammenhängen. Wie Spiers und Maguire (2004) zeigen konnten, sind parahippocampale Kortexareale an einem impliziten Mapping von salienten lokalen Landmarken beteiligt. Diese erlauben während der Navigation Entscheidungsprozesse im Sinne eines Reiz-Reaktionsverhaltens, obwohl keine bewussten Erinnerungen an diese Landmarken existieren.

Neben der dauerhaften und expliziten kognitiven Karte werden demnach während des Place-Learnings auch dynamische, egozentrische und implizite Wissenstrukturen aufgebaut. Diese Art Online-Repräsentationen sind notwendig, um Place-Learning Verhalten im MWM vollständig beschreiben zu können.

2.1.3 Pfadintegration

In der Verarbeitung sensorischer Informationen zu implizitem Raumwissen über Richtungen, Geschwindigkeiten, Raumgeometrien und Positionen sehen Pfadintegrationsmodelle die eigentliche Aufgabe der Ortszellen und weiterer spezieller Zellverbände im Hippocampus (Richtungs- und Gitterzellen; Fyhn, Molden, Witter, Moser & Moser, 2004; Sargolini et al., 2006). Diese Informationsverarbeitungsprozesse werden in Aufgaben, die dynamische Ortsveränderung, d. h. Navigationsfähigkeiten, erfordern, besonders wichtig.

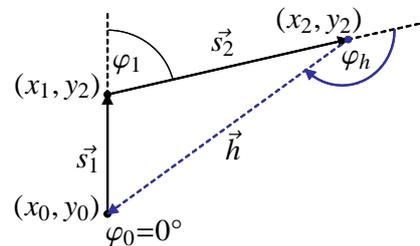
Navigation ist eine räumliche Fähigkeit, sich selbst von einem Punkt im Raum zu einem Zielpunkt zu bewegen. Bei einer Navigationsaufgabe sind in der Hauptsache folgende Teilprobleme zu lösen: (1) Bestimmung der eigenen Position im Raum, (2) Ermitteln des Weges zum Ziel sowie (3) die

Führung zum Ziel auf einem vorgesehenen Weg, d. h. Halten des optimalen Kurses. In allen drei Teilaufgaben sind es Pfadintegrationsprozesse, die zur Orts- und Richtungsbestimmung eingesetzt werden.

Während wir uns bewegen, erhält unser Gehirn ständig sensorische Signale über die Veränderungen von Richtung, Geschwindigkeit und Beschleunigung. Diese Signale können idiothetische (interne) sein, zu denen propriozeptive, vestibuläre und gravitative Informationen, aber auch Efferenzkopien der Bewegung und spezielle visuelle Signale gehören. Letztere sind aus der Bewegung resultierende Veränderungen perspektivischer Ansichten der Umgebung. Ausgehend von einem initialen Punkt im Raum wird durch Pfadintegration eine fortlaufende Aktualisierung der mentalen Repräsentation der eigenen Position durch Akkumulation der Geschwindigkeits- und Beschleunigungsvektoren durchgeführt. Dazu können auch Richtungs- und Distanzschätzungen zu sichtbaren salienten Landmarken, im Sinne allothetischer (externe) Signale, herangezogen werden.

Abbildung 2.2: Pfadintegration als Basis für *Homing* und *Dead Reckoning*

Zur Bestimmung der aktuellen Position durch Dead-Reckoning wird auf die letzte bekannte Position die zurückgelegte Teilstrecke addiert. Die Heimkehrichtung φ_h ergibt sich als Heimkehrvektor \vec{h} aus der Summe aller Teilstrecken auf dem gesamten Weg. Zur Ermittlung der Teilstrecken ist die Integration der Geschwindigkeits- bzw. auch Beschleunigungsvektoren aus Translation und Rotation notwendig. Bei der Pfadintegration werden diese Größen über die Verarbeitung interner Signale wahrgenommen.



$$\text{Heimkehrichtung: } \vec{h} = \sum_{i=1}^n \vec{s}_i$$

$$\text{Dead-Reckoning: } \begin{pmatrix} x_{i+1} \\ y_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \vec{s}_i$$

$$\text{Pfadintegration: } s_i = \int v dt \text{ (Translation)}$$

$$\varphi_i = \int \omega dt \text{ (Rotation)}$$

Pfadintegration (PI) ist also eine Berechnungsmethode, die auch ein technisches Pendant hat, das unter der Bezeichnung Dead-Reckoning besonders in der Nautik bekannt ist (s. Abb. 2.2). Sie ist ein außerordentlich wichtiger Bestandteil erfolgreicher Navigation. Insbesondere das Halten eines optimalen Kurses bedarf einer fortlaufenden Positionsaktualisierung durch PI, um aus dem Abgleich mit der mentalen Repräsentation der Umgebung Abweichungen vom berechneten Kurs zu ermitteln und eventuell Verhaltensänderungen zu generieren. Dabei scheint ein Zusammenhang zwischen PI und Place-Learning zu bestehen.

Im Morris-Water-Maze konnte Wishaw (1998) zeigen, dass hippocampektomierte Ratten in der Lage sind, räumliche Assoziationen zwischen Landmarken und der versteckten Plattform zu lernen und abzurufen. Sie lernten zwar deutlich langsamer als Tier einer Kontrollgruppe, zeigten aber in Testdurchgängen des MWM deutliches, auf Place-Learning schließendes, Verhalten. Jedoch erwies sich dieses Ortswissen als besonders lösungsresistent in Versuchsbedingungen, in denen Landmarken mehrdeutig oder uninformativ waren. Die Ratten mit Hippocampuschädigungen waren nicht in der Lage, die Diskrepanzen zwischen den veränderten Umgebungsreizen und ihrer menta-

len Repräsentation zu erkennen. Kontrolltiere konnten in diesen Situationen auf die ideothetische Signalverarbeitung der PI zurückgreifen und sich schnell umorientieren. Es scheint also, dass PI der Mechanismus ist, der Spatial-Mapping im Hippocampus erst ermöglicht.

Dass Pfadintegrationsprozesse auf hippocampaler Informationsverarbeitung beruhen, vertreten u. a. auch Samsonovich und McNaughton (1997). In einer aktuellen Arbeit erläutern McNaughton, Battaglia, Jensen, Moser und Moser (2006) die Bedeutung hippocampaler Ortszellen und enthorinaler Gitterzellen für die Pfadintegration. Sie sehen in diesen Zellensembeln einen „Pfadintegrator“, der echtes Mapping relativer räumlicher Positionen ohne externe Hinweisreize und nur durch die Integration von Bewegungsinformationen ermöglicht.

Über Pfadintegration gelangen also ganz wesentliche und in ihrer Art implizite Informationen in eine kognitive Karte. Roche, Mangaoang, Commins und O’Mara (2005) fügten deshalb dem Modell der Cognitive-Map-Theorie das Pfadintegrationssystem hinzu. Ähnlich wie Routenwissen basiert PI hier auf egozentrischen Strategien und kann nur durch Bewegung erworben werden. Anders als beim Routenwissen, das episodisch oder semantisch encodiert sein kann, ist die Ausgabe des PI-Systems jedoch eine rein funktionelle Repräsentation des Raumes. Nach Roche et al. (2005) existieren alle Systeme parallel und ihre Ausgaben münden in eine gemeinsame kognitive Karte, die aus einem Netzwerk aus Hippocampusformation, parahippocampalem Cortex und *Gyrus lingualis* besteht.

2.2 Räumliche Präferenzen

Ausgehend vom Strategiemodell stellen auch räumliche Präferenzen eine für die Person typische Art der Informationsverarbeitung dar. Für Putz-Osterloh und Lüer (1979) hängt die Wahl der Bearbeitungsstrategie in einem Raumvorstellungstest von der Art der Aufgabe ab. Köller, Rost und Köller (1994) und Schultz (1991) interpretieren dagegen Unterschiede in der Strategiewahl als Personeneigenschaft und zeigen, dass auch bei gleicher Aufgabenstellung individuelle Bearbeitungsstrategien zum Einsatz kommen. Tatsächlich kann nur mit einer bestimmten Wahrscheinlichkeit unterschieden werden, ob die Entscheidungsprozesse intern oder extern, d. h. von Personen- oder Stimuluseigenschaften, getriggert waren. Zusammen mit den für die Lösung der Aufgabe erforderlichen Kompetenzen und den mentalen Repräsentationen des Testmaterials bestimmen räumliche Präferenzen die Auswahl der Bearbeitungsstrategien.

Die Anwendung von Strategien stellt aus testtheoretischer Sicht eher ein Problem dar. Latente, d. h. nicht direkt beobachtbare, Eigenschaften (Fähigkeiten) einer Person werden i. d. R. durch psychometrische Test operationalisiert. Die Kompetenz bezüglich einer latenten Fähigkeit manifestiert sich dann in der Testleistung. Indem Items existieren, die unterschiedliche Strategien zu lassen, kann nicht mehr von der Messung einer eindimensionalen Fähigkeit ausgegangen werden. Die Konstruktvalidität ist dadurch gefährdet. In einem problemlösungsorientierten Ansatz kann aber gerade die Iteminhomogenität bestimmter Raumvorstellungstests zur Abbildung der unterschiedlichen Strategien genutzt werden. Aus deren Kenntnis kann wiederum auf die verwendeten mentalen Re-

präsentationen und somit auf die typischen Präferenzen im Umgang mit räumlichen Informationen geschlossen werden.

2.2.1 Analytisch vs. holistisch

Die Bedeutung effizienter Bearbeitungsstrategien in Raumvorstellungstests ist bereits seit längerem bekannt. Man unterscheidet dabei zwischen holistischen (ganzheitlichen) oder analytischen (einzelheitlichen) Strategien (Barratt, 1953; Cooper, 1976; Hosenfeld et al., 1997). Versuchspersonen, die holistische Strategien einsetzen, sind in der Lage, sich Figuren aus einem Test als ganze dreidimensionale Objekte vorzustellen. Versuchspersonen, die analytische Strategien einsetzen, repräsentieren und analysieren jeweils nur Teile der Figuren.

In Raumvorstellungstests, wie den Untertests *Würfelaufgaben* und *Figurenauswahl* aus dem Intelligenz-Struktur-Test (I-S-T; Amthauer, Brocke, Liepmann & Beauducel, 1999) oder dem *Mental Rotations Test* (MRT; Peters et al., 1995), sind zur Lösung der Aufgaben in unterschiedlichem Maße räumlich-perzeptive und räumlich-kognitive Fähigkeiten, z. B. in der Erkennung räumlicher Relationen und der mentalen Rotation, erforderlich. Aufgrund struktureller Besonderheiten einiger Testitems können diese aber auch ohne mentale Rotation gelöst werden. Im Test *Würfelaufgaben* betrifft das die unterschiedlichen Aufgabentypen „Raumwürfel“ und „Flächenwürfel“ (Putz-Osterloh, 1977). Im MRT sind es Itemtypen, die durch die Art der Distraktoren in „Spiegel-“ oder „Strukturitems“ unterteilt werden (Geiser, 2004; Voyer & Saunders, 2004). Nur „Raumwürfel“ und „Spiegelitems“ erfordern mentales Rotieren, also die Anwendung holistischer Strategien. Alle anderen Aufgaben können auch durch analytische Detailvergleiche gelöst werden. Auch im Test „Figurenauswahl“ gibt es Items, die nur durch Zusammenschieben der Figurenteile und ohne jegliche mentale Rotation gelöst werden können (Bühner, Ziegler, Krumm & Schmidt-Atzert, 2006).

Mit Klassenanalysen der Item-Response-Theorie (Rost, 1996) gelingt es Personengruppen zu identifizieren, die in Raumvorstellungstests Präferenzen für holistische bzw. analytische Strategien zeigen. Dabei werden anhand probabilistischer Mehr-Klassen-Modelle (Mixed-Rash-Modell oder Latent-Class-Modell) aus den beobachteten Patternhäufigkeiten Subgruppen abgeleitet, die sich durch unterschiedliche Antwortprofile auszeichnen. Es lassen sich dann, unter Ausnutzung der strukturellen Besonderheiten des Raumvorstellungstests und der Klassenprofile, Personen identifizieren, die die Aufgaben wahrscheinlich mit Hilfe holistischer, also ganzheitlicher, Strategien lösen (Köller et al., 1994; Hosenfeld et al., 1997; Glück, Machat, Jirasko & Rollett, 2001; Geiser et al., 2006). Manche Klassen zeichnen sich dadurch aus, dass ihre Versuchspersonen offensichtlich holistische Strategien anwenden, dies aber auf langsamere, ineffizientere Weise, so dass sie durch das Zeitlimit im Test weniger Items bearbeiten können (Geiser et al., 2006). Versuchspersonen, die gezielt Items, wie die „Strukturitems“ im MRT oder die „Flächenwürfel“ der *Würfelaufgaben*, herausuchen und bearbeiten, wenden höchstwahrscheinlich keine holistischen, sondern analytische Strategien an.

Die holistische vs. analytische Präferenz spiegelt unmittelbar unterschiedliche Organisationsprinzipien des Arbeitsgedächtnisses wieder. Die Anwendung analytischer Strategien kann als sukzessive

Abspeicherung aller, für relevant befundenen, räumlichen Details angesehen werden. Zusätzlich können darunter Transformationen, z. B. in verbale Repräsentationen, verstanden werden. Bei Anwendung entsprechender Strategien auf sehr detailreiche räumliche Materialien können deshalb Arbeitsgedächtniskapazitäten schnell erschöpft sein. Daraus können wiederum in der Testperformanz hohe Verarbeitungszeiten und Fehlerraten resultieren. Dagegen funktionieren holistische Repräsentationen wie Chunks als Zusammenfassung vieler räumlicher Details in einem effizienter zu verarbeitendem Item.

2.2.2 Egozentrisch vs. allozentrisch

Noch enger an räumlich-mnestische Informationsverarbeitung geknüpft ist der Begriffe der egozentrischen und allozentrischen Strategien. Unter egozentrisch vs. allozentrisch kann eine Präferenz des zur räumlichen Orientierung bevorzugt eingesetzten Referenzsystems verstanden werden. Allozentrische Referenzsysteme verwalten von der eigenen Position unabhängige Repräsentationen der Umgebung als Überblickswissen, wie die durch das Place-Learning-Mapping-System aufgebauten kognitiven Karten. Egozentrische Bezugssysteme sind alle Informationen im Route-Learning-System, d. h. das relativ zur eigenen räumlichen Position repräsentierte Routenwissen.

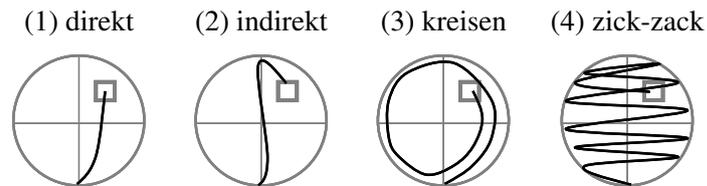
Ging man seit O’Keefe und Nadel (1978) allgemein davon aus, dass die Akquisition räumlicher Informationen letztendlich immer in kognitiven Karten mündet, argumentierten Wang und Spelke (2002), dass eher die dynamisch egozentrischen Repräsentationen eingesetzt werden. Auf welches System zurückgegriffen wird, hängt aber vielmehr von der Vertrautheit mit der Umgebung und der Zuverlässigkeit der jeweiligen Repräsentationen ab. Die Zuverlässigkeit egozentrischer Systeme kann durch Desorientierung, z. B. große Drehungen um die Körperlängsachse, empfindlich gestört werden (Wang & Spelke, 2002). In bekannten Umgebungen, in denen offensichtlich Überblickswissen zur Reorientierung genutzt werden kann, zeigt sich diese Desorientierung nicht. Sie zeigt sich ebenfalls nicht nach Rotationen von weniger als 90° (Waller & Hodgson, 2006). Egozentrische und allozentrische Repräsentationen können offensichtlich nebeneinander existieren und spontane Wechsel im Abruf zwischen den Systemen sind möglich.

Beim Abruf allozentrischer Informationen sind dann ausgedehnte Aktivierungen im rechten Hippocampus beobachtet worden, was dessen Bedeutung für diese Art der räumlichen Repräsentation unterstreicht (Jordan, Schadow, Wuestenberg, Heinze & Jäncke, 2003; Iaria, Petrides, Dagher, Pike & Bohbot, 2003). Bei Iaria et al. (2003) berichteten Probanden aber auch vermehrt einen bewussten Wechseln von der allozentrischen Strategie zu einer egozentrischen. Letztere ging mit einer erhöhten Aktivierung im *Nucleus Caudatus* einher. Jordan et al. (2003) berichteten für egozentrische Strategien eine stärkere Aktivierung im linken Thalamus. Diskrepanzen in den neuronalen Aktivierungsmustern könnten in der größeren Bandbreite egozentrischer Strategien begründet sein.

Geht man auf die Ebene der Verhaltensbeobachtung wird deutlich, dass insbesondere egozentrische Orientierungsstrategien in eine ganze Reihe unterscheidbarer Bewegungsmuster auffächern. In Navigationsexperimenten, wie der Morris-Water-Maze-Aufgabe, können Strategien auch durch

Klassifizierung der Trajektorien unterschieden werden. Durch eine manuelle Analyse der Hauptkomponenten in den Spurdaten erkannten Astur, Tropp, Sava, Constable und Markus (2004) vier typische Verhaltensmuster: (1) direktes Anschwimmen der Plattform; (2) Anschwimmen einer Landmarke bis zum Beckenrand und von dort Umorientierung zur Plattform; (3) Schwimmen entlang des Beckenrandes auf einer kreisförmigen Bahn bis zur Plattform; (4) Durchsuchen des gesamten Beckens auf einem Zick-Zack-Kurs (s. Abb. 2.3).

Abbildung 2.3: Bewegungsmuster der vier von Astur et al. (2004) in einer Water-Maze-Aufgabe identifizierten Suchstrategien.



Nur Probanden, die die direkte Strategie (1) anwenden, verfügen offensichtlich über stabile allozentrische mentale Repräsentationen der Umgebung und wenden deshalb per Definition als einzige eine allozentrische Strategie an. Die Probanden mit der indirekten Strategie (2) steuern zunächst auf eine mit dem Ziel assoziierte Landmarke zu, um sich dann von dort aus lokal zum Ziel zu orientieren. Diese Strategie beruht zwar auf relationalem Wissen über den Raum, ist aber eine zusammengesetzte Strategie aus Reiz-Reaktions-Verhalten und der Durchführung von einstudierten Bewegungssequenzen. Dieses Verhalten kann ausschließlich über egozentrische Referenzsysteme generiert werden.

Probanden mit Strategie drei oder vier zeigen egozentrische Verhaltensmuster, die auf der wiederholten Anwendung von Bewegungssequenzen beruhen. Typisch und häufig beobachtet ist die Klassifizierung in kreisende oder geradlinige Vor-und-Zurück-Bewegungen (Kallaia et al., 2005; Brody & Holtzman, 2006). Beide Strategien werden zum systematischen Durchsuchen der Umgebung eingesetzt und deshalb auch von den beiden ersten Strategien durch die Bezeichnung als nicht-räumliche Strategien abgrenzt. Eine besondere Form einer kreisenden Strategie im MWM ist Thigmotaxis¹, bei der in unmittelbarer Nähe der äußeren Begrenzung des Pools eine bestimmte Strecke zurückgelegt wird, die in die unmittelbare Nähe des Ziels führt (Kallaia et al., 2005).

In Astur et al. (2004) konnten von 51 Probanden 23 (45 %) mit einer Präferenz für die direkte, d. h. allozentrische Strategie, identifiziert werden. Nur 27 % wendeten egozentrische Strategien an. Es wurde aber auch klar, dass mit der manuellen Methode mehr als ein Viertel der Versuchspersonen (28 %) nicht genauer klassifiziert werden konnten, da sie Misch-Strategien anwendeten oder undifferenzierbare Bewegungsmuster zeigten.

¹Thigmotaxis: durch Berührungsreiz ausgelöste Orientierungsbewegung von Tieren

2.2.3 Unimodal vs. bimodal

Wechselnde Strategien können aber auch entdeckt werden, da sie typische Verhaltensmuster in den Daten erzeugen. Lernkurven aus dem MWM weisen zu Beginn oft eine alternierende Performanz in aufeinander folgenden Durchgängen auf. Dieses Auf und Ab in den Leistungsmaßen wird als bimodales Verhalten bezeichnet und lässt auf instabile mentale Raumrepräsentationen schließen (Burke & Barnes, 2006). Nähern sich die Leistungsmaße einem stabilen Wert an, spricht man von unimodalem Verhalten. Typische Lernkurven zeigen nun einen charakteristischen Wechsel zwischen bimodalem und unimodalem Verhalten.

Begründet ist dieses Verhalten vermutlich in einer eigentlich antagonistischen Eigenschaft kognitiver Karten. Sie müssen einerseits extrem stabil und andererseits äußerst flexibel sein. Ortszellen zweier verschiedener Regionen im Hippocampus (CA1 und CA3) sind offensichtlich in der Lage, diese funktionelle Dissoziation zu verwirklichen (Barnes, Suster, Shen & McNaughton, 1997). Tierexperimente im Morris-Water-Maze zeigten, dass die kognitiven Karten in der CA1-Region einiger Ratten über lange Zeit stabil bleiben. Kehren diese Tiere in eine bekannte Umgebung zurück wird die gleiche Karte abgerufen, die zuvor mit dieser Umgebung encodiert wurde. Bei neuen Ratten feuern in einer solchen Situation plötzlich andere unabhängige Neuronen der CA1-Region, d. h. die Karten waren lediglich innerhalb der Trainingsepisode stabil. Stabil sind dagegen bei diesen Tieren kognitive Karten der CA3-Region, die fatalerweise unverändert auch in neuen Umgebungen feuern. Bei den flexibleren Tieren existierten dagegen während der Exploration unabhängige Repräsentationen der neuen und der alten Umgebung in den CA1- bzw. CA3-Regionen.

Diese Einschränkung mnestischer Systeme erzwingt also unmittelbar die rigide Anwendung bestimmter räumlicher Strategien. Im Verhalten sind diese als nicht vorhandene allozentrische Repräsentationen zu beobachten, obwohl die Fähigkeit zum Erwerben dieser durchaus bestehen kann. Aus dem Verhalten kann also nicht eindeutig geschlossen werden, ob die entsprechende Repräsentation nicht existiert oder einfach nicht angewandt wurde. Es kann aber als Präferenz für eine rigidere gegenüber einer flexibleren Verhaltensgenerierung interpretiert werden.

3 Alters- und Geschlechtsdifferenzen in der Raumkognition

3.1 Altersdifferenzen

Zweifelloos steht biologisches Altern mit Veränderungen auch kognitiver Fähigkeit in Zusammenhang. Quantitative Unterschiede zwischen Altersgruppen zeigen sich in fast allen Teilbereichen. Die Frage, ob diese Unterschiede als Alters- oder Entwicklungseffekte zu interpretieren sind, ist nicht trivial und wird hier beantwortet, indem neuropsychologische und entwicklungspsychologische Aspekte zusammenhängend betrachtet werden. In der hier zugrunde liegenden Ansicht wird Altern als Menge aller auf- und abbauenden Entwicklungsprozesse von neuroanatomischen Strukturen und kognitiven Funktionen über die gesamte Lebensspanne aufgefasst.

In geriatrischen Entwicklungsabschnitten zeigen sich Abnahmen der Verarbeitungsgeschwindigkeit, des Gedächtnisses und der kognitiven Flexibilität. Diese werden eher als Defizite angesehen, die die allgemeine Leistungsfähigkeit älterer Probanden herabsetzen. Jedoch beobachtet man in Aufgaben mit raumkognitiven Anforderungen besondere Auffälligkeiten, die mit dieser allgemeinen Leistungsabnahme allein nicht zu erklären sind (Newman & Kasznaik, 2000; Wilkniss, Jones, Korol, Gold & Manning, 1997)

Ältere Versuchspersonen um das sechzigste Lebensjahr benötigen in einer virtuellen Version des Morris-Water-Maze deutlich mehr Lerndurchgänge, um zielgerichtet die versteckte Plattform zu erreichen als jüngere Versuchspersonen (Driscoll et al., 2003; Driscoll, Hamilton, Yeo, Brooks & Sutherland, 2005; Laurence et al., 2002; Moffat & Resnick, 2002). Obwohl ältere Erwachsene weniger Praxis mit virtuellen Umgebungen haben (näheres dazu im Kap. 4), können sie diese explorieren und in Form einer kognitiven Karte encodieren. Positionen von distalen Landmarken und deren Relationen untereinander werden von jüngeren und älteren Versuchspersonen gleich gut erinnert (Laurence et al., 2002).

Allerdings haben ältere Erwachsene keine präzise Vorstellung, wo sich die versteckte Plattform befindet. In Versuchsbedingungen, in denen die Plattform entfernt wird, halten sich jüngere Versuchspersonen deutlich länger im ursprünglichen Zielquadranten auf als ältere. Ist die Plattform sichtbar gibt es keine Leistungsunterschiede (Driscoll et al., 2005), d. h. die beobachteten Unterschiede sind sehr wahrscheinlich auf eine altersbedingte Abnahme der räumlichen Fähigkeit Place-Learning zurückzuführen. Allerdings nimmt mit dem Alter auch die Leistungsvarianz sehr stark zu (Moffat & Resnick, 2002).

In Fähigkeiten, die mit der Bearbeitung der Water-Maze-Aufgabe zusammenhängen, wie die Positionsbestimmung durch Pfadintegration, zeigen sich meist keine Alterseffekte (Lemay, Bertram & Stelmach, 2004). Lediglich hinsichtlich der Reaktionszeiten gibt es Unterschiede zu Gunsten jüngerer Probanden. Auch Facetten der räumlichen Vorstellung sind zirka ab dem 65. Lebensjahr geringer ausgeprägt (Driscoll et al., 2005; Geiser, 2004). Bei mentaler Rotation z. B. bestehen aber schwächere Zusammenhänge zwischen biologischem Alter und Testleistung als in der Water-Maze-Aufgabe (Geiser, Lehmann & Eid, in Press).

Vielmehr finden sich in einigen Raumvorstellungstests starke Unterschiede innerhalb früherer Entwicklungsphasen, nicht zuletzt deshalb, da diese im Zusammenhang mit Geschlechtsdifferenzen eingehend untersucht wurden (näheres in dazu in Abs. 3.2). Während der Adoleszenz scheinen maßgebliche Veränderungen in den Facetten mentale Rotation, räumliche Wahrnehmung und Visualisierung stattzufinden (Voyer, Voyer & Bryden, 1995; Linn & Petersen, 1985).

Kinder im Alter von fünf bis zwölf Jahren unterscheiden sich dagegen sehr deutlich in ihrer Fähigkeit zur Pfadintegration von jungen Erwachsenen (Jansen-Osmann & Fuchs, 2006). In einer Aufgabe zur Schätzung der Heimkehrrichtung schätzten junge Erwachsene im Alter von 25 Jahren deutlich genauer als Kinder der Altersklasse fünf bis sieben Jahre. Sehr deutlich sind allerdings auch die Differenzen zwischen jüngeren (sieben Jahre im Mittel) und älteren Kindern (elf Jahre im Mittel). Jüngere Kinder scheinen noch nicht in der Lage zu sein, korrekte Orientierungsreaktionen in einer komplexen Umgebung zu zeigen.

Diese Unterschiede fallen sehr genau mit Entwicklungsabschnitten zusammen, die sich auf die Verfügbarkeit mentaler Repräsentationen des Raumes beziehen. So sind Kinder vor dem zehnten Lebensjahr meist nicht in der Lage auf allozentrische Raumrepräsentation zurückzugreifen. Sie zeigen in der Mehrzahl egozentrische Orientierungsreaktionen, die bei Störungen der Umwelt (z. B. Löschen lokaler Hinweisreize) versagen. Ältere Kinder (zehn bis zwölf Jahre) dagegen sind in ihrem räumlichen Verhalten bereits unabhängig von diesen lokalen Hinweisen und dadurch in der Lage erfolgreichere Orientierungsstrategien einzusetzen (Lehning, Haaland, Pohl & Leplow, 2001; Leplow, Höll, Zeng & Mehdorn, 2000).

Der Aufbau allozentrischen Raumwissens scheint dagegen im höheren Erwachsenenalter prinzipiell möglich zu sein, jedoch wird die Leistungsfähigkeit in einer Water-Maze-Aufgabe durch unflexiblen Umgang mit diesem Wissen negativ beeinflusst. Versuchspersonen im jungen Erwachsenenalter zeigen in den Lernkurven den typischen Wechsel zwischen bimodalem und unimodalem Verhalten. Einige haben bereits nach einem einzigen Durchgang eine stabile kognitive Karte der Umgebung erworben (One-Trial-Learning). Ältere Erwachsene behalten dagegen bis zum Schluss ein bimodales Muster in ihrem Lernverhalten bei (Driscoll et al., 2003).

Neuropsychologische Erklärungsmodelle führen die mit Altern und Entwicklung verbundenen komplexen Verhaltensveränderungen auf die Dissoziation zweier neuronaler Informationssysteme zurück. In einem Mehr-Faktoren-Modell unterscheidet Buckner (2004) zwischen einem fronto-striatalen System und der medial-temporalen Schleife. Das fronto-striatale System steht in Zusammenhang mit Aufmerksamkeit und Exekutivfunktionen. Es reagiert sensitiv auf Verände-

rungen der grauen und weißen Substanz und einiger Neurotransmitter. Das Volumen der frontalen grauen Substanz ist bereits mit dem elften Lebensjahr auf dem Level eines Erwachsenen. Im Laufe weiterer sechs Jahre reift nur noch weiße Substanz heran (Caviness, Kennedy, Richelme, Rademacher & Filipek, 1996). Im Frontalkortex ist auch die stärkste altersbedingte Atrophie grauer Substanz zu beobachten. In lateral-frontalen Arealen geht diese mit einer direkten Abnahme exekutiver Funktionen einher (Zimmerman et al., 2006). Diese können durch einen Rückgang strategischer Fähigkeiten zu einer leichten Beeinträchtigung mnestischer Leistungen führen. Veränderungen im fronto-striatalen System werden von Buckner (2004) als die mit normalem Altern in Verbindung stehenden Abbauprozesse angesehen.

Mnestische Abbauprozesse in der medial-temporalen Schleife sind dagegen pathologische Erscheinungen. Im Zentrum des medial-temporalen Speichersystems steht der Hippocampus, dessen Volumen über die Lebensspanne erstaunlich konstant bleibt. Im zweiten bis dritten Lebensjahr erreicht der Hippocampus bereits seine maximale Größe, die bis wenigstens zum 18. Lebensjahr stabil bleibt. Altersbedingte Volumenabnahmen dieser Hirnstruktur liegen zwar vor, konnten jedoch in einer Metaanalyse durch Van Petten (2004) nicht in einen direkten Zusammenhang zu altersbedingten Gedächtnisdefiziten gebracht werden. Aufgrund solcher Ergebnisse geht man mittlerweile davon aus, dass eine generelle Abnahme der Neuronendichte nicht mit dem normalen Alterungsprozess, sondern mit pathologischen Veränderungen einhergeht.

Vielmehr scheinen anatomische und biochemische Veränderungen einzelner innerer Strukturen im Hippocampus mnestische und insbesondere räumliche Fähigkeiten über die gesamte Lebensspanne zu beeinflussen (Driscoll et al., 2003). Diese sind mit Volumetrie nicht nachweisbar. Als Ursache reduzierter Place-Learning-Fähigkeiten vermuten Burke und Barnes (2006) eine veränderte Interaktion zwischen Nervenzellen. So ist eine Abnahme der neuronalen Plastizität im medialen Temporallappen vermutlich mit verantwortlich für altersbedingte Defizite im Mechanismus der Langzeitpotenzierung.

3.2 Geschlechtsdifferenzen

Leistungsunterschiede in raumkognitiven Fähigkeiten werden häufig in Zusammenhang mit Geschlechtsdifferenzen berichtet (Cahill, 2006; Kimura, 1999). Obwohl es Hinweise auf Geschlechtsunterschiede in allen Facetten der Raumvorstellung gibt, zeigen sich diese bezüglich der mentalen Rotation am deutlichsten. Metaanalysen bestätigten Mittelwertsunterschiede zwischen männlichen und weiblichen Probanden als reliable und große Effekte (Linn & Petersen, 1985; Masters & Sanders, 1993; Voyer et al., 1995). Mit d^1 als Effektstärkemaß berichtete Geschlechtsdifferenzen in der mentalen Rotation bewegen sich im Bereich zwischen $d=0.56$ und $d=0.86$ (Linn & Petersen, 1985; Voyer et al., 1995; Collins & Kimura, 1997). Besonders in der Operationalisierung durch den *Mental Rotations Test* (MRT: Vandenberg & Kuse, 1978; Peters et al., 1995) ergeben sich teilweise

¹Cohens d : Differenz von zwei Gruppenmittelwerten normiert an der Standardabweichung der Gesamtstichprobe (Cohen, 1977)

noch größere Unterschiede. Im MRT werden oft Differenzen bis zu einer Standardabweichung und mehr berichtet (Rönicke, 2003; Lehmann, 2000).

Dieser Unterschied ist über die gesamte Lebensspanne nachweisbar (Linn & Petersen, 1985) und ist am ausgeprägtesten während der Adoleszenz. Voyer et al. (1995) berichten von einer kontinuierlichen Zunahme der Effektstärken von $d=0.33$ für Kinder unter 13 Jahren bis zu $d=0.66$ für junge Erwachsene ab 18 Jahren. Im Altersbereich zwischen 13 und 18 Jahren ergaben sich die größten Differenzen in der Testleistung zwischen Mädchen und Jungen für die Altersgruppe der 14-jährigen (Geiser, 2004).

Geschlechtsunterschiede im Place-Learning wurden mit Water-Maze-Aufgaben mehrfach untersucht (Driscoll et al., 2005; Astur et al., 2004, 1998; Sandstrom, Kaufman & Huettel, 1998). Die Ergebnisse lassen sich zu folgenden Erkenntnissen zusammenfassen: Männer finden schneller und präziser sowie auf kürzeren Wegen die versteckte Plattform in den Lern- und Trainingsdurchgängen. In Testdurchgängen, in denen die Plattform entfernt oder verschoben wurde, suchen sie intensiver an der ursprünglichen, erlernten Position der versteckten Plattform. Dieser Vorteil bleibt über die gesamte Lebensspanne stabil (Driscoll et al., 2005, 2003; Moffat & Resnick, 2002). Die Geschlechtseffekte sind dabei je nach verwendetem Leistungsmaß zwischen $d=0.65$ und $d=1.04$ angesiedelt (Astur et al., 1998). Diese Leistungsmaße korrelieren mit Testleistungen zur mentalen Rotation (Summenscore des MRT) signifikant um Werte von $r=-0.34$ bis $r=-0.48$ (Driscoll et al., 2005; Astur et al., 2004).

Auch in Heimkehr- und Zeigeaufgaben als Operationalisierung von Pfadintegrationsfähigkeiten bestehen Geschlechtsunterschiede, die jedoch nicht immer konsistent nachgewiesen werden konnten. Lawton und Morrin (1999) fanden, dass Männer im Mittel, unabhängig von der Länge der gelaufenen Wege, um 20° genauere Schätzungen der Heimkehrrichtung in einem virtuellen Labyrinth abgaben. Keinen Geschlechtsunterschied in der Genauigkeit der Richtungsschätzungen fanden dagegen Golledge, Ruggles, Pellegrino und Gale (1993), Montello und Pick (1993) in einer realen Umgebung. Silverman et al. (2000) fanden wiederum bei einer Heimkehraufgabe durch eine waldrreiche reale Gegend eine leichte Überlegenheit männlicher Versuchspersonen. Von den raumkognitiven Maßen, die Silverman et al. (2000) erhoben hatten, korrelierte nur der MRT signifikant mit dem Leistungsmaß aus der Heimkehraufgabe zu $r=-0.29^2$.

Korrelative Zusammenhänge zwischen Place-Learning, Pfadintegration und mentaler Rotation sowie die gleichermaßen vorhandenen Geschlechtsdifferenzen bei diesen räumlichen Fähigkeiten lassen hier ähnliche Wirkmechanismen vermuten. Zur Erklärung der Geschlechtsunterschiede haben wahrscheinlich mehrere Faktoren interagierende oder vermittelnde Eigenschaften. Insbesondere Geschlechtshormone wurden hier auf Grund ihrer organisierenden und aktivierenden Wirkung auf Hirnstrukturen und -funktionen als Einflussfaktor vermutet (Burkitt, Widman & Saucier, 2007; Driscoll et al., 2005; Miller, Grieco & Klibanski, 2005; Bell & Saucier, 2004; Hausmann, Slabbekoorn, Van Goozen, Cohen-Kettenis & Güntürkün, 2000; Phillips & Silverman, 1997). Neben den endokrinen Einflüssen werden aber auch evolutionsbiologische (Silverman & Eals, 1992), neu-

²Partial Korrelationskoeffizient mit Geschlecht als Kontrollvariable

ropsychologische (Kimura, 1999), sozialisations- und kulturbedingte Ansätze diskutiert. Im Zuge dieser Diskussion taucht immer wieder das Problem der Anwendung unterschiedlicher Lösungs- und Bearbeitungsstrategien auf.

Bei Raumvorstellungstests spricht vieles dafür, dass Frauen in der Mehrzahl die Aufgaben anders bearbeiten als Männer. Eine Einteilung der Versuchspersonen in fünf latente Klassen ergab eine überproportional häufigere Besetzung von Klassen, die höchstwahrscheinlich holistische Strategien einsetzten, mit männlichen Versuchspersonen (Geiser et al., 2006). Mehr als 50 % der männlichen Probanden gehörten diesen Klassen an, dagegen nur 25 % der weiblichen Versuchspersonen. Einer Klasse, die besonders wenige Items löste, gehörten mehr als 35 % der weiblichen und nur ca. 15 % der männlichen Personen an. Das Profil dieser Klasse ließ vermuten, dass die MRT-Aufgaben für diese Probanden generell zu schwierig waren. Eine eindeutige Aussage, ob dies auf die Anwendung ineffizienter analytischer Strategien zurückzuführen sei, konnte nicht getroffen werden. Stattdessen konnte durch die Klassenanalyse eine Subgruppe ermittelt werden, deren Personen gezielt Items herausuchten, die analytische Strategien bevorzugen. Eine Differenzierung nach Geschlechtszugehörigkeit ergab eine, wenn auch gering, größere Häufigkeit weiblicher Personen in dieser Klasse (20 % zu 15 %).

Im MRT haben, neben den in Geiser et al. (2006) betrachteten, weitere Eigenschaften der Items Einfluss auf die Geschlechtsdifferenzen. In Voyer und Hou (2006) ist Okklusion, d. h. die perspektivische Darstellung verdeckter Teile einer Figur, ausschlaggebend für die Gesamtleistung und somit verantwortlich für die Geschlechtsdifferenzen in diesem Test.

Neurowissenschaftliche Befunde untermauern die Annahme, dass weibliche Versuchspersonen andere Verarbeitungsmechanismen und Strategien in Aufgaben zur mentalen Rotation einsetzen. Weiblichen Probanden aktivieren während der Verarbeitung von Rotationsstimuli zusätzlich okzipitale Areale und größere Areale im intra-parietalen Sulcus (IPS) als männliche Versuchspersonen. Die mehr ventralen Aktivierungen in okzipitalen Arealen könnten für eine in diesem Verarbeitungspfad stattfindende Analyse von Objektdetails, die ausgedehntere dorsale Aktivierung im IPS für eine größere Speicher- und Verarbeitungslast räumlicher Informationen stehen.

Eine andere Art der Verarbeitung räumlicher Informationen spielt auch bei den Geschlechtsdifferenzen in Place-Learning- und Pfadintegrationsaufgaben eine Rolle. Prinzipiell gibt es keine Anhaltspunkte dafür, dass die mentale Repräsentation des Raumes in Form von Vektoren, kognitiven Karten oder Bewegungssequenzen geschlechtsspezifischen Einschränkungen unterworfen sein sollte. Es scheint aber für Frauen und Männer unterschiedliche Ausprägungen in den Präferenzen für Umgebungsmerkmale zugeben. So bevorzugen weibliche Versuchspersonen eher saliente Landmarken zur räumlichen Orientierung. Männliche Versuchspersonen verwenden dagegen eher Informationen über die Geometrie der Umgebung, Distanzen und Peilungen zu Fixpunkten (Sandstrom et al., 1998).

Diese Präferenzen könnten zu einer frühen Filterung in der Wahrnehmung räumlicher Informationen beitragen, die den Aufbau allozentrischer oder egozentrischer Raumrepräsentationen begünstigt bzw. verhindert. Beobachtbar wird dies in den Mustern der Suchstrategien, z. B. in den Water-Maze-

Aufgaben. In Astur et al. (2004) konnten mehr als 65 % der männlichen Versuchspersonen mit direkten Suchstrategien identifiziert werden. Die Mehrzahl der weiblichen Versuchspersonen (60 %) verwendeten tatsächlich indirekte oder unbekannte Misch-Strategien.

Geschlechtsspezifische Verarbeitung räumlicher Informationen zeigt sich auch in der Aktivierung unterschiedlicher neuronaler Netzwerke. Weibliche Versuchspersonen aktivieren während einer Navigationsaufgabe stärker als Männer rechtsseitige Areale des räumlichen Arbeitsgedächtnisses: inferior-parietale und dorsolaterale Kortexanteile. Während der gleichen Aufgabe aktivieren Männer stärker Anteile des medial-temporalen Systems: posteriorer Hippocampus linksseitig (Grön, Wunderlich, Spitze, Tomczak & Riepe, 2000).

4 Raumkognition in virtuellen Umgebungen

4.1 Räumliche Vorstellung virtueller Umgebungen

Virtuelle Umgebungen (VUs) eröffnen in der Raumkognitionsforschung eine elegante Möglichkeit zur Durchführung von Verhaltensexperimenten ohne tatsächliche Bewegungen der Versuchsperson. Dadurch gelingt auch eine exakte zeitliche Kopplung in der Erhebung physiologischer und verhaltensbezogener Daten (Iaria et al., 2003; Jordan et al., 2003; Grön et al., 2000). Generell bietet die Methode gegenüber einem realen Setup eine besseren Kontrollierbarkeit und gegenüber einem „klassischen“ Raumvorstellungstest eine höhere ökologische Validität. Aus diesen Gründen ist sie in der Raumkognition als vielseitige Forschungsmethode etabliert.

Man geht davon aus, dass in virtuellen Umgebungen gewonnene räumliche Informationen in den gleichen mentalen Repräsentation gespeichert werden, wie diese aus der realen Umwelt. So belegen Arbeiten von Waller (2000) und Höll, Lepow, Schönfeld und Mehdorn (2002), dass zuvor in einer VU trainierte Personen in realen Testsituationen besser abschneiden als Untrainierte. Diese Transfereffekte legen nahe, dass in der VU entsprechende Wissenstrukturen aufgebaut und in der realen Umgebungen auch abgerufen werden können.

Implementierungen eines virtuellen Morris-Water-Maze demonstrieren zudem die Möglichkeit des Place-Learnings in einer virtuellen Umgebung (Sandstrom et al., 1998; Astur et al., 1998; Moffat & Resnick, 2002; Laurance, Learmonth, Nadel & Jacobs, 2003; Skelton, Ross, Nerad & Livingstone, 2006). Obwohl es sich nur um Desktop-Implementierungen der von Morris (1984) entworfenen Aufgabe handelte, sind Probanden in der Lage, den Ort der versteckten Plattform zielgenau anzu-steuern. Zudem sind, konsistent mit den vorhergesagten Defizit aus dem Tiermodell, im virtuellen Water-Maze Patienten mit Schädigungen des Hippocampus und angrenzender Bereiche beeinträchtigt in ihrer Fähigkeit zum Place-Learning (Brandt et al., 2005; Astur, Taylor, Mamelak, Philpott & Sutherland, 2002; Skelton, Bukach, Laurance, Thomas & Jacobs, 2000).

Es gibt aber auch Kritiken an der Methode. Ein Hauptargument sind die fehlenden tiefensensiblen Signale, welche bei der Erfahrung des extrapersonellen Raumes eine wichtige Rolle spielen (Wang & Spelke, 2002; Klatzky et al., 1990). Außerdem muss man sich im Klaren sein, dass die virtuelle Umgebung in ihrer räumlichen Ausdehnung nur in der Vorstellung der Versuchsperson existiert. In diesem Zusammenhang zeigte Waller (2000), dass die Performanz in einer virtuellen Navigationsaufgabe stark mit Fähigkeiten der Raumvorstellung korreliert. Zum genaueren Verständnis dieser Zusammenhänge wird im folgenden Abschnitt das allgemeine Funktionsprinzip einer VU kurz erläutert.

4.2 Virtuelle Umgebungen

Mit dem Begriff virtuelle Umgebungen ist im Allgemeinen jede künstliche computergenerierte Umwelt gemeint, die dem Nutzer Interaktionen in Echtzeit erlaubt. Echtzeit bedeutet, dass die Initiierung von Abläufen, die Durchführung von Aktionen sowie die Veränderung von Objekten in der VU realen zeitlichen Mustern folgen. Die technischen Voraussetzungen für diese Interaktionen werden durch Computersystem der virtuellen Realität (VR) geschaffen. Die Eigenschaften eines VR-Systems werden im Wesentlichen durch den Grad der Immersion und die Präsenz des Nutzers innerhalb der VU beschrieben (Held & Durlach, 1992).

Der Begriff der Immersion dient der Klassifizierung von VR-Systemen anhand der eingesetzten Technologien und bezeichnet das „Eintauchen“ in die virtuelle Umgebung, so dass diese als real wahrgenommen wird. Erreicht wird dies durch spezielle Ein- und Ausgabegeräte, die Informationen über die reale Welt ausblenden und künstliche Sinneseindrücke über die virtuelle Umwelt vermitteln (Bülthoff & van Veen, 1999; Loomis, Klatzky, Golledge & Philbeck, 1999). Der Grad der Immersion hängt von der Anzahl der angesprochenen Sinnesmodalitäten und der Bandbreite der generierten Signale ab.

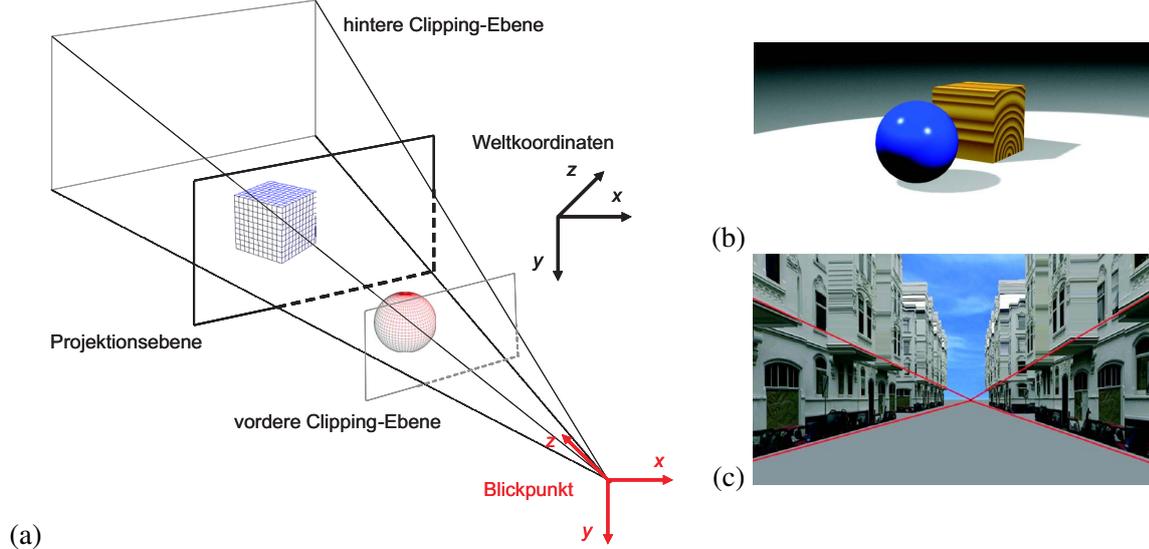
Ziel dieser immersiven Implementierung ist die Erzeugung einer Präsenz beim Nutzer. Als Präsenz wird die Wahrnehmung einer Person bezeichnet, sich an einem Ort zu befinden, obwohl sie sich physisch an einem anderen Ort aufhält. Genauer bezeichnet es, dass eine Person, die durch die Technik vermittelte virtuelle Umgebung nicht als Illusion wahrnimmt, sondern als realen Raum und in diesem authentisch agiert (Herbelin, Vexo & Thalmann, 2002).

Obwohl die Präsenz vom Grad der Immersion abhängt, gibt es Anwendungen die ohne immersive VR-Techniken beim Nutzer den Eindruck einer Präsenz erzeugen können. Diese als Desktop-VRs bezeichneten Applikationen vermitteln ausschließlich visuelle Sinneseindrücke über Computerbildschirme und verwenden Standardeingabegeräte. Eine realistische Visualisierung der virtuellen Umwelt durch Computergrafik ist dadurch um so wichtiger. 3D-Grafiksoftware setzt dazu auch auf wahrnehmungspsychologische Effekte, um wirkungsvolle Abbildungen einer real nicht existenten Umgebung und darin enthaltener Objekte zu erzeugen.

Grafische Komponenten einer virtuellen Umgebung sind 3D-Szenen (s. Abb. 4.1 a). Unter Angabe von Lage und Position (Weltkoordinaten) werden dreidimensionale Modelle von Objekte darin platziert. Durch die Wahl von Blickpunkt und Clipping-Ebenen wird dem Betrachter ein Ausschnitt dieser virtuellen Umwelt gezeigt. Verändert sich der Blickpunkt verändert sich auch die abgebildete Szene.

Perspektivische Projektionsverfahren, Okklusion, Textur, Schattenwurf und Reflexion erzeugen dabei den Tiefeneindruck. Licht und Schatten in Verbindung mit bestimmten Objekteigenschaften (Texturen) dienen der monokularen Raumwahrnehmung, so dass Objekte als räumliche Gebilde erscheinen (s. Abb. 4.1 b). Texturen sind Oberflächendaten (matt, glänzend, transparent), die die Plastizität virtueller Objekte erhöhen. Verdeckung von Objektteilen durch davor liegenden Objekte (Okklusion) erzeugt den Eindruck einer tatsächlichen räumliche Anordnung (Kulissenbildung).

Abbildung 4.1: (a) Szenen- und Objektbeschreibung. (b) Textur, Okklusion, Licht und Schattenwurf. (c) Zentralprojektion.



In der Zentralprojektion scheinen alle in der Tiefe des Raumes liegenden Kanten sich in einem Fluchtpunkt zu vereinigen. Dadurch entsteht bei parallel in die Tiefe laufenden Kanten der Effekt der stürzenden Linien. Aufgrund des Wissens, dass diese in Wirklichkeit parallel sind, wird eine solche Abbildung als räumliche Situation interpretiert (s. Abb. 4.1 c).

Wie jedoch individuelle raumkognitive Fähigkeiten die Wahrnehmung der Präsenz beeinflussen und *vice versa*, konnte noch nicht einmal ansatzweise geklärt werden. Problematisch ist dabei, dass Präsenz durch Faktoren bestimmt wird, die sowohl Eigenschaften der VR als auch Fähigkeiten der interagierenden Person betreffen (Witmer & Singer, 1998; Albuquerque, Melo & Velho, 2003; Schubert, Friedmann & Regenbrecht, 2001).

4.3 Faktoren virtueller Präsenz

Witmer und Singer (1998) definieren in einem Modell zwei wichtige Komponenten zur Erzeugung einer virtuellen Präsenz. Die Einbindung in Inhalte und Aufgaben der VR (*Involvement*) ist die erste wichtige Komponente und erzeugt durch die Fokussierung der Aufmerksamkeit auf die Reize im virtuellen Raum eine selektive Ausfilterung der Informationen über die reale Umgebung. Auf diese Weise entsteht eine Immersion auf Grund der Relevanz und Salienz der virtuellen Stimuli. Zur zweiten Komponente zählen Witmer und Singer (1998) die Faktorengruppen Kontrolle, Sensorik, Distraktoren und Realismus.

Kontrolle umfasst alle Interaktionsmöglichkeiten mit der VR. Sie erhöht die Wahrnehmung der Präsenz (Warren & Wertheim, 1990) und ist, neben den Raumvorstellungsfähigkeiten einer Person, der zweite bedeutende Prädiktor für die Performanz in der VR (Waller, 2000). Insbesondere die Bewegungskontrolle in VRs sollte dabei immer möglichst intuitiv erfolgen können (Bowman, Koller

& Hodges, 1998). Dazu eignen sich Techniken zur direkten Zielauswahl, die Begrenzung der Bewegung auf zwei Dimensionen und die Animation von Bewegung (statt Teleportation). In VRs, in denen der Nutzer räumliche Informationen erfassen soll, müssen Steuertechniken trainiert werden. Ist Training nicht möglich, sind direkte Zielauswahltechniken vorzuziehen.

Faktoren der Sensorik hängen am engsten mit den technischen Möglichkeiten einer VR zusammen und betreffen alle Bereiche der künstlichen Vermittlung sensorischer Eindrücke. Zu Distraktoren einer virtuellen Umgebungen zählen Artefakte, wie duplizierte Objekte, nicht differenzierbare Landmarken (Veränderungsblindheit), Irrelevanz der Stimuli, Verzerrungen in der Darstellung, fehlende oder unrealistische Manipulierbarkeit von Objekten und Feedback-Verzögerungen (Loomis, Blascovich & Beall, 1999). Unter Realismus einer VR versteht man den Abstraktionsgrad einer virtuellen Umwelt. In diesem Zusammenhang ist wenig über den notwendigen und hinreichenden Realismus zur Erzeugung virtueller Präsenz bekannt. Lediglich einige Arbeiten, die VR zur Konfrontationstherapie bei Phobien einsetzen, konnten zeigen, dass bereits sehr einfache virtuelle Umgebungen starke reale Eindrücke vermitteln und physische Reaktionen hervorrufen können (Schuemie & Mast, 2001).

5 Zusammenfassung und Fragestellung

5.1 Zusammenfassung

Gute Leistungen in Raumvorstellungstests und Navigationsexperimenten gehen vermutlich stärker auf die Anwendung erfolgreicher Strategien zurück als auf die Ausprägung einzelner eindimensionaler Fähigkeiten. Räumliche Strategien sind dabei Ausdruck eines effizienten, aufgabenbezogenen Informationsverarbeitungsprozesses. Dieser wird maßgeblich durch den Aufbau und die Nutzung mentaler Repräsentationen über die räumlichen Strukturen bestimmt. Innerhalb des hier verwendeten Stratiemodells wird dieser Prozess als Zusammenwirken von Kompetenzen und Präferenzen verstanden, wobei Kompetenzen als eindimensionale räumliche Fähigkeiten aufgefasst werden und Präferenzen die Dichotomie der beteiligten mnestischen Systeme widerspiegeln.

Bezüglich der ausgewählten räumlichen Fähigkeiten Place-Learning, Pfadintegration und mentale Rotation sind gruppenspezifische Leistungsunterschiede (Geschlecht, Altersgruppen) häufig berichtet worden. In einigen Fällen konnten diese Differenzen durch die Anwendung unterschiedlicher Lösungs- und Bearbeitungsstrategien aufgeklärt werden. Die Anwendung unterschiedlicher Strategien geht offensichtlich mit der Aktivierung spezieller neuronaler Netzwerke einher. Dies unterstreicht die Bedeutung von Strategien als qualitativen Unterschied in der Verarbeitung räumlicher Informationen.

Unter anderem zur Untersuchung dieser neuronalen Netzwerke werden in der Raumkognitionsforschung häufig Methoden der virtuellen Realität angewandt. Sie sollen in der Praxis aufwendige Navigationsexperimente ersetzen, Verhaltensexperiment und neurophysiologische Datenakquisition koppeln und komparative Transspeziesforschung ermöglichen. Theoretisch wurde jedoch ein neuer Bereich raumkognitiver Fähigkeiten geschaffen, der sich durch die Vorstellung und Erkundung real nicht existierender virtueller Objekte und Räume auszeichnet. Diese Methode steht somit zwischen realen Verhaltensexperimenten und „klassischen“ Raumvorstellungstests. Tendenzen in die eine oder andere Richtung hängen wiederum gleichermaßen von technischen Faktoren sowie von Personeneigenschaften ab, die offensichtlich in komplexer Interaktion den Grad der virtuellen Präsenz beeinflussen.

5.2 Ableitung der Fragestellungen

Im Mittelpunkt steht die Untersuchung qualitativer Alters- und Geschlechtsunterschiede in der Raumkognition. In diesem Zusammenhang wird davon ausgegangen, dass sich Bearbeitungsstrategien in bestimmten, unterscheidbaren Verhaltensmustern äußern. Diese qualitativen Unterschiede sind verantwortlich für die quantitativen Mittelwertsdifferenzen zwischen den Gruppen. Außerdem wird vermutet, dass die Zugehörigkeit zu einer Geschlechts- bzw. Altersgruppe vermittelnden Charakter für eine bestimmte Art der Informationsverarbeitung hat. Qualitative Effekte werden als Anwendung unterschiedlicher räumlicher Strategien verstanden und in ein Konstrukt räumlicher Präferenzen eingeordnet. Hier soll die Frage geklärt werden, unter welchen Umständen diese Strategien eingesetzt werden. Unklar ist bisher, in welchem Umfang die Wahl von Strategien durch externe Signale oder interne Prozesse beeinflusst wird. Durch die Variation des Testmaterials und durch crossmodale Strategieanalysen sollen hier Einflussfaktoren aufgedeckt werden.

Auf explizit, d. h. bewusst, angewandte Strategien kann durch entsprechende Fragebögen zugegriffen werden. Es muss aber davon ausgegangen werden, dass implizite und explizite Prozesse bei der Generierung von Verhalten eine Rolle spielen. Ein Selbstreport kann also nur explizite, deklarative Inhalte wiederspiegeln. Eine Analyse der Verhaltensmuster aus Trajektorien, Antwortpattern etc. ist deshalb notwendig. Diese Art der Analysen wurde in Raumvorstellungstests bereits durch probabilistische Mehr-Klassen-Modelle der Item-Response-Theorie durchgeführt. Ein Nachteil dieser Methoden im Bereich experimenteller Untersuchungen ist die dafür notwendige Stichprobengröße, die im Bereich mehrerer hunderter Versuchspersonen liegen müsste.

Automatische Mustererkennung nach Methoden des unsupervidierten Lernens und Clusteralgorithmen arbeiten auch auf der Basis sehr kleiner Datenmengen mit zufriedenstellendem Erfolg (Webb, 2002; Duda, Hart & Stork, 2000). Die Aussagen dieser Analysen sind allerdings als exploratorische Verfahren immer nur im Kontext der jeweiligen Stichprobe zu sehen. Trotzdem sollen im Rahmen dieser Arbeit erstmals zwei dieser Data-Mining-Verfahren in der Verhaltensforschung angewandt und an realen empirischen Daten erprobt werden.

Es wurden für die vorliegende Arbeit folgende konkreter Fragestellungen formuliert:

1. Lassen sich mit Verfahren der automatischen Mustererkennung Strategien in Raumvorstellungstests und Suchstrategien in Bewegungsdaten analysieren?
2. Bestehen Zusammenhänge zwischen der Strategiewahl und der Zugehörigkeit zu Alters- bzw. Geschlechtergruppen?
3. Haben die Faktoren Alter und Geschlecht in diesem Zusammenhang lediglich einen vermittelnden Charakter für die Anwendung unterschiedlicher Strategien in Aufgaben mit räumlichen Anforderungen?
4. Können die identifizierten räumlichen Strategien den vorgeschlagenen Konstrukten räumlicher Präferenzen (egozentrisch vs. allozentrisch, analytisch vs. holistisch) zugeordnet werden?

5. Gibt es Zusammenhänge in den Ausprägungen zwischen beiden räumlichen Präferenzen?
6. Bleiben diese Präferenzen über verschiedene Testsituationen stabil, d. h. sind sie materialinvariante Personeneigenschaften?

5.2.1 Hypothesen zu den Fragestellungen

Zur Bearbeitung der formulierten Fragestellungen wurden drei Experimente durchgeführt, in denen raumkognitive Fähigkeiten über verschiedene Tests operationalisiert wurden. Unabhängige Variablen sind in jedem der Experimente die kategorialen Faktoren Altersgruppe, Geschlecht und Strategie.

Aus der Vielfältigkeit raumkognitiver Fähigkeiten ergibt sich auch in der vorliegenden Arbeit die Frage nach der Operationalisierung ökologisch valider Testsituationen. Es wurde der Fokus auf die Raumvorstellungsfacette mentale Rotation und die enger mit Navigationsfähigkeiten in Verbindung stehenden Fähigkeiten Place-Learning und Pfadintegration gelegt. In der Implementierung entsprechender Tests wurde auf computergestützte Methoden der virtuellen Realität besonderer Wert gelegt. Im Rahmen der Arbeit wurden ein selbst erstellter Computertest des *Mental Rotations Test* (CMRT), sowie eine selbst entwickelte virtuelle Kompass-Zeigeaufgabe (VPT von *Virtual Pointing Task*) und ein virtuelles Water-Maze (VWM) implementiert.

Experiment 1. Primäres Ziel des ersten Experiments ist die Untersuchung von Lösungs- und Bearbeitungsstrategien in verschiedenen Raumvorstellungstests zur mentalen Rotation. Dazu werden zunächst quantitative Alters- und Geschlechtsunterschiede in Testleistungen aus dem *Mental Rotations Test* (MRT) und den Würfelaufgaben beschrieben.

Für den Geschlechtseffekt werden im MRT und dem Computertest wenigstens $d=0.7$ zugunsten männlicher Versuchspersonen erwartet. In den Würfelaufgaben werden dagegen keine bedeutsamen Differenzen erwartet. Positive Zusammenhänge zwischen beiden Leistungsmaßen werden aber dennoch vermutet.

Hinsichtlich der Altersunterschiede wird angenommen, dass diese zwar vorliegen, der Effekt jedoch gering ausgeprägt ist. Der Geschlechtseffekt bleibt über die Lebensspanne erhalten und nimmt vermutlich zwischen dem 13. und 15. Lebensjahr zu.

Wichtiger sind andere Faktoren, die sich oft unbemerkt als Störfaktoren in Studien befinden. Das sind z. B. Schulleistung, Studienfach, Computererfahrung (besonders Computerspiele) und Vorerfahrung mit ähnlichen Raumvorstellungstests. Da dieser Einfluss vermutet wird, werden die entsprechenden Faktoren kontrolliert.

Der Faktor Strategie wird exploratorisch mit einem alternativen Analyseverfahren auf den Antwortmustern des MRT und der Würfelaufgaben untersucht. Diese Tests sind bereits mit Modellen der IRT analysiert worden. Eine Konvergenz mit den in Geiser (2004), Geiser et al. (2006) und

Hosenfeld et al. (1997) berichteten Resultaten soll hier als erste Überprüfung der Validität des eingesetzten Klassifizierungsverfahrens angesehen werden.

Dabei wird davon ausgegangen, dass sich Versuchspersonen hinsichtlich der Antwortmuster klassifizieren lassen. Diese Muster repräsentieren typische Lösungsstrategien, die wiederum unter Ausnutzung struktureller Besonderheiten der Tests in ihrer Bedeutung interpretiert werden können. Dazu wird angenommen, dass bestimmte Itemtypen innerhalb eines Tests die Anwendung einer bestimmten Strategie erfordern und sich Personen verschiedener Klassen diesbezüglich unterscheiden.

Auf Grund des Zusammenhangs zwischen dem Konstrukt räumlicher Präferenzen und Strategien wird von wenigstens einer dichotomen Klassifizierung in beiden Tests ausgegangen. Jedoch werden in Anlehnung an die Befunde in Geiser (2004) nicht mehr als fünf Klassen für plausibel gehalten.

Weiterhin konnten in Geiser (2004) Zusammenhänge zwischen der Art der verwendeten Strategie und dem Geschlecht bereits gezeigt werden. Im MRT werden sich deshalb auch hier männliche Versuchspersonen eher in Klassen mit effizienten holistischen Strategien und weibliche Personen in Klassen mit langsameren und eher analytischen Strategien befinden. Für die Würfelaufgaben wird dagegen nach Hosenfeld et al. (1997) nicht von einer geschlechtsspezifischen Strategiewahl ausgegangen.

In Geiser (2004) wurden ebenfalls sehr geringe Zusammenhänge zwischen Strategien und dem Alter der Versuchspersonen nachgewiesen. Nicht genau betrachtet wurde dort die Verteilung der Strategien in der Spanne zwischen dem 13. und 15. Lebensalter. Hier wird generell eine Zunahme effizienter holistischer Strategien erwartet. Darüber hinaus müssten auf Grund der Zunahme der Geschlechtseffekte auch eine überproportionale Zunahme dieser Strategien bei männlichen Versuchspersonen zu beobachten sein.

Zu den Zusammenhängen zwischen der Strategiewahl in beiden Tests gibt es bisher keine Erkenntnisse. Im MRT konnte gezeigt werden, dass es zwischen den Testhälften nur wenige „Klassenwechsler“ gibt. So wird vermutet, dass Personen in beiden Tests entweder eher holistisch oder eher analytisch vorgehen.

Neben dem bekannten MRT wird als drittes Maß zur mentalen Rotation auch eine eigene Adaption des Tests als Computerversion eingesetzt. Dazu sollen als sekundäres Ziel der Untersuchung erste Erkenntnisse zur konvergenten und diskriminanten Validität des Computertests gesammelt werden. Besonders bei einem Computertest mit Zeitbegrenzung könnte sich die Vorerfahrung im Umgang mit Computern als Faktor auf die Testleistung auswirken. Es wird aber angenommen, insofern dieser Faktor kontrolliert wird, dass sich die Papier-Bleistift-Version und die Computerversion nicht unterscheiden.

Experiment 2. In diesem Experiment wird das Problem untersucht, inwieweit Pfadintegrationsprozesse durch einen Computertest operationalisiert werden können, der keine aktive Bewegung des Probanden erfordert. Es wird insbesondere der Frage nachgegangen, ob virtuelle Umgebungen

entworfen werden können, in denen „reine“ Pfandintegration angewendet wird, d. h. keine Nutzung von Landmarken und kein Erwerb von Überblickswissen möglich ist.

Dazu werden folgende Annahmen zur Operationalisierung des Konstrukts Pfandintegration (PI) gemacht: Die PI-Leistung kann über Fehler in der Schätzung der Heimkehrrichtung gemessen werden. Die PI-Leistung wird beeinflusst durch die Komplexität der Pfade. Bei Anwendung von Pfandintegration sind die Schätzfehler systematisch, d. h. um einen bestimmten Mittelwert normalverteilt. Beim Raten sind die Fehler unsystematisch, d. h. gleichverteilt. Multimodale Verteilungen der Schätzfehler lassen auf mögliche Subgruppen schließen, die entweder Pfandintegration oder andere Strategien einsetzen.

Die Komplexität eines Pfades wird über die Anzahl der Abzweigungen und Winkel sowie die Darbeitungsgeschwindigkeit variiert. Es wird angenommen, dass die Schätzfehler mit wachsender Komplexität generell zunehmen. Reaktionszeiten und Schätzfehler steigen vermutlich mit dem Alter der Versuchspersonen an. Es wird auch mit Geschlechtsdifferenzen in der Ausprägung gerechnet, wobei die Schätzfehler weiblicher Versuchspersonen im Mittel um 20° über denen männlicher Personen liegen müssten (Lawton & Morrin, 1999).

Es wird vermutet, dass PI in virtuellen Umgebungen und andere Facetten der Raumvorstellung, wie der mentalen Rotation, dissoziierte Fähigkeiten sind. Leistungsmaße aus dem VPT korrelieren vermutlich deshalb nur gering mit Raumvorstellungstest.

Es wird weiterhin angenommen, dass sich auch in der VPT unterschiedliche Lösungsstrategien identifizieren lassen. Mit dem von Waller und Hodgson (2006) beschriebenen Desorientierungseffekt sollte es möglich sein, zwischen egozentrischen und allozentrischen Orientierungsstrategien zu differenzieren. In Versuchsbedingungen mit Drehungen um 135° sollte dieser Effekt auftreten. Es wird erwartet, dass Personen, die in diesen Bedingungen keine signifikante Erhöhung der Schätzfehler zeigen, zusätzlich zu den egozentrischen, durch PI aufgebauten Raumrepräsentationen, über eine allozentrische kognitive Karte der Umgebung verfügen, um damit die Desorientierung zu kompensieren. Dieser Unterschied soll zur Klassifizierung ausgenutzt werden.

Experiment 3. Im abschließenden Experiment der Arbeit werden Zusammenhänge zwischen den räumlichen Fähigkeiten Place-Learning, Pfandintegration und mentaler Rotation untersucht. Zunächst wird dazu die Operationalisierung der Fähigkeit Place-Learning durch die eigene Versuchsanordnung des VWM überprüft. Die Navigation im VWM erfordert die Integration räumlich-mnestischen Fähigkeiten und der effizienten Bewegungssteuerung in einer virtuellen Umgebung. Durch Aggregation verschiedener Merkmale aus den VWM-Spurdaten sollen diese Dimensionen quantifiziert werden. Für diese quantitativen Maße wird zunächst überprüft, ob bekannte Alters- und Geschlechtseffekte mit der eigenen virtuellen Versuchsanordnung reproduziert werden können. Für die Geschlechtsdifferenzen werden laut Astur et al. (1998) je nach verwendetem Leistungsmaß Effekte zwischen $d=0.75$ und $d=1.0$ erwartet. Altersunterschiede müssten sich im präzisen Wiedererinnern der Ziellokation und in speziellen Verhaltensmustern in der Lernphase zeigen.

Da davon ausgegangen wird, dass ein Teil der Alters- und Geschlechtsunterschiede im VWM auf

die Anwendung unterschiedlicher Strategien zurückzuführen ist, sollen automatische Musteranalysen der Spurdaten zur Identifikation dieser Strategien genutzt werden. Es wird erwartet, dass sich mit den genannten Analyseverfahren Subgruppen identifizieren lassen, die sich voraussichtlich zu unterschiedlichen Anteilen aus Kombinationen von räumlichen und nicht-räumlichen, egozentrischen und allozentrischen Suchstrategien zusammensetzen.

Place-Learning, Pfadintegration und mentale Rotation werden weitgehend als dissoziierte räumliche Fähigkeiten angesehen, so dass nur schwache Zusammenhänge erwartet werden. Nichtsdestotrotz soll in einem exploratorischem, bisher noch nicht versuchten, Vorgehen nach stärkeren Zusammenhängen zwischen diesen Teilbereichen auf der Ebenen der eingesetzten Strategien gesucht werden.

5.2.2 Fragen der Softwareentwicklung

Da über die empirischen Untersuchungen dieser Arbeit hinaus die eingesetzten computergestützten Verfahren in weiteren Studien Anwendung finden sollen, liegt auf der Implementierung der Tests als Software ein weiterer Schwerpunkt. Als Zielsetzung für diesen Teil der Arbeit wurden besondere Anforderungen an die zu entwickelnden Programme gestellt: Die Programme stellen parametrisierte Setups virtueller Umgebungen für psychometrische Tests bzw. Experimente zur Verfügung. Sie erlauben dadurch eine Modifizierung von Material, Aufbau und Ablauf eines Versuchs ohne besondere Programmierkenntnisse. Dazu sind Schnittstellen zu schaffen über die der Versuchsleiter die Anpassungen vornehmen kann.

Außerdem sollten alle computergestützten Experimente, insbesondere solche, die Fortbewegungen in virtuellen Umgebungen erfordern, komfortable Eingabegeräte und Kontrollmechanismen verwenden. Dadurch soll eine hohe Akzeptanz der Aufgaben bei den Versuchspersonen gewährleistet werden. In diesem Zusammenhang sind, falls erforderlich, zu einzelnen Tests und Experimenten Übungsaufgaben bzw. Trainingsdurchgänge zu implementieren.

Neben einer vollständigen Protokollierung aller im Test anfallenden Rohdaten sollen häufig gebräuchliche Messwerte in gängige, intuitiv verständliche Datenformaten abgelegt werden, z. B. als Tabellen für Microsoft[®] EXCEL (Microsoft Corp.) oder SPSS[®] (SPSS, Inc.). Für Experimente mit komplexeren Rohdaten, wie die Spurdaten aus dem VWM, sind Hilfsprogramme (Tools) zur Visualisierung und Aggregation weiterer quantitativer Leistungsmerkmale zu programmieren.

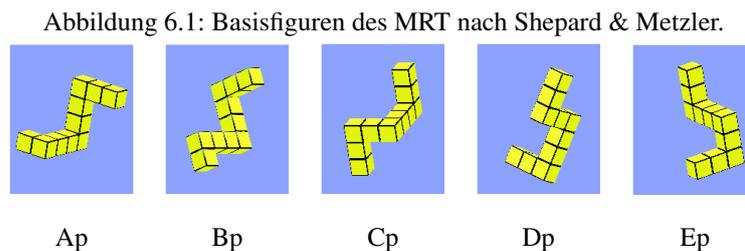
Methodenteil

6 Softwareentwicklung

6.1 Computer Mental Rotations Test

6.1.1 Material und Aufbau

Für die Implementierung des MRT (Peters et al., 1995) als Computertest war es zu nächst notwendig, sich genauer mit dem verwendeten Testmaterial auseinander zusetzen. Der Originaltest basiert auf Figuren aus einem Experiment von Shepard und Metzler (1971) (Shepard-Metzler-Figuren). Diese bildeten auch die Basis des Computertests. Abbildung 6.1 zeigt sie in der Darstellung für die Computerversion.



Es handelt sich dabei um aus einzelnen Würfeln zusammengesetzte, dreidimensionale Objekte. Shepard und Metzler (1971) verwendeten die in der Abbildung 6.1 dargestellten fünf Basisfiguren (Ap, Bp, Cp, Dp, Ep) und fünf spiegelbildlich zusammengesetzte Figuren (An, Bn, Cn, Dn, En). Jede dieser Figuren bestand aus zehn identischen Würfeln der Kantenlänge eins. Auf Grund dieser Systematik konnten alle Figuren aus den Replikas eines einzigen Würfels zusammengesetzt werden. Das 3D-Modell diese Würfels wurde mit 3d Studio Max (Autodesk Inc.) entworfen.

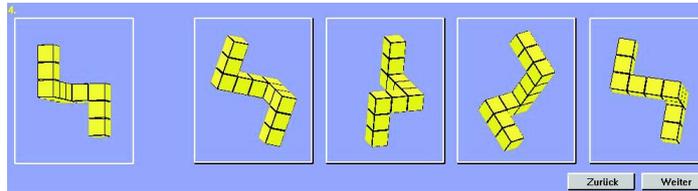
Das Zusammensetzen der Figuren geschah zur Laufzeit des Programms und wurde durch eine einfache Konfigurationsdatei gesteuert. Es war somit für den Versuchsleiter möglich, die Computerversion des MRT um zusätzliche, anders zusammengesetzte Basisfiguren zu erweitern. Nähere Erläuterungen zu diesen technischen Möglichkeiten befinden sich im Anhang 15.

Zur Operationalisierung der mentalen Rotation leiteten Shepard und Metzler (1971) zweidimensionale Abbildungen von jeder der Figuren aus sieben verschiedenen Perspektiven ab. Zur Projektion der 3D-Objekte wurde dazu die Parallelperspektive gewählt. Sie veränderten den Blickpunkt auf die Figur, indem sie sie schrittweise um die Körperlängsachse rotieren ließen. Beginnend von einer Ausrichtung der Figur im Raum, aus der diese gut zu erkennen war, drehten sie die Figur in den

Schrittweiten 20°, 80°, 80°, 20°, 40°, 80° und 40°. Durch diese Systematik entstanden insgesamt 70 Abbildungen, die auch das Material für den MRT bildeten.

Im MRT wurden allerdings die Abbildungen der Figuren offensichtlich weiteren Transformationen unterzogen. Insbesondere wurden einige Abbildungen zusätzlich in der Bildebene rotiert. Dies wurde in der Implementierung des Computertests ebenfalls berücksichtigt.

Abbildung 6.2: Ein Item aus der Computerversion des MRT.



Die Items im MRT bestehen aus jeweils fünf solcher Abbildungen, die jeweils in einem kleinen Rahmen dargestellt sind (s. Abb. 6.2). Auf der linken Seite befindet sich die Zielfigur, auf der rechten Seite, etwas davon abgesetzt, vier Alternativfiguren. Zwei der Alternativfiguren zeigen die Zielfigur aus einer anderen Perspektive und stellen die richtige Lösung des Items dar. Die beiden anderen Figuren sind Distraktoren, die nach unterschiedlichen Gesichtspunkten ausgewählt wurden. Zeigen beide Distraktoren eine gespiegelte Figur der Zielfigur, handelt es sich um ein „Spiegel-“ oder Typ-I-Item. Sind beide Distraktoren jedoch anders zusammengesetzte Figuren, ist diese ein Typ-II-Item bzw. „Strukturitem“. Gemischte Items bzw. Typ-III-Items beinhalten beide Arten von Distraktoren.

Tabelle 6.1 zeigt die Klassifizierung der 24 Items aus dem MRT-A und MRT-B (Peters et al., 1995), wie sie bei der Digitalisierung des originalen Testmaterials ermittelt wurden. Entgegen den Arbeiten von Geiser (2004) und Voyer und Hou (2006) enthält der MRT-A nur ein Typ-III-Item und nicht wie bisher publiziert wurde, drei bzw. zwei Items diesen Typs. Zum MRT-B gab es bisher noch keine Untersuchung zur Struktur der Items.

Tabelle 6.1: Itemtypen des MRT. Klassifizierung durch die Art der Distraktoren bzw. durch das Vorhandensein von Figuren mit Okklusion.

	Typ-I-Items	Typ-II-Items	Typ-III-Items
MRT-A	1, 2, 5, 6, 9, 10, 11, 12, 15, 16, 19, 20	3, 4, 7, 8, 13, 14, 17, 18, 21, 22, 24	23
MRT-B	1, 2, 3, 4, 7, 8, 9, 11, 12, 14, 16, 19, 22, 23, 24	5, 6	10, 13, 15, 17, 18, 20, 21
	Items mit Okklusion		Items ohne Okklusion
MRT-A	9, 10, 11, 14, 15, 17, 18, 19		1, 2, 3, 4, 5, 6, 7, 8, 12, 13, 16, 20, 21, 22, 23, 24
MRT-B	18, 24		1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23

In einigen Items des MRT gibt es Figuren mit Okklusion. In Anlehnung an die Definition von Voyer und Hou (2006) bedeutet dies, dass wenigstens ein Segment der Figur nicht sichtbar oder dessen Dreidimensionalität nicht eindeutig zu erkennen ist. Entgegen der Arbeit von Voyer und Hou (2006) befinden sich dann laut ihrer eigenen Definition nicht sieben sondern acht Items mit Okklusion im Test (s. Tab. 6.1).

Aus der Tabelle 6.1 wird sofort deutlich, dass entgegen der Angaben in Peters et al. (1995) die Testformen A und B nicht als Parallelversionen, in denen lediglich die Reihenfolge der Items verändert wurde, vorliegen. Hinzu kam ein offensichtlicher Fehler im Originaltestheft des MRT-B. Das Item 24 enthielt, scheinbar bisher unbemerkt, unter den vier Alternativfiguren drei Abbildungen der Zielfigur. Im Computertest wurde deshalb eine dieser Abbildungen durch einen Distraktor ausgetauscht. Im Anhang 14 kann der revidierte Test in der Papier-Bleistift-Version eingesehen werden.

Um die strukturellen Besonderheiten und Abweichungen im Material des MRT im Computertest zu berücksichtigen, wurde eine Implementierung gewählt, die eine flexible Zusammenstellung einzelner Items ermöglicht. Jede Figur wird in einem eigenen Fenster („Viewport“) gerendert. Die Größe eines Viewports, und somit auch die Größe der 3D-Figur, ist beliebig skalierbar. Außerdem kann die Position auf dem Bildschirm beliebig gewählt werden. Es ist außerdem möglich zu kennzeichnen, ob ein Viewport eine Zielfigur oder eine Alternativfigur darstellt. Viewports, die eine Alternativfigur darstellen, konnten später im Test mit einem Mausklick ausgewählt werden.

Eine Besonderheit der Computerversion ist die Möglichkeit, die Figur in einem Viewport so zu konfigurieren, dass sie über Tastatur- oder Maussteuerung rotiert werden kann. Dadurch können mit der gleichen Software auch Trainingsprogramme zur mentalen Rotation entwickelt werden. Weitere Details dazu befinden sich ebenfalls im Anhang 15.

6.1.2 Protokoll

Vor der Bearbeitung des Tests wird den Versuchspersonen der Begriff der mentalen Rotation erläutert. Dazu wird eine der Basisfiguren aus unterschiedlichen Perspektiven gezeigt. Im Computertest sind dazu zwei animierte Figuren dargestellt, die langsam um sich selbst rotieren. Erst nachdem alle Versuchspersonen sich diese Transformationen vorstellen können, werden vier Übungsitens bearbeitet. Ein Item im Computertest wird bearbeitet, indem die Versuchsperson mit der Maus auf die gewählten Alternativfiguren klickt. Eine ausgewählte Figur wird durch ein großes rotes Kreuz gekennzeichnet; erneutes Anklicken löscht das Kreuz. Die richtigen Lösungen der Übungsitens werden nach der Bearbeitung genannt. Dieser Teil des MRT ist in den Testformen A und B identisch.

Der Test selbst besteht aus zwei Testhälften; die Items eins bis zwölf bilden die erste Testhälfte und die Items 13 bis 24 die zweite Testhälfte. Die Items erscheinen einzeln und nacheinander auf dem Bildschirm. Innerhalb einer Testhälfte kann zwischen den Items geblättert werden. Dazu sind zwei Buttons mit den Beschriftungen „Zurück“ bzw. „Weiter“ vorhanden. Zwischen den Testhälften liegt eine zweiminütige Pause. Jede Testhälfte kann durch Anklicken eines Startbuttons individuell

gestartet werden. Nach Beendigung der zweiten Testhälfte wird die erreichte Punktzahl berechnet und auf dem Bildschirm ausgegeben.

Bei der Entwicklung des Computertests wurde darauf Wert gelegt, dass Testinstruktionen, Zeitlimits, Itemreihenfolge und -zusammensetzung in einfacher Weise verändert werden können. Kontrolle über diese Parameter hat der Versuchsleiter über eine Konfigurationsdatei (Erläuterungen im Anhang 15). Auf diese Weise sind im Rahmen der Softwareentwicklung verschiedene Konfigurationen des MRT-A und des MRT-B erstellt wurden. Diese Konfigurationen unterscheiden sich in den Instruktionen (Texte für Kinder oder Erwachsene) und in den Zeitlimits (drei oder fünf Minuten pro Testhälfte). Die Auswahl einer der Testversionen erfolgt über ein Menü, das direkt nach dem Programmstart erscheint.

6.1.3 Punktvergabesysteme

Für die statistische Auswertung der MRT-Leistung wird bereits während des Tests eine formatierte Protokolldatei angelegt, in der die Punktzahl für jedes Item gespeichert wird. Näherer Erläuterungen zu diesem Datenformat befinden sich im Anhang 15. Punkte werden nach einem strengen und einem liberalen Punktesystem vergeben. Im strengen Punktesystem wird nur dann ein Punkt pro Item vergeben, wenn beide Alternativfiguren richtig angekreuzt wurden. War nur eine Figur richtig und die zweite Figur falsch oder gar nicht angekreuzt, so wird dieses Item mit null Punkten bewertet. Damit sind maximal 24 Punkte im gesamten Test möglich. Im liberalen Punktesystem sind dagegen maximal 48 Punkte möglich, da hier jede richtig angekreuzte Figur mit einem Punkt bewertet wird.

6.2 Kompass-Zeigeaufgabe

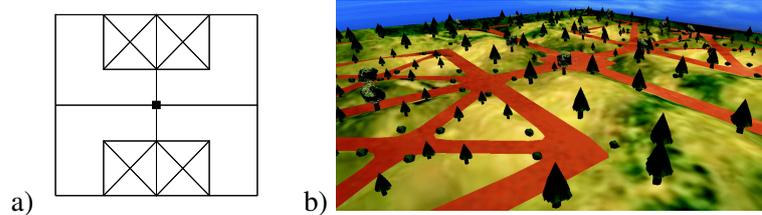
6.2.1 Virtuelle Umgebung

Für die Durchführung der Kompass-Zeigeaufgabe in der VR wurde eine stark vereinfachte Parklandschaft entworfen. Abbildung 6.3 zeigt im Grundriss und aus der Vogelperspektive die zahlreichen Wege der Parkanlage, die symmetrisch um deren Mittelpunkt angeordnet sind. Die Fortbewegung erfolgt ausschließlich auf diesen Wegen. Im virtuellen Park gibt es keine salienten Landmarken oder proximalen Hinweisreize, die zur Orientierung dienen könnten. Stattdessen soll nur durch die Wahrnehmung der Eigenbewegung aus dem optischen Fluss die eigene Position im Raum ermittelt werden.

Dazu wurden Texturen für die Wege verwendet, in der sich helle und dunkle Bereiche abwechseln. Entlang der Wege vermitteln mehrere identisch Bäume und Büsche eine strukturierte Landschaft. Abzweigungen dienen der Segmentierung der Wege. Über diese Maßnahmen wird die Abschätzung der Geschwindigkeit ermöglicht, die zur Positionsbestimmung durch Pfadintegration notwendig ist.

Um zu verhindern, dass nach wenigen Läufen durch den virtuellen Park die Versuchspersonen Überblickswissen aus der geometrischen Anordnung der Wege einander erwerben, befinden sich

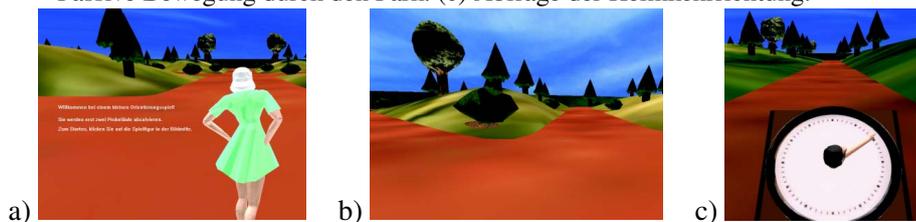
Abbildung 6.3: Virtuelle Umgebung der Parklandschaft für die Kompass-Zeigeaufgabe. (a) Grundriss. (b) Umgebung aus der Vogelperspektive.



Bodenerhebungen dazwischen, die den Blick auf benachbarte Wege versperren.

Die Fortbewegung durch den Park erfolgt ausschließlich passiv und wird aus einer egozentrischen Perspektive wahrgenommen, d. h. Versuchspersonen müssen nicht selbst navigieren, sondern nur den computeranimierten Bewegungen mental folgen. Jeder Lauf durch den Park startet vom Mittelpunkt der Parkanlage aus und beginnt mit dem Blick auf eine Spielfigur (s. Abb. 6.4 a). Dabei kann beim Start des Experimentes zwischen einer weiblichen oder männlichen Figur gewählt werden, die als Avatar der Versuchsperson während des Experiments dient.

Abbildung 6.4: Ein Lauf durch die virtuelle Parklandschaft. (a) Start durch Anklicken der Spielfigur. (b) Passive Bewegung durch den Park. (c) Abfrage der Heimkehrrichtung.



Die Versuchsperson klickt zum Start eines neuen Laufs mit der Maus auf die Spielfigur und sieht als nächstes, wie der Blickpunkt in deren Perspektive zoomt. Durch diese Animation wird nonverbal vermittelt, dass sich die Versuchsperson in dem Avatar durch die virtuelle Umgebung bewegt. Die egozentrische Perspektive wird dann für den gesamten Lauf beibehalten. Ein Lauf wird in konstanter Geschwindigkeit absolviert und führt über unterschiedlich lange Wegeabschnitte und Abzweigungen in verschiedene Richtungen.

Am Ende jedes Laufs wird ein Winkelmesser mit einem Zeiger zur Abfrage der Heimkehrrichtung eingeblendet (s. Abb 6.4 c). Der Winkelmesser ist ein kompassähnliches Instrument mit Skaleneinteilung einer analogen Uhr mit 60 Teilstrichen. Der Zeiger wird per Drag-&-Drop bedient. Durch Klicken der linken Maustaste wird der Zeiger aufgenommen, in die gewünschte Richtung gedreht und durch Loslassen der Maustaste dort fixiert. Um diese Eingabe zu bestätigen und den Lauf abzuschließen, muss mit der rechten Maustaste neben den Kompass geklickt werden. Danach wird die Versuchsperson an den Ausgangspunkt teleportiert und ein weiterer Lauf beginnt.

Technisch besteht ein solcher Lauf aus einer Menge von Spurdaten. Das sind Positionen und Blickrichtungen in diskreten Zeitabständen, die zuvor berechnet wurden. Während der Durchführung des Experiments werden diese mit der VR-Software „abgespielt“. Die Menge der Daten zu einem

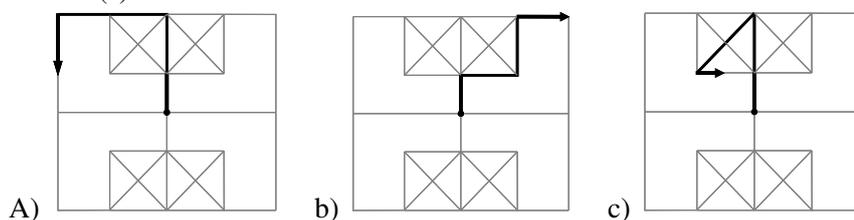
einzelnen Lauf werden als Pfad bezeichnet. 50 solcher Pfade wurden für den virtuellen Park im Rahmen der Softwareentwicklung berechnet. Diese kann ein Versuchsleiter beliebig zu einem Experiment zusammenstellen. Näheres zur Erstellung neuer Pfade und deren Einbindung in das Setup der Zeigeaufgabe erläutert der Anhang 15.

Die Software zur Zeigeaufgabe wurde mit World Up® (Sense8 Inc.) und 3d Studio Max (Autodesk Inc.) für Microsoft Windows 2000/XP erstellt. Die Zeigeaufgabe ist auf Computersystemen mit unterschiedlicher Rechenleistung einsetzbar. Eine automatische Anpassung der Geschwindigkeit an die Rechenleistung gewährleistet eine konstante Abspielgeschwindigkeit der Pfade und somit die Vergleichbarkeit der Ergebnisse von Experimenten, die auf unterschiedlicher Hardware durchgeführt wurden.

6.2.2 Protokoll

Die Durchführung eines vollständigen Versuchs mit der Kompass-Zeigeaufgabe wird unmittelbar mit Testdurchgängen begonnen. Da keine aktive Navigationsleistung erforderlich ist, wird auf Übungsdurchgänge verzichtet (Bowman, Johnson & Hodges, 2001). Die einzelnen Bedingungen variieren im der Aufbau der Pfade und der Darbietungsgeschwindigkeit. In Abbildung 6.5 sind die im Versuchsaufbau verwendeten Pfade dargestellt.

Abbildung 6.5: Pfade der 6 Versuchsbedingungen im VPT. (a) Pfad-1: $2 \times 90^\circ$ -Turns (b) Pfad-2: $3 \times 90^\circ$ -Turns (c) Pfad-3: $2 \times 135^\circ$ -Turns

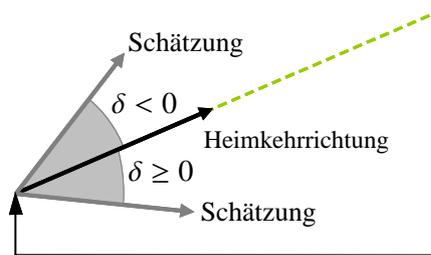


Pfad-1 und Pfad-2 variieren in der Anzahl rechtwinkliger Abzweigungen. Pfad-3 wurde in die Versuchsanordnung aufgenommen zur Untersuchung des Desorientierungseffekts bei Drehungen um mehr als 90° . Diese drei Pfade werden in den Durchgängen eins bis drei in dieser Reihenfolge gelau- fen. In den Durchgängen vier bis sechs werden sie mit der doppelten Geschwindigkeit wiederholt.

6.2.3 Messwerte

Die Bewertung der Testleistung in der Zeigeaufgabe erfolgt über die Genauigkeit der Schätzung der Heimkehrrichtung. Das Programm bestimmt dazu den Winkel zwischen der Heimkehrrichtung und der auf dem Kompass eingestellten geschätzten Richtung (s. Abb 6.6). Dieser Winkel δ ist definiert als die kleinste Winkeldifferenz zwischen den beiden Richtungen und liegt in einem Wertebereich von -180° bis 180° mit $\delta \geq 0$ für Schätzungen, die rechts der Heimkehrrichtung, und $\delta < 0$ für Schätzungen, die links der Heimkehrrichtung liegen.

Abbildung 6.6: Berechnung des Schätzfehlers δ als Winkel zwischen geschätzter (grau) und tatsächlicher Heimkehrrichtung. Schätzungen links der Heimkehrrichtung ergeben negative Fehlerwerte; rechts positive.



Für die Ermittlung des Schätzfehles wird δ noch auf die Skaleneinteilung des Kompassinstrumentes umgerechnet. Da 60 Teilstriche auf dem Vollkreis abgebildet sind, bestimmt sich der endgültige Wert aus der ganzzahligen Division von δ durch sechs. Damit ergibt sich als Messwert für den Schätzfehler ein Wertebereich von -30 bis 30. In den statistischen Auswertungen wird oft nur der Betrag des Schätzfehlers berücksichtigt, d. h. ein absoluter Schätzfehler, der Werte zwischen null und 30 annehmen kann.

Der Schätzfehler und eine Reaktionszeit in Millisekunden werden nach jedem Durchgang in eine Protokolldatei geschrieben. Die Reaktionszeit wird als Latenz zwischen dem Erscheinen des Kompasses und der Bestätigung der Eingabe mit der rechten Maustaste ermittelt.

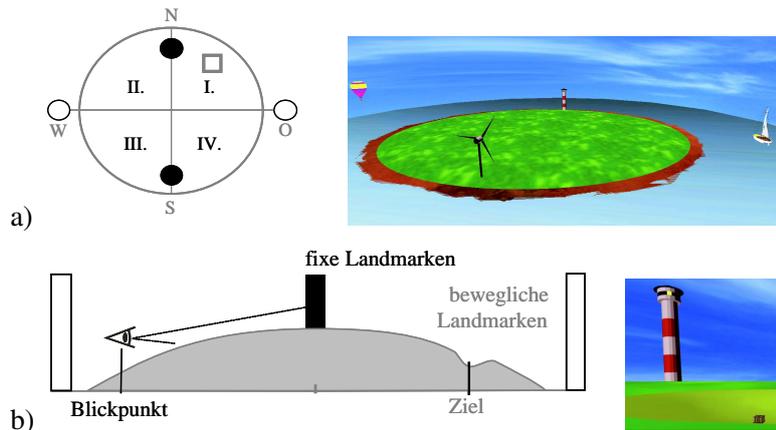
6.3 Water-Maze-Aufgabe

6.3.1 Virtuelle Umgebungen

Für die Umsetzung der Water-Maze-Aufgabe wurden insgesamt zwei virtuelle Umgebungen entworfen: ein virtueller Trainingskurs (VT) und das virtuelle Water-Maze (VWM). Der Trainingskurs wird als Übungsumgebung verwendet, in der sich die Probanden mit der Darstellung einer computergenerierten Umgebung und dem Umgang mit den Eingabegeräten vertraut machen können. Über das VWM wird dann die Aufgabe zum Place-Learning operationalisiert.

Um die ökologische Validität der VR-gestützten Aufgabe zu erhöhen, wurde das VT und das VWM als „Landversion“ gestaltet. Die Versuchsperson bewegt sich nicht in einem Wassertank, sondern auf einer kreisrunden Insel (Abb 6.7). Die versteckte Plattform ist eine Schatzkiste, in der sich ein, für eine auf einer Insel gestrandeten Person, nützlicher Gegenstand befindet. Das Programm wählt diesen zufällig aus insgesamt sieben Objekten aus. Die Kiste öffnet sich automatisch, wenn die Versuchsperson an diese herantritt. Um die Kiste, ähnlich der Plattform aus dem Tierexperiment, möglichst gut zu verbergen, wurden zwei gestalterische Mittel eingesetzt. Erstens: Der Boden der Insel ist leicht gewölbt, um den visuellen Horizont zu verkleinern, d. h. Teile der Landschaft werden erst bei näherem Herangehen sichtbar. Zweitens ist die Schatzkiste in einer kleinen Vertiefung im Boden versteckt (s. Abb. 6.7 b). Die Ziellokation ist so optimal verborgen und kann erst aus unmittelbarer Nähe wahrgenommen werden.

Abbildung 6.7: Prinzipieller Aufbau der Insel für die Water-Maze-Aufgabe. (a) Grundriss und Vogelperspektive. (b) Gestalterische Mittel zum Verstecken der Plattform und Blick auf die Ziellokation aus der Perspektive der Versuchsperson.



Weiterhin sind auf der Insel Landmarken und proximale Hinweiserreize platziert (s. Abb. 6.7 a). Im Rahmen der Parametrisierung der VR besteht die Möglichkeit, verschiedene Landmarken auszuwählen und deren Standorte beliebig zu verändern. Technische Details dazu können im Anhang 15 nachgelesen werden. In einer Art Standardversion, die auch im Experiment 3 eingesetzt wurde, befanden sich vier Landmarken an den folgenden festen Positionen: im Norden ein Leuchtturm, im Süden ein Windrad, im Osten eine Segelyacht und im Westen ein Heißluftballon. Die Landmarken unterscheiden sich hinsichtlich ihrer Semantik in „fixe“ und „bewegliche“ (in Abb. 6.7 schwarz und weiß dargestellt). Die Schatzkiste befindet sich im I. Quadranten der Insel in der Mitte der Vertiefung. Die zur Ziellokation nächste Landmarke ist der Leuchtturm im Norden der Insel.

In der Umgebung des VT sind keine Landmarken vorhanden. Stattdessen befinden sich sechs identische Schatzkisten in einem Parcours innerhalb des II. Quadranten auf der Insel. Die Vertiefung liegt diesmal ebenfalls in diesem Quadranten und versteckte eine siebente Schatzkiste. Die Kisten sind mit darüber schwebenden rotierend Zahlen gekennzeichnet, die aus jeder Position gut sichtbar sind. Diese Zahlen legen die Reihenfolge fest, in der diese Kisten geöffnet werden sollen. In den Kisten befindet sich jeweils ein anderes Objekt. Position und Inhalte der Kisten können im Trainingsparcours ebenfalls beliebig konfiguriert werden (s. Anhang 15).

Probanden laufen mit einer wahrgenommenen Geschwindigkeit von konstant¹ zwei Meter pro Sekunde über die Insel. Legt man diese Geschwindigkeit zugrunde, hat die gesamte Insel eine wahrgenommene Ausdehnung von ca. 180 Metern im Durchmesser. Die Vertiefung hat einen Durchmesser von acht Metern und eine Tiefe von einem Meter. Die Schatzkiste erscheint in einer Größe von ca. 50 × 50 × 30 Zentimetern und Landmarken mit einer Höhe von 15 bis 20 Metern.

In dieser virtuellen Umgebung trägt die Textur des Bodens entscheidend zur Wahrnehmung von Bewegung und Geschwindigkeit bei. Es wurde eine detailarme, aus unregelmäßigen ineinanderfließenden hellen und dunklen Flächen bestehende Grafik verwendet (s. Abb. 6.7 a rechts). Die

¹Konstante Geschwindigkeit wurde durch automatische Schrittweitenanpassung an die Rechenzeit (Framerate) gewährleistet.

Farbwechsel bilden ein wiederkehrendes Muster, so dass die Versuchsperson aus der Ansicht des Untergrunds nicht auf ihre Position auf der Insel schließen kann.

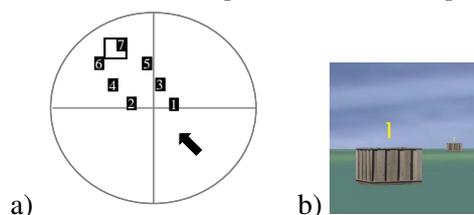
Um den realistischen Eindruck der Umgebung zu erhöhen sind, einige Objekte mit typischen Bewegungen animiert wurden. Segelyacht und Boje bewegen sich in schlingernden Schwimmbewegungen, das Windrad rotiert und am Leuchtturm kreist ein Spotlicht. Über der Insel befindet sich eine rotierende Kuppel mit Wolkentextur, so dass der Eindruck scheinbar über den Himmel ziehender Wolken entsteht.

Die Software zu den virtuellen Umgebungen wurde mit dem VR-Autorensystem World Up[®] (Sense8 Inc.) und Visual Basic[®] (Microsoft Inc.) für die Verwendung auf PCs mit Microsoft Windows-Betriebssystemen erstellt. Alle 3D-Objekte sind mit 3d Studio Max (Autodesk Inc.) entworfene Modelle. Die Programme VT und VWM sind auch auf verschiedenen Computersystemen mit unterschiedlicher Rechenleistung bei Gewährleistung der Vergleichbarkeit der Ergebnisse einsetzbar. Dies wird durch die automatische Anpassung der Geschwindigkeit an die Rechenleistung gewährleistet.

6.3.2 Protokoll

Trainingsdurchgang. Vor dem eigentlichen Versuch im VWM führen die Versuchspersonen ein Training durch. Sie beginnen dazu im IV. Quadranten der Insel mit Blick in Richtung eines Kistenparcours im II. Quadranten (s. Abb. 6.8). Die Aufgabe besteht darin, die Kisten in der Reihenfolge eins bis sieben zu öffnen. Damit soll eine präzise Ansteuerung der Ziellokation im anschließenden Water-Maze-Experiment trainiert werden. Die Instruktion der Aufgabe erhalten die Probanden

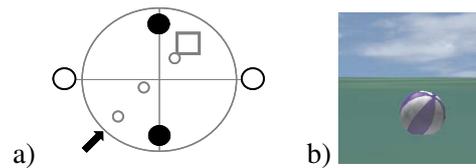
Abbildung 6.8: Virtuelle Trainingsaufgabe. (a) Grundriss mit Startposition (Pfeil) und Positionen der Kisten. (b) Blick auf eine Kiste aus der Perspektive der Versuchsperson.



verbal vom Versuchsleiter und zusätzlich schriftlich über den Computerbildschirm. Maximal kann sechs Minuten geübt werden. Nach Ablauf dieser Zeit erscheint wieder eine Anzeige, jedoch können sich die Versuchspersonen auch darüber hinaus länger auf der Insel bewegen.

Explorationsdurchgang. In der ersten Phase der Water-Maze-Aufgabe erkunden die Versuchspersonen die Umgebung des VWM. Dazu beginnen sie am Rand der Insel im III. Quadranten mit Blick in Richtung des Zielquadranten. Alle vier Landmarken und drei proximale Hinweisreize (blau-weiße Bälle) sind in dieser Bedingung sichtbar (s. Abb. 6.9). Die Versuchspersonen werden instruiert, nach der versteckten Schatztruhe zu suchen. Dazu sollen sie, wenn möglich, den Bällen

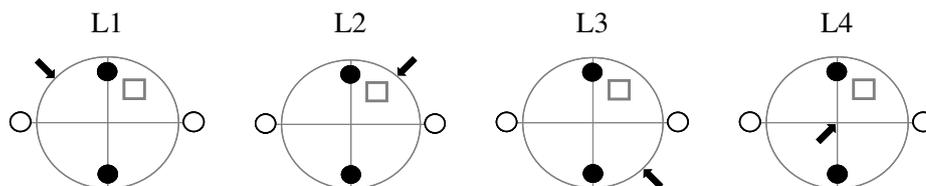
Abbildung 6.9: Explorationsdurchgang der Water-Maze-Aufgabe. (a) Grundriss. (b) Proximale Hinweise (Bälle) befinden sich auf dem direkten Weg vom Start zum Ziel.



folgen und sich die Umgebung und die Relationen und die Distanzen zu den einzelnen Objekten und der Objekte untereinander genau einprägen. Für die Erkundung der Insel steht beliebig Zeit zur Verfügung. Nachdem die Truhe auffindig gemacht wurde, können die Probanden sich innerhalb eines Radius von ca. 17 m um die Ziellokation eine beliebige Zeit frei bewegen. Der Explorationsdurchgang ist beendet, wenn dieser Bereich verlassen wird.

Lerndurchgänge. Nach der ersten Exploration wird in insgesamt vier Durchgängen das Wissen um die räumlichen Anordnungen gefestigt. Dazu starten die Versuchspersonen in jedem Durchgang von einer anderen Position mit Blickrichtung zur Mitte der Insel (s. Abb. 6.10). Die Versuchsperso-

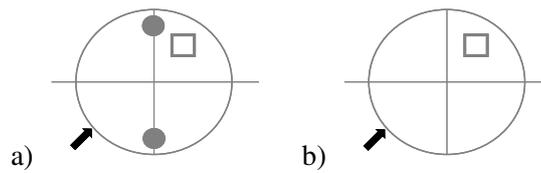
Abbildung 6.10: Lerndurchgänge L1 bis L4 der Water-Maze-Aufgabe.



nen werden instruiert, so schnell und zielgenau wie möglich zur versteckten Schatztruhe zu gehen. Ein Durchgang ist beendet, wenn die Versuchspersonen die Truhe geöffnet haben. Danach können sie sich in unmittelbarer Nähe der Truhe noch frei bewegen. Wird das Versteck der Truhe verlassen, wird die Versuchsperson an die neue Startposition teleportiert und ein weiterer Durchgang beginnt. Für jeden Lerndurchgang stehen maximal vier Minuten zur Verfügung.

Testdurchgänge. Direkt im Anschluss an die Lernphase folgen zwei Testbedingungen, in denen die Umgebung gezielt verändert wird, um das Verhalten des Probanden auf die Anwendung bestimmter Strategien zu testen. In der ersten Bedingung werden die beweglichen Landmarken Segelyacht und Boje im Osten bzw. Westen der Insel entfernt (s. Abb. 6.11 a). Damit kann geprüft werden, ob die gesamte Konstellation der Landmarken für die Orientierung berücksichtigt wird, oder ob nur die Relation Ziellokation-Leuchtturm gelernt wurde. Die Versuchspersonen beginnen ihre Suche in dieser Bedingung vom Rand der Insel aus dem III. Quadranten. Für diesen Test stehen maximal fünf Minuten zur Verfügung. Im zweite Test („Pfadintegrationsbedingung“) befinden sich keine Landmarken mehr auf der Insel (s. Abb. 6.11 b). Die Versuchspersonen werden instruiert, dass sie in diesem Test von der gleichen Position wie im vorangegangenen Durchgang starten.

Abbildung 6.11: (a) Testdurchgang 1: Löschen der beweglichen Landmarken. (a) Testdurchgang 2: Löschen aller Landmarken - Pfadintegrationsbedingung.



Zurückgelegte Distanzen und Richtungsänderungen können nur durch die von der Bewegung hervorgerufenen Veränderung der Textur des Untergrundes wahrgenommen werden. Für diesen Test wurde ebenfalls eine Zeitlimit von fünf Minuten festgelegt.

6.3.3 Merkmalsaggregation

Während der Navigation in den virtuellen Umwelten werden Rohdaten der Bewegung in einem Tupel mit fünf Attributen gespeichert. Ein Zeitstempel, die x- und y-Koordinaten der aktuellen Position, die kumulierte Distanz zum Ziel und die kumulierte Pfadlänge werden immer dann protokolliert, wenn das System eine Bewegung registriert. Aus diesen Daten kann der gesamte Bewegungsverlauf rekonstruiert und zur Gewinnung qualitativer und quantitativer Aussagen über die Orientierungsleistung verwendet werden.

Zur Verarbeitung und Visualisierung der Rohdaten wurde das eigene Tool „Dataviewer“ entwickelt. Implementiert wurde es in der Programmiersprache Python™ (Python Software Foundation). Es erlaubt die Anzeige eines gesamten Datensatzes einer Versuchsperson, die Selektion einzelner Trials und die Aggregation zahlreicher vordefinierter Merkmale aus den Bewegungsrohdaten. Eine genauere Beschreibung dieses Hilfsprogramms befinden sich im Anhang 15.

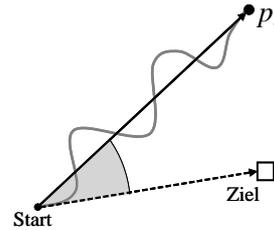
Es ist nicht möglich, eine pauschale Aussage darüber zutreffen, welche Merkmale die VWM-Leistung am besten charakterisieren. Vielmehr richtet sich die Auswahl der Merkmale nach der Fragestellung, die mit dem VWM untersucht werden soll, und dem konkreten Versuchsplan. Latenzzeiten und Pfadlängen sind jedoch die gebräuchlichsten und am einfachsten zu bestimmenden Maße. Eine vollständige Liste der im „Dataviewer“ implementierten Maße und deren Berechnungsmethoden befindet sich ebenfalls im Anhang 15.

Für die empirischen Untersuchungen der vorliegenden Arbeit wurden die in den VWM-Prozeduren beschriebenen Maße anderer Arbeitsgruppen berücksichtigt (Skelton et al., 2006; Moffat & Resnick, 2002; Hamilton et al., 2002; Sandstrom et al., 1998). Aus dem Training werden Latenzzeiten und Pfadlängen zwischen dem Öffnen zweier Kisten und die Summe dieser Werte berechnet. Darüber können mit ausreichender Genauigkeit Unterschiede im Umgang mit der VR-Technik und deren Verbesserung im Laufe des Trainings abgebildet werden.

Die Water-Maze-Aufgabe vereint mehrere Anforderungen an die Navigationsfähigkeit und das räumliche Gedächtnis. Als allgemeines Maß der VWM-Leistung dienen zunächst die Latenzzeiten und Pfadlängen bis zum Erreichen der Ziellokation aus jedem Durchgang. Spezielle Teilaspekte

Abbildung 6.12: Berechnung des initialen Heading-Errors.

Der initiale Heading-Error (grau) wird ermittelt als der Winkel zwischen den Peilungen vom Startpunkt zum Punkt p_i und zum Mittelpunkt der Ziellokation. Dabei ist p_i ein Punkt, der 20 Sekunden nach Initiierung der Bewegung erreicht wurde.



der Navigationsleistung sind die Berechnung einer optimalen Route zum Ziel und das Halten des optimalen Kurses. Ersteres kann durch den initialen Heading-Error abgebildet werden. Das ist der Winkel zwischen dem mittleren Richtungsvektor der ersten 20 Sekunden nach dem die Bewegung initiiert wurde und der direkten Peilung zum Ziel (s. Abb. 6.12). Die Verzögerung wurde so gewählt, da die Versuchspersonen zirka 20 Sekunden benötigen, um bei einem Start vom Rand eine Position auf der Insel zu erreichen, von der aus alle Landmarken zu sehen sind.

Zur Bewertung inwieweit ein optimaler Kurs während der Navigation eingehalten wurde, sind die Maße Umweg, durchschnittlicher Heading-Error und Anzahl der Umkehrpunkte implementiert wurden. Mit Umweg wird der Quotient aus zurückgelegter Pfadlänge und der kürzesten Distanz zwischen Start und Ziel (Luftlinie) bezeichnet. Der durchschnittliche Heading-Error ist ein zwischen allen diskreten Positionen auf dem Kurs akkumulierter und über die Anzahl aller Positionen gemittelter Richtungsfehler. Umkehrpunkte sind Positionen auf einem Kurs, in denen die Bewegungsrichtung sich um mehr als 90° ändert.

Die Komponente der räumlichen Gedächtnisleistung in der VWM-Aufgabe wird über die relative Verweildauer im Zielquadranten gemessen. Sie ist der Quotient aus der Navigationszeit im Zielquadranten und der Navigationszeit im gesamten Versuchsdurchgang.

7 Datenerhebung

7.1 Experiment 1 - Mentale Rotation

7.1.1 Stichprobe

Die Stichprobe im Experiment 1 bestand aus insgesamt 262 Versuchspersonen und setzte sich, wie in Tabelle 7.1 dargestellt, aus Schülern, Studenten und älteren Erwachsenen zusammensetzten. Die ca. 200 Schüler waren zum Zeitpunkt der Testung Schüler des fünften bis achten Schuljahrs (mittleres Alter=12,5; SD=1,17) an einem Gymnasium mit mathematisch-naturwissenschaftlicher Spezialisierung. Die Stichprobe der Studenten setzte sich aus Studenten des ersten bis dritten Semesters (mittleres Alter=22,6; SD=3,68) der Studiengänge Psychologie, Natur-, Rechts- und Wirtschaftswissenschaften zusammen. Die älteren Erwachsenen, Senioren mit einem mittleren Alter von 66,7 Jahren (SD=7,15), konnten unter den Mitgliedern eines Universitätssportvereins bzw. den Teilnehmern eines PC-Kurses der Volkssolidarität für die Studie gewonnen werden.

Tabelle 7.1: Stichprobe Experiment 1.

	N	N weiblich	N männlich
Gesamt	262 (100,0%)	90 (34,4%)	172 (65,6%)
Schüler	206 (78,6%)	60 (22,9%)	146 (55,7%)
Studenten	36 (13,7%)	18 (6,9%)	18 (6,9%)
Senioren	20 (7,6%)	12 (4,6%)	8 (3,1%)

Die Teilnahme am Experiment war für alle Versuchspersonen freiwillig und erfolgte ohne Vergütung. Die Untersuchung am Gymnasium wurde mit der Genehmigung des Schulamtes¹ und unter der Voraussetzung einer vorliegenden schriftliche Einwilligungserklärung der Eltern durchgeführt. Alle Daten wurden anonymisiert unter Verwendung eines Codeworts erhoben. Auf Wunsch konnten individuelle Testergebnisse von den Versuchsteilnehmern eingesehen werden.

Alle Versuchspersonen wurden bezüglich neurologischer und psychischer Erkrankungen, Schädel-Hirnverletzungen, der Einnahme psychoaktiver Medikamente und Substanzabhängigkeiten zum aktuellen Zeitpunkt oder in ihrer Vergangenheit befragt. Datensätze von Personen mit entsprechenden Auffälligkeiten wurden von der statistischen Aufwertung ausgeschlossen.

¹Reg.-Nr. 504-35/06, Landesverwaltungsamt Sachsen-Anhalt, Referat Gymnasien und Gesamtschulen, Magdeburg

7.1.2 Messinstrumente

7.1.2.1 Raumkognitive Maße

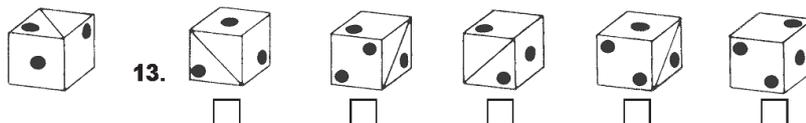
Mental Rotations Test. Drei von insgesamt vier der Operationalisierungen der Fähigkeit der mentalen Rotation sind mit dem MRT (Peters et al., 1995) durchgeführt worden. Es sind die, wie unter Abschnitt 6.1 beschrieben, Papier-Bleistift-Versionen der den Testform A (MRT-A) und B (MRT-B) und die Computerversion der Testform B (CMRT-B) eingesetzt worden. Alle Tests mit dem MRT wurden in einem zweiteiligen Speedtest durchgeführt. Das Zeitlimit betrug für Schüler und Studenten drei Minuten pro Testhälfte, für die Senioren wurde es auf fünf Minuten pro Testhälfte erhöht.

Zur Bewertung der Testleistung wurden Summenscores nach dem strengen Punktesystem aus allen Items und den beiden Testhälften gebildet und die Anzahl bearbeiteter Items ermittelt. Weiter wurden die dichotomisierten Antwortmuster aller 24 Items erfasst, d. h. ein vollständig richtig gelöstes Item wurde mit eins bewertet, ein falsch gelöstes bzw. ein nicht bearbeitetes Item mit null.

Würfelaufgaben. Mit dem Untertest *Würfelaufgaben* aus dem Intelligenz-Struktur-Test (I-S-T 2000, Amthauer et al., 1999) wurde ein weiterer Test zur mentalen Rotation aufgenommen. Innerhalb des I-S-T operationalisiert der Untertest räumlich-figurale Fähigkeiten und wird der Strukturkomponente figurale Intelligenz zugeordnet. Es handelt sich bei dem im Test verwendeten Material um perspektivische Abbildungen von Würfeln mit unterschiedlichen Symbolen auf den Seitenflächen. In der Abbildung eines Würfels sind so jeweils drei Symbole zu erkennen (s. Abb. 7.1).

Die Aufgabenstellung ist ähnlich dem MRT: Zu einem Zielwürfel soll aus fünf Alternativen der dazu passende Würfel ausgewählt werden. Es ist aber jeweils nur ein Würfel richtige. Der Lösungswürfel wurde dabei gegenüber dem Zielwürfel gedreht oder gekippt oder gedreht und gekippt; die vier anderen Abbildungen zeigen anders zusammengesetzte Würfel.

Abbildung 7.1: Testitem des Untertests *Würfelaufgaben* aus dem I-S-T 2000.



Zur Beantwortung eines Items, in der hier verwendeten leicht modifizierten Form des Tests, musste das Kästchen unterhalb des ausgewählten Würfels angekreuzt werden. Dieser Bearbeitungsmodus wurde gegenüber den ursprünglichen Instruktionen im I-S-T abgeändert, um ihn an den des MRT anzupassen. Der gesamte Test bestand aus 25 solcher Items. Nach einer Erläuterung des Testmaterials lösten die Versuchspersonen zunächst fünf der Items als Übungsaufgabe ohne Zeitbegrenzung. Die richtigen Lösungen der Übungsaufgaben waren im Testheft zur Überprüfung der Antworten angegeben. Die restlichen 20 Items waren dann als Speedtest mit einem Zeitlimit von neun Minuten zu bearbeiten. Die Testleistung wurde über den Summenscore und die Anzahl bearbeiteter Items

bestimmt. Dichotomisierte Antwortmuster aller 20 Testitems wurden für eine Analyse der Bearbeitungsstrategien ebenfalls erfasst.

Tabelle 7.2: Itemtypen des I-S-T Untertests *Würfelaufgaben*.

Flächenwürfel	Flächenwürfel-Plus	Raumwürfel
1, 3, 4, 5, 11, 12, 14, 15, 16	2, 7, 8, 13	6, 9, 10, 17, 18, 19, 20

Auch im Test *Würfelaufgaben* gibt es verschiedene Itemtypen, die die Anwendung unterschiedlicher Bearbeitungsstrategien erlauben. Anhand struktureller Merkmale können diese ebenfalls klassifiziert werden. Die 20 Testitems lassen sich nach der Bezeichnung von Putz-Osterloh (1977) in „Flächenwürfel“, „Flächenwürfel-Plus“ und „Raumwürfel“ aufteilen. Tabelle 7.2 zeigt die Einteilung der einzelnen Items aus dem verwendeten Testmaterial in diese drei Typenklassen.

Analysiert man die Struktur der Items, fällt auf, dass für die Kennzeichnung der Seitenflächen der Würfel nur vier Symbole verwendet wurden. Daraus ergeben sich bei deren Abbildung auf drei Seitenflächen vier mögliche Kombinationen (d. h. ohne Beachtung der räumlichen Anordnung). Während bei Items mit „Flächenwürfel“ der Lösungswürfel die gleiche Kombination wie der Zielwürfel aufweist, und somit durch einen Detailvergleich der Symbole gelöst werden kann, führt diese Bearbeitungsstrategie bei einem „Raumwürfel“ zu einem Fehler. Bei diesem Itemtyp unterscheiden sich die Abbildungen des Ziel- und des Lösungswürfels jeweils in einem Symbol und sind nur durch mentales Rotieren miteinander zu vergleichen.

Die Lösung eines „Flächenwürfel-Plus“ kann wiederum mit mehreren Detailvergleichen erfolgen. Bei diesen Items stimmt die Kombination der Symbole auf dem Zielwürfel mit jeweils zwei Alternativwürfeln überein. Sind diese beiden identifiziert, müssen sie nur noch in Distraktor und Lösung differenzieren werden. Dazu muss nun die räumliche Anordnung der Symbole auf den Würfeln berücksichtigt werden. Das kann durch mentales Rotieren, aber auch durch einen weiteren Detailvergleiche erfolgen.

7.1.2.2 Fragebögen

Neben den demografischen Daten Alter, Geschlecht, Händigkeit, Muttersprache wurde bei Schülern die letzte Mathematiknote und bei Erwachsenen der Bildungsstand über Schulabschluss und Berufsausbildung erfasst. Weiterhin wurden allen Versuchspersonen zu Vorerfahrungen mit Raumvorstellungstests befragt. Dieses Merkmal wurde dreifach gestuft in „keine Erfahrung“, „allgemeine Erfahrungen mit Raumvorstellungstest“ und „spezielle Erfahrungen mit dem MRT“.

Computererfahrung. Erfahrungen im Umgang mit Computern und Computerspielen wurden ebenfalls anhand eines Fragebogens erfasst. Im Fragebogen wurde die täglich verwendete Zeit für die Nutzung von Anwendungsprogrammen und Computerspielen ermittelt. Diese wurden in den

sechs Abstufungen „nie“, „nicht täglich“, „täglich weniger als eine Stunde“, „täglich mehr als eine Stunde“, „täglich mehr als zwei Stunden“ oder „täglich mehr als drei Stunden“ erfasst.

Räumliche Strategien. Neben der automatischen (impliziten) Analysen der Antwortmuster wurde zur Beurteilung der im MRT verwendeten Lösungsstrategien auch eine explizite Methode eingesetzt. Dazu wurden 24 dichotome Items in einem Fragebogen zusammengestellt, die jeweils eine mögliche Strategie betrafen. 22 der Fragen richteten sich an die verwendeten Lösungsstrategien und erfassten vier Komplexe: mentale Repräsentation und Manipulation der 3D-Figuren, Systematik des Vorgehens innerhalb eines Items, Flexibilität und Genauigkeit beim Einsatz der jeweiligen Strategien. Im Anhang 14 ist dieser Fragebogen abgebildet.

7.1.2.3 Allgemeine kognitive Maße

Prüfsystem für Schul- und Bildungsberatung. Mit vier Untertests aus dem PSB-R (Horn, Lukesch, Mayrhofer & Kormann, 2003) wurden allgemeine Fähigkeiten zum schlussfolgernden Denken (Reasoning) erfasst, die auch bei der Bearbeitung von Raumvorstellungstests von Bedeutung sind. Im PSB-R werden dazu die vier materialspezifischen Aufgabengruppen *Reihen mit Buchstaben* (AG 2), *Reihen mit Zahlen* (AG 3), *Reihen mit Figuren* (AG 4) und *Zählen von Flächen* (AG 6) verwendet (s. Abb. 7.2). In den Aufgaben *Reihen* sind Folgen vorgegeben, die nach bestimmten Regeln aufgebaut sind. Ein Eintrag in jeder Reihe verletzt die jeweilige Regel und ist durchzustreichen. Jeder Untertest aus dieser Aufgabengruppe besteht aus jeweils 18 Items. Die Buchstabenreihen mussten innerhalb von fünf Minuten, die Zahlenreihen in vier und die Figurenreihen in drei Minuten bearbeitet werden. Der Summenscore aller Reihenaufgaben wird im PSB-R zu einem Maß der „induktiven Denkfähigkeit“ zusammengefasst.

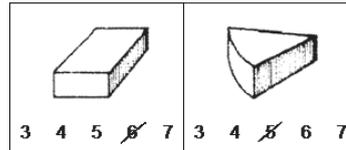
Abbildung 7.2: Beispielaufgaben aus dem PSB-R 6-13. Aufgabengruppen *Reihen* (AG 2, 3 und 4) mit unterschiedlichen Modalitäten.

2	2	2	2	2	2	2	2	2
a	b	a	b	a	b	a	a	a
	+		+		+		 	

Durch die Aufgabengruppe 6 wird im PSB-R die Komponente Raumvorstellung erfasst. Diese Aufgabengruppe enthält 42 Items, die in drei Minuten zu bearbeiten waren. Ein Item zeigt eine perspektivische Abbildung eines geometrischen Körpers mit sichtbaren und verdeckten Flächen. Die Versuchspersonen sollen alle Flächen des Körpers zählen und durch Streichen der jeweiligen Zahl ihre Antwort geben (s. Abb. 7.3). Jedes richtig beantwortete Item wurde mit einem Punkt bewertet. Der Summenscore bildet zusammen mit der „induktiven Denkfähigkeit“ im PSB-R den Reasoning-Faktor (R-Faktor).

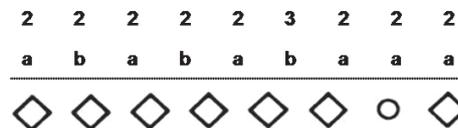
Der PSB-R wurde auf Grund der vorliegenden Normierung für Schüler und junge Erwachsenen nur zur Befragung der Schüler und Studenten eingesetzt.

Abbildung 7.3: Beispielaufgabe aus dem PSB-R. Aufgabengruppe *Raumvorstellung* (AG 6)



Leistungsprüfsystem 50+. Zur Erfassung vergleichbarer kognitiver Maße in der Stichprobe der Senioren wurden drei Untertests aus dem LPS 50+ (Sturm, Willmes & Horn, 1993) verwendet. Die Untertests 3 und 4 operationalisieren schlussfolgerndes Denken ebenfalls über das Erkennen von Regelverletzungen innerhalb von Zahlen-, Buchstaben- und Figurenreihen (s. Abb. 7.4). Jeder Untertest besteht aus 42 Items. Der Bearbeitungsmodus ist mit denen des PSB-R identisch. Die Bearbeitungszeit für beide Untertests betrug fünf bzw. acht Minuten.

Abbildung 7.4: Beispielaufgaben *Reihen* aus dem LPS 50+.



Untertest 9 aus dem LPS 50+ ist identisch zu den Aufgaben *Flächenzählen* aus dem PSB-R. Die Anzahl der Items beträgt ebenfalls 42, die in drei Minuten zu bearbeiten waren.

7.1.3 Technische Apparatur

Zur Durchführung des MRT in der Computerversion sind mehrere, aber untereinander vergleichbare PC-Systeme eingesetzt worden. Es handelte sich dabei um PCs aus Computerpools der Universität oder Schule, an denen die Testungen stattfanden. Alle Computer verfügten über handelsübliche CRT-Monitore mit 15 Zoll Bildschirmdiagonale in einer Auflösung von 1024 × 786 Pixeln mit 32-Bit Farbtiefe. Auf allen PCs waren das Microsoft® Windows® XP -Betriebssystem (SP 2), der OpenGL®-kompatible VR-Player World UP® (Sense8, Inc.) und die Computerversion des MRT installiert. Die Rechenleistung aller PCs war für die Durchführung des Tests ausreichend und spielte praktisch keine Rolle, da die 3D-Objekte der Shepard-Metzler-Figuren in dieser Version nur statisch dargestellt wurden.

7.1.4 Durchführung

Das Experiment wurde in einem Zweigruppen-Prä-Post-Design mit Parallelisierung auf der Gruppenebene durchgeführt. Im Prätest wurden demografische Daten, der Fragebogen zur Computererfahrung, die beschriebenen Untertests aus dem PSB-R bzw. LPS 50+ und dem MRT-A in der Papier-Bleistift-Version erhoben. Parallelisiert wurde nach Alter, Geschlecht, MRT-Leistung (Summenscore) und Computererfahrung. Im Posttest wurden der MRT-B, der Fragebogen zu den Lö-

sungsstrategien und der Test *Würfelauflagen* erhoben. Eine Gruppe absolvierte den MRT in der Papier-Bleistift-Version und die andere führte den Computertest durch.

Zwischen Prä- und Posttest lagen vier Wochen. Alle Testungen wurden als Gruppentest durchgeführt. Die Schüler wurden in ihren Schulklassen während des regulären Unterrichts getestet. Die Klassenstärken betrug zwischen 15 und 27 Schülern pro Klasse. Studenten und Senioren nahmen an vereinbarten Terminen in Gruppen zwischen vier bis 18 Personen am Experiment teil.

7.2 Experiment 2 - Pfadintegration

7.2.1 Stichprobe

Die Stichprobe zum Experiment 2 wurde als Datenpool aus vier einzelnen Studien, die die virtuelle Kompass-Zeigeaufgabe (VPT) einsetzten, gebildet. Drei Altersgruppen wurden *post hoc* nach folgenden Kriterien festgelegt: jüngere Erwachsene im Alter zwischen 20 bis 29 Jahren, Erwachsene im Alters von 38 bis 43 Jahren und ältere Erwachsene zwischen 50 und 59 Jahren.

Für die Gruppe der jüngeren Erwachsenen ergab sich ein mittleres Alter von 24,9 Jahren (SD=2,4 Jahre), für die Erwachsenen mittleren Alters von 40,2 Jahren (SD=1,6 Jahre) und für die Gruppe der älteren Erwachsenen von 53,9 Jahren (SD=2,19 Jahre). Frauen hatten ein mittleres Alter von 36,6 Jahren (SD=13,0 Jahre) und Männer ein mittleres Alter von 37,6 Jahren (SD=13,8 Jahre). Tabelle 7.3 zeigt den Anteil von Frauen und Männern in den drei Altersgruppen.

Tabelle 7.3: Stichprobe Experiment 2.

	N	N weiblich	N männlich
Gesamt	200 (100%)	88 (44%)	112 (56%)
jüngere Erwachsene	103 (51,5%)	45 (22,5%)	58 (29,0%)
mittlere Erwachsene	27 (13,5%)	16 (8,0%)	11 (5,5%)
ältere Erwachsene	70 (35,0%)	27 (13,5%)	43 (21,5%)

Die jungen Erwachsenen waren Studenten verschiedener Studiengänge der Universität Halle. Die älteren und mittleren Erwachsenen waren Familienangehörige oder Bekannte von nicht am Experiment teilnehmenden Studenten. Alle Versuchspersonen wurden bezüglich neurologischer und psychischer Erkrankungen, Schädel-Hirnverletzungen, der Einnahme psychoaktiver Medikamente und Substanzabhängigkeiten zum aktuellen Zeitpunkt oder in ihrer Vergangenheit befragt. Datensätze von Personen mit entsprechenden Auffälligkeiten wurden von der weiteren statistischen Auswertung ausgeschlossen.

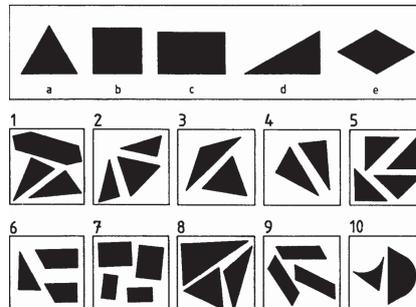
7.2.2 Messinstrumente

7.2.2.1 Raumkognitive Maße

Virtuelle Kompass-Zeigeaufgabe. Die Erhebung empirischer Daten zu Pfadintegration Umgebungen mit Hilfe der im Abschnitt 6.2 beschriebenen Methode steht im Mittelpunkt des Experiment 2. Die Versuchspersonen absolvierten die virtuelle Kompass-Zeigeaufgabe, wie sie unter Abschnitt 6.2 beschrieben wurde. Für die statistische Auswertung wurden die absoluten Schätzfehler aus jedem Durchgängen und der Mittelwert der Antwortlatenz über alle Durchgänge verwendet.

Figurenauswahl. Die Fähigkeit zur mentalen Rotation wurde mit dem Untertest *Figurenauswahl* aus dem Intelligenz-Struktur-Test (I-S-T 2000, Amthauer et al., 1999) erfasst. Der Test besteht aus insgesamt 20 Items, die in Form von zwei Aufgabenblöcken präsentiert werden. Abbildung 7.5 zeigt einen Aufgabenblock (Item 1 bis 10) aus dem Test.

Abbildung 7.5: Aufgabenblock aus dem Untertest *Figurenauswahl* des I-S-T 2000.



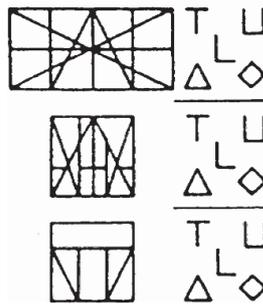
Es sind fünf Musterfiguren und zehn Abbildungen von Figuren, die in mehrere Teile zerschnitten wurden, vorgegeben. Die Versuchspersonen müssen herausfinden, welche der fünf Musterfiguren (a-e) aus den einzelnen Stücken zusammengesetzt werden kann, ohne dass Ecken überstehen oder Platz dazwischen bleibt. Diese Aufgabe ist nur durch mentale Rotation und Verschiebung der Einzelteile zu lösen. Die Bearbeitungszeit für den gesamten Test betrug sieben Minuten. Die Summe aller richtig gelösten Items wurde als Leistungsmaß verwendet.

Leistungsprüfsystem. Aus dem LPS (Horn, 1983) ist der Untertest 10 (LPS10) für die Untersuchung entnommen worden. Dieser wird im LPS der Facette Form- und Gestalterfassung zugeordnet und dient zusammen mit weiteren Untertests der Ermittlung technischer Begabungen. Die Aufgaben des LPS10 wurden ausgewählt, da sie auch der Beurteilung von Fähigkeiten einer Person dienen, inwieweit diese in der Lage ist aus einem komplexen geometrischen Muster Zahlen oder Buchstaben herauszulösen. Im weitesten Sinne werden also auch Fähigkeiten der räumlichen Wahrnehmung operationalisiert.

Dieser Test kann aber auch mit dem von Witkin et al. (1954) aufgestellten Konstrukt der Feldabhängigkeit in Verbindung gebracht werden. So geht es nach Horn (1983) um das „Erkennen des

ht

Abbildung 7.6: Aufgabenblock aus dem Untertest 10 des LPS.



Wesentlichen trotz ablenkender Einzelheiten“. Eine Fähigkeit, die Witkin et al. (1954) als Entscheidungsfindung unabhängig von Einflüssen aus dem umgebenden Feld beurteilten. In Abbildung 7.6 ist eines der Items des LPS10 dargestellt. Insgesamt waren 40 Items zu beantworten, die zu einem Summenscore zusammen gefasst wurden. Die Bearbeitungszeit war auf vier Minuten begrenzt.

Rey-Oesterrieth-Complexfigur. Visuell-räumliche Gedächtnisleistungen für zwei-dimensionale Abbildungen wurden mit dem *Rey-Oesterrieth Complex Figur Test* erfasst (RCF; Rey, 1941; Osterrieth, 1944). Zur Erfassung visuo-konstruktiver Fähigkeiten zeichneten die Versuchspersonen zunächst die Figur ab (RCF-copy). Nach 30 Minuten sollte die Figur aus dem visuell-räumlichen Gedächtnis spontan erinnert werden (RCF-recall). Die Leistung wurde über ein Rating der korrekt wiedergegebener Anteile der Zeichnung nach dem Scoringssystem von Lezak (1995) bewertet.

7.2.2.2 Allgemeine kognitive Maße

Mehrfachwahl-Wortschatztest. Zur Erfassung des allgemeinen Intelligenzniveaus wurde der Mehrfachwahl-Wortschatztest in der Form B (MWT-B) nach Lehrl, Triebig und Fischer (1995) eingesetzt. Es handelt sich bei diesem Test um einen Worterkennungstest der primär verbale Fähigkeiten erfasst. Ein Testitem besteht jeweils aus fünf Wörtern, wobei nur eines der Wörter sinnvoll ist. Die vier anderen Wörter sind Nicht-Wörter. Die Aufgabe besteht darin, das sinnvolle Wort zu erkennen. Der Schwierigkeitsgrad zwischen sinnvollem Wort und Nicht-Wörtern zu differenzieren steigt mit jedem weiteren Item. Insgesamt sind 37 Items im Test vorgegeben, die ohne Zeitlimit bearbeitet wurden. Für die statistische Auswertung wurde ein Summenscore aus allen richtig beantworteten Items gebildet. Dieser gilt darüber hinaus als zuverlässige Schätzung der kristallinen Intelligenz.

Trail Making Test. Dieser Test besteht aus den beiden Teilen A und B. Im Teil A (TMT-A) müssen die Versuchspersonen die Zahlen 1 bis 24, die in einer willkürlichen Ordnung auf einem Blatt Papier verteilt sind, in aufsteigender Reihenfolge miteinander verbinden. Im Teil B (TMT-B) müssen zwölf Zahlen und zwölf Buchstaben in alternierender Reihenfolge (1-A, 2-B, etc.) miteinander verbunden werden. Als Testscore in beiden Teilen dienen die Bearbeitungszeiten in Sekunden.

Auf diese Weise erfasst der TMT-A visuo-motorische Leistungsfähigkeit und der TMT-B kognitive Flexibilität (Boll & Reitan, 1973).

Befindlichkeitsskala. Das momentane subjektive Befinden wurde über die Befindlichkeitsskala (Bf-S) von Zerssen (1976) erfasst. Sie ist eine Eigenschaftswörterliste, bestehend aus jeweils 28 Gegensatzpaaren (z. B.: frisch - matt), für die die Versuchspersonen den Pol angeben sollten, der ihrem augenblicklichen Zustand am ehesten entspricht. Außerdem kann eine neutrale Auswahl getroffen werden. Dem positiven Pol wurden null Punkte, dem negativem Pol zwei Punkte und der neutralen Antwort ein Punkt zugeordnet. Die Summe der Punktwerte bildete den Testscore.

7.2.3 Technische Apparatur

Die Daten aus der VPT wurden mit drei Computersystemen erhoben. 110 Versuchspersonen, das entspricht 55 % der Gesamtstichprobe, wurden mit einem Notebook (3.0 GHz Pentium-IV®, nVIDIA® GeForce™ Go 5700) der Firma Toshiba getestet, das über ein 17 Zoll Breitbild TFT-Farbdisplay (Bildschirmauflösung 1280 × 720 Pixeln) verfügte. Die Bildschirmgröße entsprach hier einem FoV von 44° × 28°.

66 Versuchspersonen (33% der Stichprobe) absolvierten die Aufgabe an einem Notebook vergleichbarer Rechenleistung mit einem 15 TFT-Display (Bildschirmauflösung: 1024×768 Pixel). Die verbleibenden 24 Versuchspersonen (12% der Stichprobe) benutzten einen PC mit einem Bildschirm gleicher Größe. Die Bildschirmgröße entsprach hier einem FoV von 34° × 26°.

Auf allen Computersystemen waren das Microsoft® Windows® XP-Betriebssystem (SP 2) und der OpenGL®-kompatible VR-Player World UP® (Sense8, Inc.) installiert.

7.2.4 Durchführung

Die Testsessions in den Einzelstudien wurden, wenn auch mit spezifischen Modifikationen, nach einem vergleichbaren Ablaufplan durchgeführt. Alle Versuchspersonen wurden zunächst in schriftlicher und mündlicher Form über Inhalt und Verlauf des Gesamtexperimentes informiert und ihre demografische Daten wurden erfasst. Zuerst absolvierten die Versuchspersonen immer zwei VR-gestützte Experimente. Die virtuelle Zeigeaufgabe war jeweils einer dieser Tests. Nach den VR-gestützten Tests wurden alle weiteren Papier-Bleistift-Tests bearbeitet.

7.3 Experiment 3 - Place-Learning

7.3.1 Stichprobe

Im Experiment 3 wurden Daten von jungen Erwachsenen (Alter: M=24.2 Jahre, SD=3.5 Jahre) und älteren Erwachsenen (Alter: M=53.5 Jahre, SD=2.2 Jahre) erhoben. Frauen hatten ein mittleres

Alter von 39.0 Jahren (SD=15.0 Jahre) und Männer ein mittleres Alter von 39.8 Jahren (SD=15.1 Jahre). Tabelle 7.4 zeigt den Anteil von Frauen und Männern in den beiden Altersgruppen.

Tabelle 7.4: Stichprobe Experiment 3.

	N	N weiblich	N männlich
Gesamt	58 (100%)	29 (50%)	29 (50%)
jüngere Erwachsene	28 (48%)	14 (24%)	14 (24%)
ältere Erwachsene	30 (52%)	15 (26%)	15 (26%)

Die jungen Erwachsenen waren Studenten des ersten und zweiten Semesters des Studiengangs Psychologie der Universität Halle. Die älteren Erwachsenen waren Familienangehörige von nicht am Experiment teilnehmenden Studenten. Alle Versuchspersonen wurden bezüglich neurologischer und psychischer Erkrankungen, Schädel-Hirnverletzungen, der Einnahme psychoaktiver Medikamente und Substanzabhängigkeiten zum aktuellen Zeitpunkt oder in ihrer Vergangenheit befragt. Personen mit entsprechenden Auffälligkeiten wurden vom Experiment ausgeschlossen.

7.3.2 Messinstrumente

7.3.2.1 Raumkognitive Maße

Virtuelles Water-Maze. Zur Beurteilung der Fähigkeit Place-Learning wurde die Water-Maze-Aufgabe eingesetzt, wie sie unter Abschnitt 6.3 beschrieben wurde. Alle Versuchspersonen absolvierten die dort beschriebenen Versuchsbedingungen Training, Explorationsdurchgang, Lern- und Testdurchgänge. Gemessen wurden während des Trainings Pfadlängen und Navigationszeiten bis zum Öffnen jeder der vorhandenen Kisten. Merkmale zur Charakterisierung des Explorationsverhaltens waren die Pfadlänge, die durchschnittliche Geschwindigkeit und die Anzahl der Umkehrpunkte.

In den Lerndurchgängen wurde erfasst, ob ein Durchgang innerhalb des Zeitlimits erfolgreich beendet wurde. Daraus wurde ein Summenscore der Anzahl der erfolgreichen Lerndurchgänge gebildet. Weiter wurde aus den Verhaltensdaten der vier Lerndurchgänge folgende Merkmale zur weiteren statistischen Verarbeitung aggregiert: der gelaufene Umweg, der initiale Heading-Error und die durchschnittliche Geschwindigkeit. Daraus wurden Mittelwerte über alle Lerndurchgänge gebildet.

Aus den beiden Testbedingungen wurden die relative Verweildauer im Zielquadrant und der initiale Heading-Error berechnet. Eine dichotome Variable wurde erhoben, aus der hervorgeht, ob das Ziel gefunden wurde.

Virtuelle Kompass-Zeigeaufgabe. Im Experiment 3 wurde ebenfalls die Zeigeaufgabe in der virtuellen Parklandschaft eingesetzt. Die Versuchspersonen absolvierten die sechs Durchgänge der

Kompass-Zeigeaufgabe, wie sie im Abschnitt 6.2 beschrieben wurden. Für die statistische Auswertung wurden absolute Schätzfehler aus jedem Durchgang und der Mittelwert der Reaktionszeiten über alle Durchgänge verwendet.

Mental Rotations Test. Mentale Rotation wurde mit dem *Mental Rotations Test* in der Testform A (MRT, Peters et al., 1995) erfasst. Der MRT wurde als zweiteiliger Speedtest mit einem Zeitlimit von drei Minuten pro Testhälfte durchgeführt. Der Summenscore über alle Items und über Items der verschiedenen Typen diente als Maß zur Bewertung der Testleistung. Die dichotome Antwortmuster wurden für die qualitativen Analyse ebenfalls protokolliert.

Figurenauswahl und Spiegelbilder. Zweidimensionale mentale Rotation wurde mit dem Untertest *Spiegelbilder* aus dem Wilde-Intelligenz-Test (WIT, Jäger & Althoff, 1994) und dem Untertest *Figurenauswahl* (I-S-T 2000; Amthauer et al., 1999) erhoben.

Der Test *Figurenauswahl* wurde bereits im Experiment 2 eingesetzt und im Abschnitt 7.2.2.1 beschrieben. Die Bearbeitungszeit für den gesamten Test betrug auch hier sieben Minuten. Der Summenscore über alle richtige gelösten Items diente als Leistungsmaß.

Abbildung 7.7: Aufgaben *Spiegelbilder* aus dem Wilde-Intelligenztest.



Der Test *Spiegelbilder* besteht aus insgesamt 26 Items, in denen jeweils fünf Abbildungen von Strichmännchen dargestellt sind. Vier dieser Abbildungen zeigen das gleiche, nur in der Bildebene rotierte, Strichmännchen. Eine Abbildung zeigt ein Strichmännchen, das gespiegelt wurde (s. Abb. 7.7). Die Aufgabe der Versuchspersonen bestand darin, das gespiegelte Strichmännchen zu kennzeichnen. Für die Bearbeitung des Tests standen insgesamt fünf Minuten zur Verfügung. Als Leistungsmaß wurde der Summenscore über alle richtig beantworteten Items gebildet.

Leistungsprüfsystem. Für die Beurteilung der Fähigkeit zur Wahrnehmung räumlicher Muster diente, wie in Experiment 2, der Summenscore aus dem Untertest 10 des LPS (LPS10; Horn, 1983 und Abschn. 7.2.2.2). Die Bearbeitungszeit für diesen Test betrug vier Minuten.

Rey-Oesterrieth-Complexfigur. Es wurde zur Erfassung visuo-konstruktiver Fähigkeiten und visuell-räumlicher Gedächtnisleistungen für zwei-dimensionale Abbildungen, wie in Experiment 2, der *Rey-Oesterrieth Complex Figures Test* (Rey, 1941; Osterrieth, 1944) in den Bedingungen „Copy“ und „Recall“ mit dem Scoringssystem von Lezak (1995) eingesetzt (s. Abschn. 7.2.2.2).

7.3.2.2 Fragebögen

Selbsteinschätzungen der Fähigkeiten zur räumlichen Orientierung, der Vorerfahrung mit Raumvorstellungstest und dem Umgang mit Computer wurden durch eigene Fragebögen erfasst. Anhand von elf Items (dichotome, Rating- und freie Antworten) wurde die Vorerfahrung im Umgang mit Computern und den beschriebenen Testverfahren erhoben. Ein Item erfasste die Fähigkeit, sich in neuen Umgebungen zurechtzufinden; ein weiteres Item das selbsteingeschätzte Landkartenverständnis. Die Items 1 und 2 erfassen die Häufigkeit des Umgangs mit Computern bzw. Computerspielen und werden von null bis vier codiert. Mit den dichotomen Items 4, 5, 6 und 9 werden diese zu einem Summenscore „Vorerfahrung“ zusammengefasst.

Die von den Versuchspersonen in der Water-Maze-Aufgabe verwendeten Strategien (Place-Learning) wurden in einem weiteren Fragebogen mit insgesamt 13 Items erfasst. Aus insgesamt sieben Items bestand ein weiterer Strategiefragebogen für die Kompass-Zeigeaufgabe. Alle Fragebögen sind im Anhang 14 einzusehen. Die Fragebögen wurden auf Itemebenen deskriptiv und korrelativ ausgewertet. Sie dienten zur Kreuzvalidierung der explizit angegebenen Lösungsstrategien mit denen aus einer automatischen Analyse gewonnen Strategien.

7.3.2.3 Allgemeine kognitive Maße

Es wurden die bereits im Experiment 2 (s. Abschn. 7.2.2.2) beschriebenen psychometrischen Verfahren des Mehrfach-Wahl-Wortschatztest (MWT-B; Lehrl et al., 1995) zur Schätzung der kristallinen Intelligenz, der Trail Making Test (TMT-A/B; Boll & Reitan, 1973) zur Messung der visuo-motorische Leistungsfähigkeit und kognitive Flexibilität und die Befindlichkeitsskala (Bf-S; Zerssen, 1976) zur Erfassung des momentanen subjektiven Befindens eingesetzt.

7.3.3 Technische Apparatur

Die VPT und das VWM innerhalb des Experiments 3 waren zwei computergestützte Aufgaben. Zur Präsentation wurde ein Toshiba Satellite-Notebook mit 3.0 GHz Intel® Pentium® IV-Prozessor und nVIDIA® GeForce™ FX™ Go 5700 (64MB VRAM)-Grafikkarte (Toshiba Corp.) eingesetzt. Das Notebook verfügte über ein Breitbild TFT-Farbdisplay mit 17 Zoll Bildschirmdiagonale, auf dem die virtuellen Umwelten in einer Auflösung von 1280 × 720 Pixeln mit 32-Bit Farbtiefe dargestellt wurden. Die Bildschirmgröße entspricht bei einem ungefähren Auge-Display-Abstand von 45 cm einen FoV von 44° × 28°. Auf dem Notebook war Microsoft® Windows® XP -Betriebsystem (SP 2) und der OpenGL®-kompatible VR-Player World UP® (Sense8, Inc.) installiert.

Eingaben zur Bewegungskontrolle während der Place-Learning-Aufgabe wurden durch ein Gamepad mit D-Pad² vorgenommen. Das D-Pad hat die Form eines Kreuzes und erlaubt diskrete Bewegungen auf zwei Achsen: Translation (Vorwärts-Rückwärts) und Rotation (Links-Rechts). Das

²Directional Pad: Digitales 4-Richtungen-Steuergeräte für Videospiele.

Gamepad wird dazu in beiden Hände gehalten und die Bedienung des D-Pads erfolgt mit dem linken Daumen. Eine Regulierung der Geschwindigkeit, wie bei einem analogen Joystick, ist nicht möglich. Dafür ist die Bewegungssteuerung sehr viel einfacher und akkurater. In der Kompass-
Zeigeaufgabe wurden alle Eingaben mit der Maus ausgeführt.

7.3.4 Durchführung

Alle Versuchspersonen wurden zunächst in schriftlicher und mündlicher Form über Inhalt und Verlauf des Gesamtexperimentes informiert. Danach absolvierte die Hälfte der Versuchspersonen zuerst die VPT und danach das virtuelle Training gefolgt von der Water-Maze-Aufgabe. Die andere Hälfte der Versuchspersonen begann mit dem Training und dem Water-Maze. Direkt nach Abschluss des VPT bzw. des VWM wurden die Strategiefragebögen zur Pfadintegration bzw. zum Place-Learning vorgelegt. Nach den VR-gestützten Tests wurden alle weiteren Papier-Bleistift-Tests, beginnend mit dem MRT, bearbeitet. Zum Abschluss der Testsession füllten die Versuchspersonen den Fragebogen zur Test- und Computererfahrung aus.

8 Datenanalyse

Zur Bearbeitung der formulierten Fragestellungen wurden quantitative und qualitative Methoden zur Datenanalyse eingesetzt. Die quantitativen Analysen, d. h. deskriptive und inferenzstatistische Verfahren, wurden mit der Statistik- und Analysesoftware SPSS 15.0 (SPSS Inc.) durchgeführt. Einige der Verfahren sind auch mit Hilfe eigener Skripte für MATLAB 6.5 und der Statistik-Toolbox (The MathWorks, Inc.) berechnet worden. Mit MATLAB wurden auch die exploratorischen Verfahren zur Klassifizierung implementiert. Für die Vorverarbeitung komplexerer Rohdaten wurden eigene Tools in den Programmiersprachen Perl (The Perl Foundation) und Python™ (Python Software Foundation) entwickelt. Zur Visualisierung der Spurdaten und der Aggregation gebräuchlicher Leistungsmaße aus dem Water-Maze kam das Tool „Dataviewer“ zum Einsatz (s. Abschn. 6.3.3).

8.1 Deskriptive Verfahren und Inferenzstatistik

Von allen erhobenen Maßen wurden Häufigkeiten berechnet, deren Verteilung durch Histogramme dargestellt und mit dem Kolmogorov-Smirnov-Anpassungstest auf Normalverteilung geprüft wurden. Spezielle Histogramme für Winkelschätzfehler wurden in Form von Rose-Plots erstellt. Als Zusammenhangsmaße wurden je nach Skalenniveau Produkt-Moment-Korrelationen nach Pearson, punkt-biseriale Korrelationen und Korrelationskoeffizienten nach Spearson berechnet. Signifikanztests der Korrelationskoeffizienten wurden zweiseitig durchgeführt. Zur Einschätzung der Größe der Zusammenhänge wurde der von Cohen (1988) gegebene Einteilung für Korrelationsmaße in schwache ($|r| \approx 0.1$), mittlere ($|r| \approx 0.3$) und starke ($|r| \approx 0.5$) Effekte gefolgt (zit. n. Wirtz & Nachtigall, 2006).

Gruppenunterschiede wurden auf der Basis von Mittelwertsvergleichen zwischen den Faktorstufen der unabhängigen Variablen durchgeführt. Für Vergleiche zwischen zwei Gruppen (z. B. Geschlecht) wurden t-Tests für unabhängige Stichproben eingesetzt und für Vergleiche zwischen mehreren Gruppen (Altersgruppen) einfaktorielle Varianzanalysen (ANOVA). Zur Berücksichtigung der Interaktionen zwischen den Faktorstufen zweier unabhängiger Variablen (UVs) wurden zweifaktorielle Varianzanalysen berechnet.

Einfaktorielle Kovarianzanalysen (ANCOVA) wurden zur statistischen Kontrolle des Einflusses der Vortestleistung auf die abhängigen Variablen in Prä-Post-Testdesigns verwendet. Als zusätzliche Voraussetzung gegenüber der Varianzanalyse fordert die Kovarianzanalyse eine Überprüfung der Homogenität der Regressionskoeffizienten. In einigen Fällen genügte dieser Test bereits dem Er-

kenntnisgewinn, da es sich lediglich um die Überprüfung von Hypothesen zur Wirkung der Kovariaten in verschiedenen Treatments handelte.

Die Wirkung verschiedener Treatmentstufen innerhalb eines Experiments wurde durch Varianzanalysen mit Messwiederholung (RM ANOVA) bestimmt (z. B. Versuchsbedingungen der Zeigeaufgabe). Bei Verletzung der Forderung nach unkorrelierten Bedingungen wurde eine Korrektur der Freiheitsgrade nach Greenhouse-Geisser durchgeführt.

Die Quantifizierung der Gruppeneffekte erfolgte durch unterschiedliche Effektstärkemaße. Cohens d diente zur Beurteilung des Unterschieds zwischen zwei Gruppen und erlaubt eine Einschätzung der Größe des Effekts in klein ($d=0.2$), mittel ($d=0.5$) und groß ($d=0.8$). Bei Vergleichen von Gruppen sehr unterschiedlichen Umfangs wurde die von Cohen (1977) vorgeschlagene Korrektur der Gesamtvarianz angewendet (zit. n. Wirtz & Nachtigall, 2006). ETA-Quadrat wurden aus der Überprüfung der Modelle der Varianzanalysen gewonnen und erlauben die Beurteilung der Varianzaufklärung bezüglich einzelner Faktoren und Prädiktorvariablen.

In Fällen, in denen die Voraussetzungen für parametrische Tests verletzt waren, wurden Mittelwertsunterschiede mit nicht-parameterischen Verfahren auf ihre statistische Bedeutsamkeit getestet. Für unabhängige Stichproben waren dies der Mann-Whitney U-Test (zwei Stichproben) bzw. der Kruskal-Wallis H-Test (mehrere Stichproben), für abhängige Stichproben Wilcoxon's Rangsummentest bzw. der Friedman-Test (Bortz & Lienert, 2003).

Da die Analyse individueller Bearbeitungsstrategien und deren Zusammenhang mit den unabhängigen Variablen untersucht werden sollten, waren Kreuz- bzw. Kontingenztabelle und verschiedene Kontingenzmaße notwendig. Im Wesentlichen wurden dazu die Prüfgröße des Chi-Quadrat-Test und der korrigierte und normierte Kontingenzkoeffizient C_{kor} berechnet. Alle diese Maße quantifizieren die beobachteten Häufigkeiten gegenüber den erwarteten Häufigkeiten in den Zellen einer Kontingenztabelle. Sie unterscheiden sich hinsichtlich der Unabhängigkeit ihrer Aussagekraft von der Stichprobengröße. Zusätzlich wurden korrigiert-standardisierte Residuen berechnet, mit denen Aussagen über signifikante Abweichungen von der erwarteten Häufigkeit auf einzelnen Faktorstufen getroffen werden können.

8.2 Exploratorische Verfahren der Musteranalysen

Klassifizierende Verfahren wurden im Rahmen der Identifikation von Lösungs- und Bearbeitungsstrategien eingesetzt. Dazu wird angenommen, dass sich eine Stichprobe aus Personen zusammensetzt, die unterschiedliche Strategien anwenden. Personen mit ähnlichen Strategien lassen sich in einer Subgruppe zusammenfassen. Der Ansatz zur Identifikation der Strategien geht davon aus, dass durch die Beschreibung des Lösungsverhaltens der Gruppe Aussagen über die Art der angewendeten Strategien getroffen werden können. Dieser Ansatz und die verwendete mathematische Methode wird im Folgenden genauer erläutert.

Die Anwendung bestimmter Strategien stellt ein latentes, d. h. nicht direkt beobachtbares, Merkmal einer Person dar. Beobachtbar ist nur die Manifestation der Strategieanwendung in spezifischen Ver-

haltensmustern. Diese Muster können bei mehreren Versuchspersonen gleichermaßen beobachtete Merkmalsausprägungen oder aber besonders auffällige Anteile im Gesamtverhalten einer Person sein. Zur Erkennung solcher Muster dienen strukturentdeckende multivariate Analyseverfahren, die zum einen komplexe Verhaltensdaten klassifizieren und zum anderen faktorisieren.

Ziel einer Klassifizierung ist die Partitionierung der Menge aller Versuchspersonen in exhaustive und disjunkte Klassen, d. h. jede Person wird genau einer Klasse zugeordnet. Innerhalb einer Klasse sollen sich dabei Personen mit möglichst ähnlichen Merkmalen befinden. Dagegen sollen sich Personen verschiedener Klassen hinsichtlich dieser Merkmale deutlich unterscheiden. Clusteranalysen ermitteln Subgruppen (Cluster) von Personen, die hinsichtlich bestimmter Eigenschaften Ähnlichkeiten bzw. Unähnlichkeiten aufweisen (Duda et al., 2000).

Faktorisierung ist die Zusammenfassung von Variablen oder Merkmalen zu wenigen orthogonalen Faktoren mit dem Ziel der Dimensionsreduktion bei minimalem Informationsverlust. Spezielle Methoden zur Dekomposition von Matrizen eignen sich dazu und haben darüber hinaus die Eigenschaft in komplexen Daten wiederkehrende Anteile zu entdecken, die als grundlegende Verhaltensmuster (Building-Blocks) interpretiert werden können.

Zur Analyse der Bearbeitungsstrategien aus den Verhaltensdaten der hier durchgeführten Experimente wird ein Verfahren verwendet, das eine Heuristik zur Matrix-Faktorisierung mit einem speziellen Clusteralgorithmus kombiniert. Die Methoden mit den Namen Consensus-Clustering und Nicht-negative Matrix-Faktorisierung wurden zu unterschiedlichen Anteilen von mehreren Arbeitsgruppen entwickelt und werden in den folgenden beiden Abschnitten genauer beschrieben. Das ursprüngliche Anwendungsgebiet war die Klassifizierung von Tumorzellen anhand von Genexpressionsdaten. Diese Daten zeichnen sich durch hochdimensionale Merkmalsvektoren bei gleichzeitig geringem Stichprobenumfang aus. Eine Eigenschaft, die auch die Spurdaten aus dem Water-Maze haben. Aber das Verfahren konnte auch auf die dichotomen Antwortmuster ausgewählter Raumvorstellungstests angewendet werden.

8.2.1 Nicht-negative Matrix-Faktorisierung

Matrix-Faktorisierung ist allgemein die Suche nach einer Zerlegung einer Matrix X in ein Produkt aus zwei oder mehreren Matrizen. Zwei der grundlegenden Verfahren sind die QR-Zerlegung und die Singulärwertzerlegung, auf die sich die meisten anderen Dekompositionsverfahren als Spezialfall zurückführen lassen. Eines dieser speziellen Verfahren, das auf Grund günstiger Eigenschaften als Methode in dieser Arbeit angewendet werden soll, ist die nicht-negative Matrix-Faktorisierung.

Nicht-negative Matrix-Faktorisierung (NMF) ist eine ganze Gruppe von Faktorisierungsalgorithmen, die eine Matrix X in zwei Matrizen W und H zerlegen. Zeilen und Spalten von W und H werden dabei so bestimmt, dass das Produkt $W \cdot H$ eine Approximation von X ist. Der Unterschied zu anderen Methoden ist, dass die Elemente aller drei Matrizen immer größer oder gleich Null, d. h. *nicht-negativ*, sind. Das ist ein großer Vorteil für die Interpretierbarkeit der Werte. Ein weiterer Vorteil ist, dass die Zerlegung als Repräsentation der Daten aus ihren grundlegenden Mustern

(Building-Blocks) verstanden werden kann. Bekannt wurde das Verfahren zunächst durch Lee und Seung (1999), die es zur Faktorisierung von Bilddaten einsetzten. Es ist aber prinzipiell auf eine Vielzahl von Problemen der statistischen Analyse multivariater Daten anwendbar.

Dazu stellt man sich einen multivariaten Datenvektor x_i der Länge n vor, der Spaltenvektor der $n \times m$ Matrix X wird. Die Anzahl m dieser Spaltenvektoren ist dann die Stichprobengröße. Diese Matrix wird approximativ in eine $n \times k$ Matrix W und eine $k \times m$ Matrix H zerlegt. Die Gleichung für die volle Dekomposition lautet dann $X = W \cdot H + R$. Das Ergebnis ist eine Reduktion der Originaldaten auf eine Matrix W mit k Basisfunktionen (Muster), eine Koeffizientenmatrix H und eine Residuenmatrix R . Zu beachten ist, dass R dabei negative Werte enthalten kann. Der Modellparameter k wird so gewählt, dass dieser klein gegen m und n ist. Als obere Schranke wird $k \leq m \cdot n / (m+n)$ festgelegt.

Das Problem der Faktorisierung einer Matrix ist nicht direkt lösbar, sondern kann nur durch Heuristiken angenähert werden. Es gibt für die NMF verschiedene Approximationsalgorithmen, je nachdem, ob sie auf der Anwendung von Divergenzfunktionen (Kostenfunktion) oder auf Regularisierungstechniken beruhen.

Lee und Seung (2001) verwendeten u. a. die euklidische Distanz als Divergenzfunktion und entwickeln daraus Update-Regeln für einen iterativen Lösungsansatz.

Regularisierungstechniken fassen das Problem der Matrix-Faktorisierung als Lösung linearer Gleichungssysteme auf. Deren Lösung kann aber nicht direkt berechnet werden, da durch $k < n$ überbestimmte und somit schlecht-konditionierte Systeme vorliegen. In diesem Fall kann durch den Einsatz der Tikhonov-Regularisierung (Tikhonov & Arsenin, 1977; Neumaier, 1998) das schlecht gestellte Problem durch ein lineares Optimierungsproblem ersetzt und numerisch gelöst werden.

Aus einem Wechsel zwischen der Anwendung der Regularisierungstechnik und der Update-Regeln lässt sich eine NMF-Heuristik implementieren, die die Vorteile beider Methoden kombiniert. Die Idee, W zu fixieren und dann das Minimierungsproblem bezüglich H zu lösen, wurde von verschiedenen Arbeitsgruppen unter dem Begriff *sparse Non-negative Matrix Factorization* (SNMF) umgesetzt (Hoyer, 2004; Shahnaz, 2004; Gao & Church, 2005). Abbildung 8.1 zeigt den SNMF-Algorithmus von Gao und Church (2005).

Zur Eingabe erhält der Algorithmus die Datenmatrix X , die Anzahl der Basisfunktionen als Modellparameter k , den Tichonov-Faktor λ und die Anzahl der durchzuführenden Iterationen N . Der Algorithmus konvergiert schnell, innerhalb weniger Hundert Iterationen ($N=500$). Der Faktor λ wird auf einen sehr kleinen positiven Werte nahe Null gesetzt ($\lambda = 0.01$). Durch diese Erweiterung wird eine numerische Lösung des linearen Minimierungsproblems ermöglicht (s. Abb. 8.1, Zeile 2.a). Für eine Approximation der h_j verwendet die Mathematiksoftware MATLAB[®] (The MathWorks, Inc) numerische Verfahren zur Bestimmung der Pseudoinversen einer Matrix (Moore, 1920; Penrose, 1955; Anderson et al., 1999). Mit dieser Software wurde der Algorithmus für die Datenanalyse der vorliegende Arbeit umgesetzt. Das Matlab-Skript von Gao und Church (2005) ist im Anhang 16 abgedruckt.

Abbildung 8.1: Der SNMF-Algorithmus nach Gao & Church (2005).

Input: X, k, λ, N **Result:** W, H 1. **Initialize** W with random pos. values2. **Do** N iterationsa. **Solve for each column** j

$$\min_{h_j} \left\{ \|x_j - Wh_j\|^2 + \lambda^2 \|h_j\|^2 \right\}$$

b. **If** $h_{ij} < 0$ **then** $h_{ij} = 0$ c. **For each** w_{ij}

$$w_{ij} = w_{ij} \frac{(XH^T)_{ij}}{(WHH^T)_{ij}}$$

Eine wichtige Eigenschaft des SNMF-Algorithmus ist die Erzeugung „dünn besetzter“ Matrizen W und H , d. h. mit möglichst wenig von Null verschiedenen Einträgen. Kontrolle über die Anzahl der Nullen in H erhält man bei der Lösung des Minimierungsproblems durch den Faktor λ . Größere λ führen hier zu einer dünner besetzten Matrix, allerdings auf Kosten der Exaktheit der Approximation. Diese lässt sich wiederum durch die Anwendung der Update-Regel auf die Einträge der Matrix W lokal verbessern (s. Abb. 8.1, Zeile 2.c). Ist H dünn besetzt, werden dadurch auch in W möglichst viele Nullen propagiert, wodurch sich die Zahl orthogonaler Basisvektoren erhöhen lässt. Auf diese Weise findet der Algorithmus zuverlässig eine auf Grundmustern basierende Faktorisierung $W \cdot H$, deren Güte mit $\|R\|^2 / \|X\|^2$ (relatives Residuum) abgeschätzt werden kann.

8.2.2 Consensus-Clustering

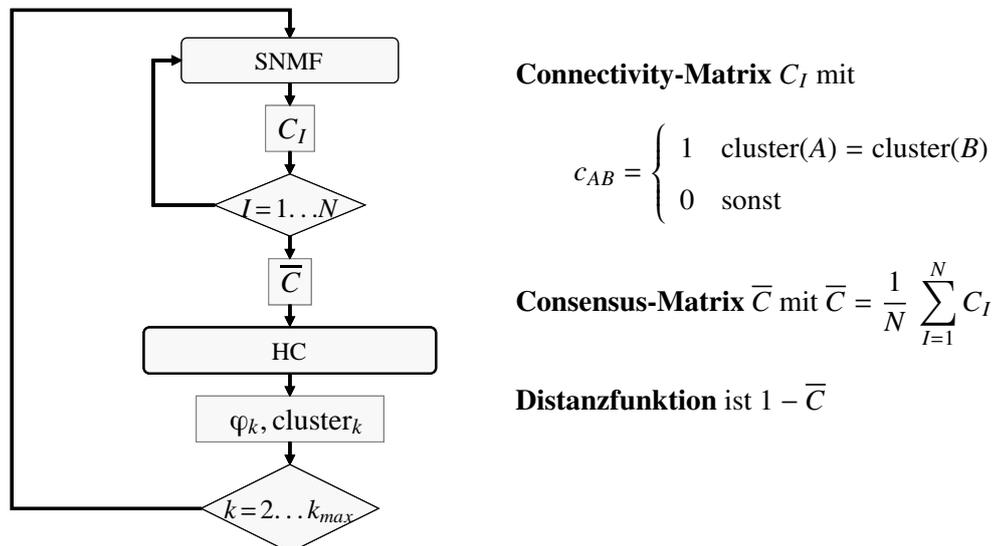
In einer Clusteranalyse werden Personen durch Merkmalsvektoren repräsentiert, die einen Vektorraum aufspannen, in dem jede Dimension einem Merkmal entspricht. Die Personen sind dann als Punkte innerhalb dieses Raumes abgebildet, deren Abstände zueinander durch Distanz- oder Ähnlichkeitsmaße definiert werden (z. B. durch die Euklidische Distanz). Im agglomerativ hierarchischen Clustern, das hier nur von Interesse ist, erfolgt die Gruppierung durch die sukzessive Zusammenfassung der Datenpunkte zu immer größeren Clustern. Dabei entsteht eine Hierarchie, in der auf jeder Ebene Cluster der darunter liegenden Ebene vereinigt werden. Zur Zusammenfassung sind Aggregationsmethoden notwendig, die der Bewertung des Abstands zweier Cluster dienen. Die Methode *average linkage* berechnet dazu den durchschnittlichen Abstand aller Elementpaare aus zwei Clustern.

Monti et al. (2003) stellten unter dem Begriff Consensus-Clustering ein Framework zur Validie-

zung von Clusterlösungen und zur automatischen Klassenentdeckung vor. Besonders Heuristiken mit randomisierten Startbedingungen, wie die NMF-Methode, können auf ihre Stabilität hin untersucht werden. Brunet, Tamayo, Golub und Mesirov (2004) und Gao und Church (2005) kombinierten erfolgreich den vorgestellten SNMF-Algorithmus mit dieser Methode zu einem automatischen Klassifizierungsverfahren.

Abbildung 8.2 zeigt den dazugehörigen Algorithmus als Flussdiagramm. Zentral ist eine agglomerative hierarchische Clusterheuristik mit *average linkage* als Aggregationsmethode (HC). Diese erhält als Eingabe die Consensus-Matrix \bar{C} , eine Ähnlichkeitsmatrix, die die Häufigkeit angibt mit der zwei Personen ein und derselben Klasse zugeordnet wurden. Die Zuordnung ergibt sich aus der Faktorisierung des NMF-Algorithmus. Dazu wird die Koeffizientenmatrix H zur Bestimmung der Klassenzugehörigkeit genutzt. Person j wird Cluster i zugewiesen, wenn h_{ij} der größte Eintrag in der Spalte h_j ist. Der Hintergrund dieser Auswahl ist die Bedeutung der Einträge der Koeffizienten h_{ij} als Ausprägung der Basisfunktion i in der Person j , d. h. die Klasse einer Person wird durch seine „dominanteste“ Basisfunktion beschrieben.

Abbildung 8.2: Flussdiagramm Consensus-Clustering kombiniert mit SNMF (Gao & Church, 2005).



Da der SNMF-Algorithmus mit randomisierter Startbedingung arbeitet, wird dieser N -mal aufgerufen, um \bar{C} zu ermitteln. In jedem Lauf I werden zunächst die Clusterzuordnungen in eine Connectivity-Matrix C_I eingetragen. Werden zwei Personen A und B dem gleichen Cluster zugewiesen, wird der entsprechende Eintrag c_{AB} in der Matrix auf 1 gesetzt. Eine Mittelung über alle Matrizen C_I ergibt dann die Consensus-Matrix. Die HC-Heuristik arbeitet aber nicht mit den Ähnlichkeiten sondern mit den Distanzen, die sich leicht aus \bar{C} berechnen lassen (Abb. 8.2).

Wiederholt man diese Schritte für verschiedene Modellparameter k , erhält man quasi durch „Probieren“ eine bestmögliche Klasseneinteilung. Die Güte einer Klassifizierung lässt sich durch den kophenetischen Korrelationskoeffizienten φ_k und durch die Consensus-Matrix einschätzen. Der kophenetische Korrelationskoeffizient prüft den Zusammenhang zwischen den gegebenen und während der Agglomeration bestimmten Distanzen. Er kann Werte zwischen 0 und 1 annehmen. Werte

für φ nahe 1 deuten auf einen geringen Informationsverlust während des Clusters. Für die Wahl des Modellparameters k betrachtet man entweder nur die Lösung mit dem größten φ oder alle Lösungen vor einem rapiden Abfall der Kurve in einem φ - k -Diagramm.

Aufschluss über die Eindeutigkeit der Klasseneinteilung geben Plots der geordneten Consensus-Matrizen. Jeder Datenpunkt gibt die Häufigkeit an, mit der zwei Personen dem gleichen Cluster zugeordnet wurden. Gebräuchlich ist auch die Darstellung der Einträge der oberen Dreiecksmatrix von \bar{C} in einem Histogramm.

Das Verfahren konnte bei Brunet et al. (2004) und bei Gao und Church (2005) in der Klassifizierung von Tumorzellen bessere Ergebnisse als jedes andere in diesem Bereich etablierte Verfahren erzielen. Weitere Vorteile sind: Die Methode ist intuitiv und deshalb für den Anwender gut verständlich, außerdem ist die Güte der Klassifikation durch die Visualisierung der Consensus-Matrix sehr anschaulich beurteilbar.

Da weder die Methode Consensus-Clustering noch SNMF Bestandteile aktuell verfügbarer Statistikprogramme sind, wurden die entsprechenden Algorithmen anhand der Angaben aus den Publikation von Monti et al. (2003) und Brunet et al. (2004) mit MATLAB 6.5 (The Mathworks Inc.) implementiert und auf artifiziellen Benchmarks („Schwimmer“-Datensatz; Donoho & Stodden, 2003) getestet. Die original Matlab-Skripte von Brunet et al. (2004) zum Consensus-Clustering stehen Online zur Verfügung¹ und befinden sich in der für die vorliegende Arbeit modifizierten Version im Anhang 16.

¹<http://www.broad.mit.edu/cancer/pub/nmf>

Ergebnisteil

9 Experiment 1 - Mentale Rotation

9.1 Allgemeine kognitive Maße

Tabelle 9.1 fasst die erhobenen Leistungsmaße zur „induktiven Denkfähigkeit“ und zur Raumvorstellung zusammen. Da sich die Stichprobenumfänge der Altersgruppen stark unterscheiden, wurden nur innerhalb dieser bestehende Geschlechtsdifferenzen betrachtet.

Tabelle 9.1: Schulleistung bzw. -abschluss und Testleistungen in den ausgewählten Untertests des PSB-R bzw. LPS 50+.

		männlich		weiblich		t	df	p
		M	SD	M	SD			
<i>Schulleistung</i>								
Mathenote	Schüler	2,5	0,73	2,5	0,65	0,06	203	.954
<i>Induktive Denkfähigkeit (Reihen)</i>								
PSB-R AG 2,3,4	Schüler	26,6	5,18	27,3	5,64	0,90	203	.371
	Studenten	65,2	6,95	63,4	6,74	0,78	34	.441
LPS 50+ UT 3,4	Senioren	46,4	7,76	46,5	7,54	0,04	18	.972
<i>Räumliche Vorstellung (Flächenzählen)</i>								
PSB-R AG 6	Schüler	27,2	5,37	29,0	4,75	2,25	203	.026
	Studenten	29,8	4,58	27,7	5,32	1,28	34	.211
LPS 50+ UT 6	Senioren	23,5	3,63	21,3	5,53	0,97	18	.344

Es wurden keine signifikanten Mittelwertsunterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der „induktiven Denkfähigkeit“, festgestellt. Lediglich in der Fähigkeit Raumvorstellung, operationalisiert durch den Untertest *Flächenzählen*, unterschieden sich Schüler und Schülerinnen. Im Mittel erreichten die Schülerinnen sogar eine etwas höhere Punktzahl als ihre männliche Vergleichsgruppe.

Zum Screening allgemeiner schulischer Leistung wurde für die Altersgruppe der Schüler die letzte bekannte Schulnote im Fach Mathematik herangezogen. Studenten wurden nach ihrer Studienfachrichtung befragt, die sich in den Häufigkeiten zwischen den Geschlechtern unterschieden ($\chi^2=22,71$; $df=3$; $p<.001$). Die Studentinnen waren überrepräsentiert durch das Fach Psychologie, die Studenten durch naturwissenschaftliche Studiengänge.

Bei den Senioren wurde nach der Art des Schulabschlusses und der Berufsausbildung gefragt. In beiden Variablen bestanden keine signifikanten Unterschiede in den Häufigkeitsverteilung bezüglich des Faktor „Geschlecht“ ($\chi^2=1,75$; $df=4$; $p=.782$ bzw. $\chi^2=1,55$; $df=3$; $p=.671$). Die meisten der befragten Personen gaben eine mindestens zehnjährige Schulausbildung und eine zwei- bis dreijährige Berufsausbildung an.

9.2 Vorerfahrung

In der Tabelle 9.2 wurden die Ergebnisse zur Befragung der Vorerfahrungen mit Computerspielen und Anwendungsprogrammen in drei Kategorien zusammengefasst, die den zeitlichen Umfang der Nutzung widerspiegeln. Allgemein nutzten Schüler und Senioren seltener und in geringerem Umfang den Computer zum Arbeiten. Hier gab es auch keine Geschlechtsunterschiede. Lediglich Studenten beschäftigen sich häufiger mit Anwendungsprogrammen als Studentinnen.

Tabelle 9.2: Vorerfahrung Computer. Häufigkeit der Nutzung getrennte nach Anwendungsprogrammen und Spielen. Kategorien: 1: nie-selten 2: wöchentlich, weniger als 1h täglich 3: täglich mehrere Stunden.

	männlich				weiblich				χ^2	df	p
	N	1	2	3	N	1	2	3			
<i>Anwendungsprogramme</i>											
Schüler	146	75%	21%	4%	60	85%	13%	2%	2,73	2	.255
Studenten	18	-	39%	61%	18	17%	67%	17%	8,89*	2	.012
Senioren	8	63%	25%	13%	12	33%	58%	8%	2,18	2	.337
<i>Computerspiele</i>											
Schüler	146	65%	23%	12%	60	90%	7%	3%	13,21**	2	.001
Studenten	18	17%	55%	28%	18	83%	17%	-	16,77**	2	<.001
Senioren	8	88%	12%	-	12	92%	8%	-	0,09	1	.761

*<0.05, **<0.01

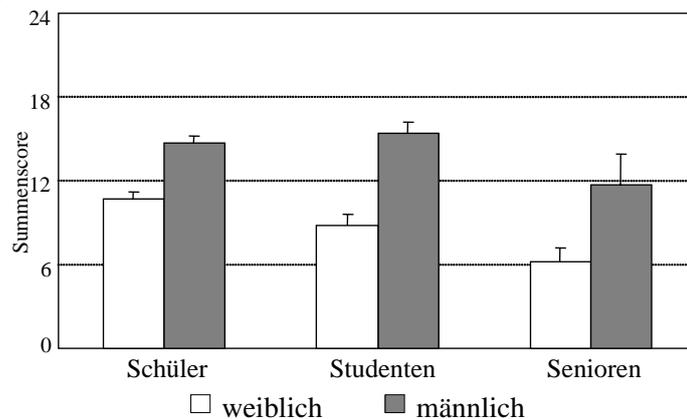
In der abhängigen Variable Computerspiele zeigten sich deutlichere Geschlechtsunterschiede in der Häufigkeit der Nutzung, sowohl für die Studenten als auch für die Schüler. Hier waren es die männlichen Versuchspersonen, die angaben, häufiger und längere Zeit mit Computerspielen beschäftigt zu sein. Bereits zwölf Prozent der Schüler und 28 % der Studenten spielte, laut eigenen Angaben, täglich mehrere Stunden am Computer. Senioren gaben hier, unabhängig vom Geschlecht, am häufigsten, an gar nicht oder nur selten Computerspiele zu verwenden.

9.3 Raumkognitive Maße

9.3.1 Mental Rotations Test A

Gruppenmittelwerte des MRT-Summscores sind in Abbildung 9.1 dargestellt. Innerhalb jeder der drei Altersgruppen konnten signifikante Geschlechtsunterschiede nachgewiesen werden. Die Scores in den drei Alterstichproben lagen normalverteilt vor (Kolmogorov-Smirnov-Test: Schüler: $Z=1,13$; $p=.154$; Studenten: $Z=1,13$; $p=.154$; Senioren: $Z=0,72$; $p=.684$). Zum Teil bestanden Varianzunterschiede zwischen den Geschlechtergruppen (Levene-Test: Schüler: $F= 5,99$; $p=.015$; Studenten: $F=0,00$; $p=.98$; Senioren: $F=5,25$; $p=.034$).

Abbildung 9.1: Geschlechtsdifferenzen im MRT-A. Mittelwertsunterschiede und SEM.

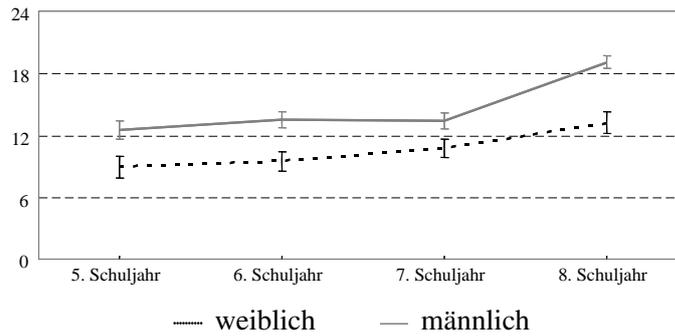


Die Mittelwertsdifferenzen zwischen männlichen und weiblichen Schülern sind mit $t(137,30)=-5,86$ ($p<.001$) signifikant und mit $d=0.82$ als großer Effekt einzustufen. Die 60 Schülerinnen erreichten im Mittel 10,7 Punkte ($SD=4,17$ Punkte), die 146 Schüler 14,7 Punkte ($SD=5,26$ Punkte). Die 18 männlichen Studenten erzielten mit $M=15,4$ Punkte ($SD=3,33$ Punkte) im Mittel den höchsten Summscore. Dadurch dass der Mittelwert der 18 weiblichen Studenten mit 8,8 Punkten ($SD=3,32$ Punkte) deutlich darunter lag, fiel der Effekt mit $d=2.0$ und $t(34)=-5,97$ ($p<.001$) extrem groß aus. Nur der Mittelwert der zwölf Seniorinnen lag noch darunter ($M=6,3$ Punkte; $SD=3,25$ Punkte). Im Vergleich erzielten die acht Senioren in der Untersuchung im Mittel 11,8 Punkte ($SD=5,97$ Punkte) im MRT-A. Auch hier war der Geschlechtseffekt mit $d=1.2$ signifikant ($t(9,80)=-2,38$; $p=.039$).

In der Mittelung der Summscores innerhalb der Stichprobe der Schüler gingen z. T. große Leistungsunterschiede zwischen den Klassenstufen unter. In Abbildung 9.2 sind die MRT-Mittelwerte in Abhängigkeit vom Schuljahr dargestellt. Die Mittelwerte unterschieden sich zwischen den Schuljahresstufen signifikant ($F(3,202)=17,28$; $p<.001$). Dieser Unterschied ging auf einen signifikant höheren Score der des achten Schuljahrs gegenüber allen anderen Jahrgangsstufen zurück. Diese galt sowohl für die Jungen als auch für die Mädchen, so dass die Geschlechtsdifferenzen über die Jahrgänge hinweg stabil blieben.

Für das fünft bis siebte Schuljahr lag der Geschlechtseffekt zwischen $d=0.59$ bis $d=0.93$. Mädchen dieser Schulstufen erzielten im Mittel zwischen 8,9 und 10,7 Punkte, Jungen zwischen 12,5 bis 13,4

Abbildung 9.2: MRT-A Summenscore in den einzelnen Schuljahren getrennt nach Geschlecht.



Punkte. Für das achte Schuljahr lagen die Mittelwerte der Mädchen bei 13,2 Punkten (SD=4,18 Punkte), die der Jungen aber noch weiter darüber (M=19,0 Punkte; SD=3,85 Punkte). So konnte ein Anstieg des Geschlechtseffekts auf $d=1.48$ beobachtet werden.

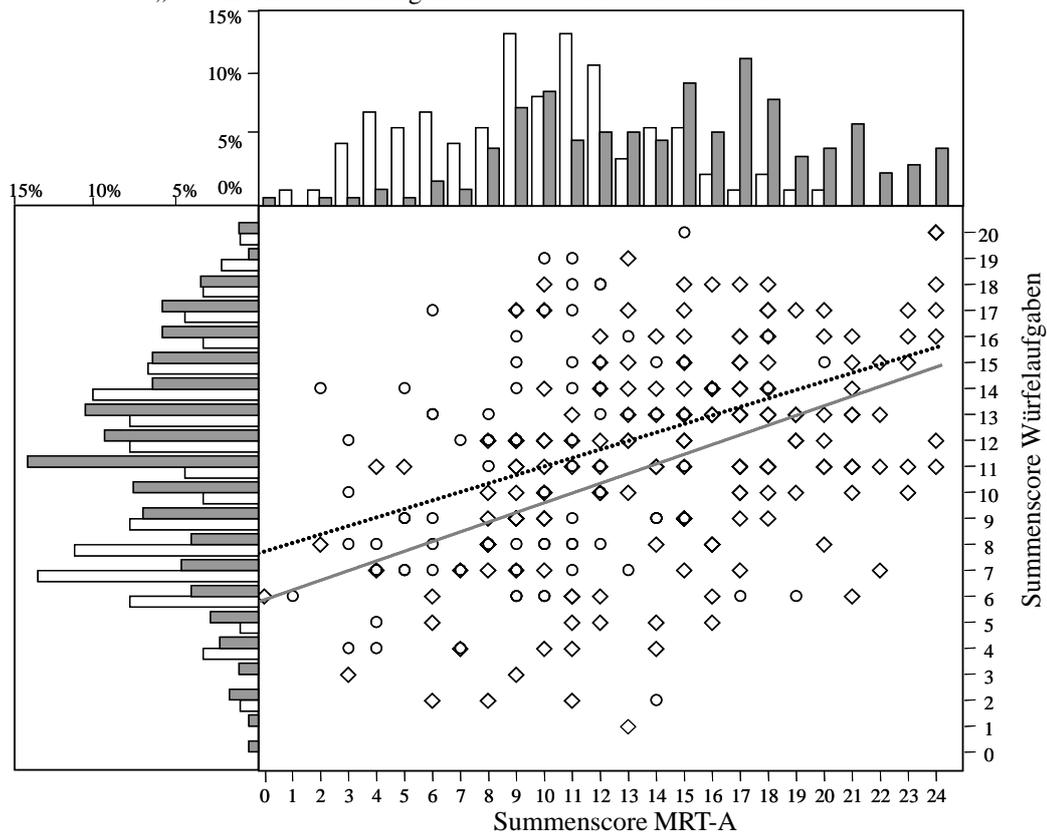
9.3.2 Würfelaufgaben

Die Summenscores aus den *Würfelaufgaben* lagen in den Alterstichprobe normalverteilt vor (Kolmogorov-Smirnov-Test: Schüler: $Z=1,26$; $p=.084$; Studenten: $Z=0,89$; $p=.414$; Senioren: $Z=0,58$; $p=.889$). Geschlechtsdifferenzen konnten in keiner der Stichproben gefunden werden. Die statistischen Test ergaben keine signifikanten Prüfgrößen (Schüler: $t(204)=0,47$; $p=.637$ bzw. $Z=-0,50$; $p=.618$; Studenten: $t(27,92)=0,30$; $p=.765$; Senioren: $t(7,80)=0,85$; $p=.420$). Im Mittel erreichten Schülerinnen und Schüler 10,9 Punkte (SD=3,62 Punkte) bzw. 11,2 Punkte (SD=3,98 Punkte) in den *Würfelaufgaben*. Studentinnen und Studenten erreichten im Mittel 13,8 Punkte (SD=4,69 Punkte) bzw. 14,2 Punkte (SD=2,81 Punkte) in den *Würfelaufgaben*. Senioren erzielten im Mittel 6,7 Punkte (SD=1,61 Punkte) bei den Frauen und 8,4 Punkte (SD=5,53 Punkte) bei den Männern.

Differenzen zwischen den Gruppenmitteln männlicher und weiblicher Versuchspersonen, wie im MRT, wurden in den *Würfelaufgaben* nicht beobachtet. Trotzdem scheinen Zusammenhänge zwischen den beiden psychometrischen Maßen zu bestehen. In Abbildung 9.3 sind diese Zusammenhänge grafisch dargestellt.

In den nach Geschlecht getrennt dargestellten Histogrammen wurden mögliche Unterschiede und Gemeinsamkeiten zwischen beider Maße deutlich. Die Summenscores im MRT lagen für beide Geschlechter in annähernd gleicher Form um einen festen Betrag gegeneinander verschoben vor. Differenzen bezüglich des Gruppenmittelwerts waren eindeutig identifizierbar. Die Verteilungsformen der Summenscores aus den *Würfelaufgaben* unterschieden sich dagegen sehr stark. Die Verteilung in der Stichprobe männlicher Versuchspersonen war unimodal normalverteilt mit Mittelwert, Modus und Median bei ca. elf Punkten. Die Verteilung der Summenscores bei den weiblichen Probanden war stark bimodal mit einer Häufung bei sieben bis acht Punkten und bei 14 Punkten. Dies könnte die Manifestation zweier unterschiedlicher Lösungsstrategien für die Lösungsaufgaben nur innerhalb der Gruppe weiblicher Versuchspersonen sein.

Abbildung 9.3: Zusammenhang zwischen Testleistung im MRT-A und den *Würfelaufgaben* getrennt für Faktor „Geschlecht“. Verteilungen und Scatter-Plot der Summenscores.



Histogramm: □ weiblich ■ männlich Beobachtet: ○ weiblich ◇ männlich
 Linear: — weiblich - - - männlich

Dieser Unterschied wurde auch in den geschlechtsspezifischen Zusammenhangsmaßen für MRT-A und *Würfelaufgaben* deutlich (s. Scatter-Plot Abb. 9.3 und Tab. 9.3). Während der prädiktive Charakter der MRT-Leistung auf das Testergebnis in den Würfelaufgabe mit $R^2=0.22$ auf einen mittleren Zusammenhang beider Maße bei männlich Versuchspersonen schließen lässt, sinkt der Anteil der Varianzaufklärung in der Stichprobe weiblicher Versuchspersonen auf ca. zehn Prozent.

9.3.3 Computer Mental Rotations Test

Für eine erste Überprüfung der Validität des selbstentwickelten Computertests wurde eine statistische Vergleich mit der bestehenden Version des MRT durchgeführt. Dazu wurden Schüler- und

Tabelle 9.3: Parameter der Regressionsanalyse (lineares Modell) zur Vorhersage der Summenscores in den *Würfelaufgaben* aus denen des MRT-A.

	R²	F	df	Sig.	Konstante	Anstieg
männlich	0.219	47,73	1;170	<.001	5,9	0,37
weiblich	0.107	10,59	1;88	.002	7,7	0,33

Studentenstichprobe in jeweils zwei Gruppen parallelisiert. Für die Altersgruppe der Senioren wurde diese Prozedur nicht durchgeführt. Sie nahmen alle am Computertest teil.

Kriterien für die Parallelisierung waren die Testleistungen im MRT-A und PSB-R, das Alter, das Geschlecht und die Erfahrung mit Computerspielen. Tabelle 9.4 fasst die entsprechenden Gruppenmittelwerte und Standardabweichungen aller für den Post-Test zur Verfügung stehenden Personen zusammen.

Tabelle 9.4: Alter, Vortestleistungen und Spieleerfahrung der parallelisierten Gruppen für den Vergleich Computertest vs. Papier-Bleistift-Test.

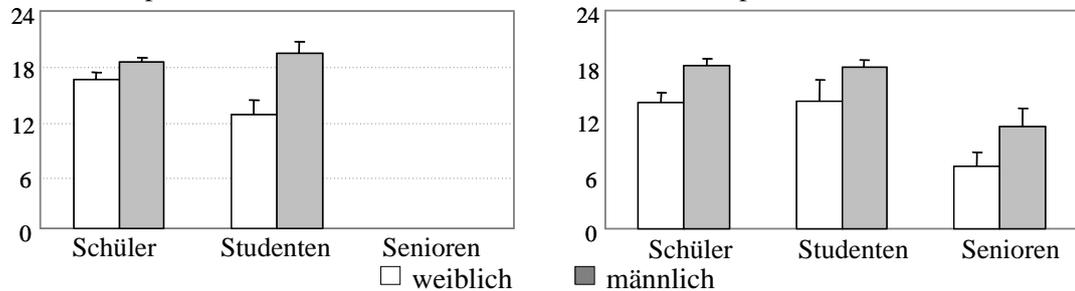
	Papier-Bleistift			Computertest			Levene-Statistik		t	df	p
	N	M	SD	N	M	SD	F	p			
<i>MRT-A</i>											
Schüler	100	13,3	5,40	106	13,8	5,20	0,07	.790	-0,68	204	.497
Studenten	25	12,7	5,10	11	10,7	3,41	2,47	.125	1,35	28,0	.188
<i>PSB-R Reihen</i>											
Schüler	100	27,1	5,04	106	26,6	5,57	0,63	.428	0,69	204	.489
Studenten	25	12,7	5,10	11	64,9	6,77	0,36	.554	-0,33	34	.742
<i>PSB-R Flächenzählen</i>											
Schüler	100	27,8	4,98	106	27,7	5,52	1,29	.258	0,19	204	.848
Studenten	25	64,1	6,95	11	28,9	7,05	3,28	.079	-0,08	12,9	.935
<i>Alter</i>											
Schüler	100	12,5	1,17	106	12,5	1,18	0,02	.888	0,06	204	.951
Studenten	25	22,8	3,52	11	22,4	4,20	0,07	.790	0,29	34	.771
<i>Geschlecht m:männlich w:weiblich</i>											
Schüler	m:72% w:28%			m:70% w:30%			$\chi^2=0,12$ p=.730				
Studenten	m:56% w:44%			m:36% w:64%			$\chi^2=1,18$ p=.278				
<i>Computerspiele 1:nie-selten 2:wöchentlich-täglich weniger 1h 3:täglich mehrere Stunden</i>											
Schüler	1:75%	2:15%	3:10%	1:70%	2:21%	3:9%	$\chi^2=1,15$ p=.561				
Studenten	1:44%	2:40%	3:16%	1:64%	2:27%	3:9%	$\chi^2=1,19$ p=.550				

Bezüglich der Kriterien gab es keine signifikanten Unterschiede zwischen der Gruppe „Papier-Bleistift-Test“ und der Gruppe „Computertest“. Jedoch nahmen signifikant weniger Studenten am Computertest teil, so dass durch die weitere Aufteilung durch den Faktor „Testversion“ die statistischen Tests auf sehr wenigen Versuchspersonen beruhen.

In Abbildung 9.4 sind Mittelwerte weiblicher und männlicher Versuchspersonen aufgeteilt nach Testversion und Altersgruppe dargestellt. Zu beobachten waren in der Gruppe der Schüler sehr hohe Gruppenmittelwerte. Im Papier-Bleistift-Test beantworteten Mädchen 16,5 (SD=4,16) Items richtig, Jungen sogar 18,3 (SD=5,13). Der Geschlechtseffekt reduzierte sich auf $d=0.38$ und war nur noch tendenziell signifikant ($t(98)=1,71$; $p=.091$). Im Computertest lösten Jungen im Mittel 18,0

Items korrekt (SD=5,86 Punkte)), Mädchen mit 13,9 Punkten (SD=6,03 Punkte) jedoch deutlich weniger. Der Geschlechtseffekt im Computertest blieb deshalb statistisch signifikant ($t(104)=3,27$; $p=.001$) und mit $d=0.69$ größer als im Papier-Bleistift-Test.

Abbildung 9.4: Geschlechtseffekte im MRT-B und CMRT-B Post-Test. Mittelwertdifferenzen und SEM. Papier-Bleistift (MRT-B)



Auch in allen anderen Untergruppen reduzierte sich der Geschlechtseffekt gegenüber der Vortestleistung im MRT-A, blieb aber weiterhin bestehen. Für die Studenten von $d=2.0$ auf $d=1.3$ ($t(23)=3,26$; $p=.003$) im Papier-Bleistift-Test und $d=0.73$ ($t(7,15)=1,53$; $p=.168$) im Computertest, für die Senioren von $d=1.2$ auf $d=0.82$ ($t(18)=1,79$; $p=.091$).

Im Gruppenvergleich zwischen Papier-Bleistift- und Computertest wurde deshalb der Faktor „Geschlecht“ mit berücksichtigt. Tabelle 9.5 fasst die Gruppenmittelwerte in dieser Richtung zusammen. Zusätzlich wurden die Anzahl bearbeiteter Items gegenübergestellt. Letztere lagen in allen Gruppen im Mittel so hoch in der Nähe der Maximalanzahl, dass von einer fast vollständigen Bearbeitung des Testmaterials ausgegangen werden konnte. Der einzige signifikante Unterschied zwischen der Computer- und der Papierversion bestand in der Gruppe der Schülerinnen. Es zeigte sich, dass Mädchen in der Wiederholung des MRT als Computertest schlechter abschnitten als vergleichbare Mädchen im Papier-Bleistift-Test. Bei den Jungen bestand kein solcher Unterschied. Deutlicher wurde diese in der Untersuchung der Prädiktorwirkung der Vortestleistung auf die Leistung im Post-Test.

In Abbildung 9.5 und Tabelle 9.6 sind unter der Annahme eines linearen Zusammenhangs Regressionsmodelle für den Summenscore des MRT-B (AV) und MRT-A (Prädiktorvariable) getrennt für die Faktoren „Geschlecht“ und „Testversion“ durchgeführt worden. Innerhalb der Gruppe männlicher Schüler konnte in beiden Subgruppen eine sehr gute Anpassung an ein lineares Regressionsmodell durchgeführt werden. Die Anstiege der Regressionsgeraden für Computertest und Papier-Bleistift-Test waren nominal beinahe identisch. Rund 39 % der Varianz in den Post-Testdaten konnte in beiden Fällen durch die Vortestleistung aufgeklärt werden. Somit war die Vortestleistung aus dem MRT-A für die männlichen Schüler ein sehr guter Prädiktor für den Post-Test.

Für die beiden Gruppen Schülerinnen konnten ebenfalls lineare Regressionsmodelle an die Daten angepasst werden. Obwohl die Prüfgröße des Signifikanztests für den Papier-Bleistift-Test auf einen schlechteren Fit hindeutet als im Computertest. Für letzteren ist der ermittelte Zusammenhang, ausgedrückt durch den quadrierten punkt-biserialen Korrelationskoeffizienten von $R^2=0.486$, fast perfekt, d. h. die Mädchen reproduzierten im Computertest beinahe identisch ihre Summenscores aus

Tabelle 9.5: Vergleich der Papier- und der Computerversion des MRT-B.

		Papier-Bleistift			Computertest			t	df	p
		N	M	SD	N	M	SD			
<i>Summenscore</i>										
Schüler	weiblich	28	16,5	4,16	32	13,9	6,03	1,95	55,1	.056
	männlich	72	18,3	5,13	74	18,0	5,86	0,39	144,0	.694
Studenten	weiblich	11	12,6	5,14	7	14,0	6,16	-0,51	16,0	.618
	männlich	14	19,3	5,00	4	17,8	1,50	0,59	16,0	.560
<i>Anzahl bearbeiteter Items</i>										
Schüler	weiblich	28	22,3	2,39	32	22,0	3,39	0,46	58,0	.648
	männlich	72	22,6	2,43	74	23,1	2,05	-1,26	138,7	.209
Studenten	weiblich	11	24,0	0,00	7	24,0	0,00		-	
	männlich	14	24,0	0,00	4	24,0	0,00		-	

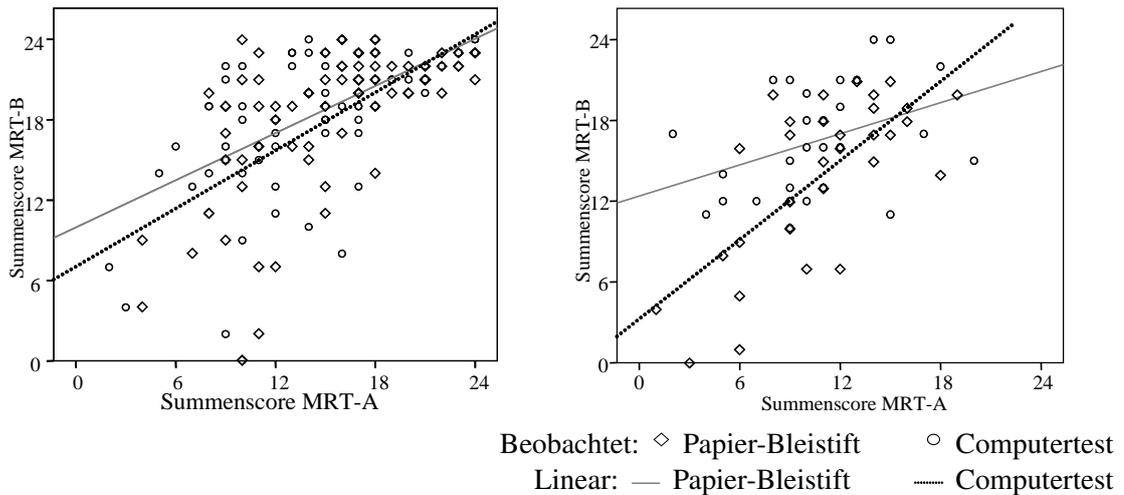
dem Vortest. Ganz anders war die Situation im Papier-Bleistift-Test. Hier konnten lediglich 14,6% der Varianz durch die Vortestleistung aufgeklärt werden. Dies ist auf den dramatischen Leistungszuwachs der Mädchen zurückzuführen, die im Vortest noch eher schwächere Leistungen zeigten. Dadurch wird auch die Regressionsgerade deutlich flacher.

Tabelle 9.6: Parameter der Regressionanalyse zur Vorhersage der Post-Testleistung aus dem MRT-A.

		R ²	F	df1	df2	Sig.	Anstieg	Konstante
männlich	Papier-Bleistift	0.388	43,80	1	69	<.001	0,59	10,0
	Computertest	0.385	45,11	1	72	<.001	0,72	7,0
weiblich	Papier-Bleistift	0.146	4,44	1	26	.045	0,38	12,4
	Computertest	0.486	28,38	1	30	<.001	0,98	3,3

Zur statistischen Prüfung der Unterschiede in den Anstiegen der Regressionsgeraden wurden Kovarianzanalysen mit MRT-A als Kovariate und Testversion als unabhängige Variable in den Ausprägungen „Papier-Bleistift-Test“ und „Computertest“ durchgeführt. Für die männlichen Schüler bestätigt sich der Eindruck des ausschließlichen Einflusses der Kovariate auf die Post-Testleistung. Zunächst bestand kein Interaktionseffekt zwischen Testversion und Kovariate ($F(1,141)=0,96$; $p=.329$), was den Eindruck der identischen Regressionsanstiege bestätigte und die Interpretation der Haupteffekte in der Kovarianzanalyse erst erlaubte. Der Effekt der Kovariate war als einziger in diesem Modell statistisch signifikant ($F(1,141)=88,24$; $p<.001$) und ist deshalb mit $\eta^2=0.385$ identisch mit dem Regressionsmodell. Der nicht signifikante Haupteffekt für den Faktor „Testversion“ ($F(1,141)=1,80$; $p=.182$), bestätigte, dass bei gleicher Vortestleistung, was durch die Parallelisierung gegeben war, die Leistungen der Jungen im Computertest und im Papier-Bleistift-Test sich

Abbildung 9.5: Zusammenhang MRT-A Vortest und MRT-B Post-Test getrennt für Geschlecht (linke und rechte Abbildung) und Testversion (innerhalb jeder Abbildung).
 männlich weiblich



nicht unterschieden.

Die Kovarianzanalysen für die Post-Tests der Schülerinnen ergaben einen signifikanten Interaktionseffekt ($F(1,56)=5,23; p=.026$). Dadurch konnten die Unterschiede in den Anstiegen der Regressionsgeraden nicht nur nominal, sondern auch als statistisch bedeutsam nachgewiesen werden. Durch die unterschiedliche Wirkung der Kovariate auf den Faktorstufen „Papier-Bleistift-Test“ und „Computertest“ waren aber auch die Haupteffekt der Kovarianzanalyse nicht eindeutig interpretierbar. Der Effekt der Kovariate war mit $F(1,56)=27,23$ ($p<.001$) signifikant und würde bei gleicher Wirkung in den Testversionen, was ja nicht gegeben war, $\eta^2=0.327$ betragen. Der Haupteffekt für den Faktor „Testversion“ betrug, bei gleicher Vortestleistung, was durch die Parallelisierung wiederum gegeben war, $\eta^2=0.141$ und war ebenfalls signifikant ($F(1,56)=9,18; p=.004$), d. h. durch die Hinzunahme der Computerversion kamen 14 % mehr Varianz in die Daten. In Bezug auf die Mittelwertsdifferenz betrug der Effekt 0.5 Standardabweichungen.

9.4 Musteranalysen

Musteranalysen zur Entdeckung gruppenspezifischer Lösungsstrategien in den Raumvorstellungstests MRT und *Würfelaufgaben* wurden mit der Methode Consensus-Clustering und SNMF auf den dichotomisierten Itemantworten für die Alterstichprobe der Schüler ($N=205$) durchgeführt.

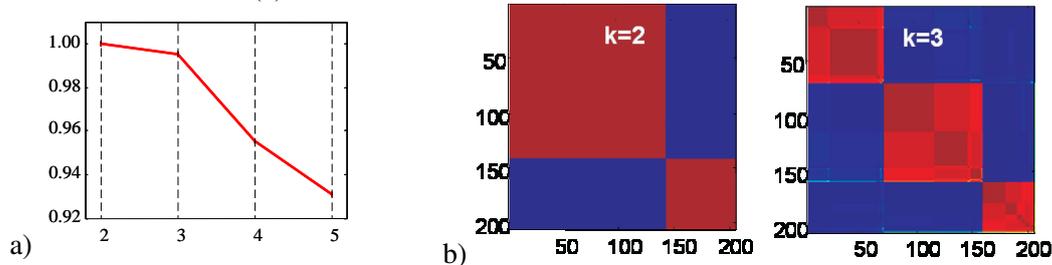
9.4.1 Mental Rotations Test A

Zur Dichotomisierung der Antwortmuster wurden im MRT-A Items, die nach dem strengen Punktesystem richtig beantwortet waren, mit Eins codiert und Items, die falsch beantwortet oder nicht bearbeitet waren, mit Null. So bestand ein Datenvektor einer Person aus $M=24$ dichotomen Variablen,

die in einer $M \times N$ Datenmatrix V zusammengefasst als Eingabe für die Matrix-Faktorisierung (NMF) diene. Zur Bestimmung des Modellparameters k (Anzahl der Subgruppen) wurde diese Heuristik, wie im Absch. 8 (S. 70) dargestellt, mit der Methode des Consensus-Clusterings kombiniert.

Zur Bestimmung des Modellparameters wurden Klassen-Lösungen zwischen $k=2$ und $k=5$ mit dem Verfahren berechnet. Eine Begrenzung auf maximal fünf Klasse wurde aus Gründen der besseren Interpretierbarkeit von Modellen mit weniger Klassen gewählt. Abbildung 9.6 stellt die aus dem Consensus-Clustering gewonnenen Kriterien zur Modellselektion dar. Der kophenetische Korrelationskoeffizienten $\varphi(k)$ für $2 \leq k \leq 5$ wurde als primäres Gütemaß der Klassifizierung aus dem hierarchischen Clusteralgorithmus gewonnen und in Abhängigkeit vom Parameter k in einem φ - k -Diagramm dargestellt (s. Abb. 9.6 a). Als Lösungen des Klassifizierungsproblems wurden diejenigen mit den höchsten Koeffizienten ausgewählt. In der Abbildung des φ - k -Diagramms für die Analyse der MRT-A-Items waren das $k=2$ mit $\varphi(k)=1,00$ und $k=3$ mit $\varphi(k)=0,996$.

Abbildung 9.6: Modellselektion für Musteranalyse beim MRT-A. (a) φ - k -Diagramm für Modellparameter $k=2$ bis $k=5$ (b) Consensus-Plots für $k=2$ und $k=3$.



Plots der Consensus-Matrizen für die 2- und 3-Klassen-Modelle sind in Abbildung 9.6 zu sehen und wurden zur Abschätzung der Stabilität der Klasseneinteilung herangezogen. Ein Matrixelement c_{ij} mit $1 \leq i, j \leq N$ enthielt dabei die Häufigkeit der Zuweisung von Person i und j zu ein und demselben Cluster. Im Plot sind Häufigkeiten nahe 1.0 rot und Häufigkeiten nahe 0.0 blau dargestellt. Die Cluster bilden bei einer optimalen Klassifizierung in der grafischen Darstellung, der nach Clusternummer sortierten Consensus-Matrix, deutlich zu erkennende rote Quadrate entlang der Hauptdiagonale. Beide Lösungen ($k=2$ und $k=3$) waren offensichtlich anhand dieser Kriterien akzeptabel.

Als weiteres Kriterium musste die Nützlichkeit der Klassifizierungen zur Differenzierung von Lösungs- und Bearbeitungsstrategien in die Auswahl mit aufgenommen werden. Aus theoretischen Überlegungen waren dies im MRT insbesondere die Unterscheidung von Personen anhand der gelösten Typ-I-Items („Spiegelitems“) und der Items mit Okklusion. Dazu wurden die individuellen Lösungen aus dem MRT-A zu Lösungshäufigkeiten und Ratio-Scores für die fünf Itemtypen des MRT zusammengefasst.

In Varianzanalysen mit der Klasseneinteilung als unabhängige Variable wurde die statistische Signifikanz der Differenzierung überprüft. Das Effektstärkemaß η^2 diente zur Abschätzung des Ausmaßes der Varianzaufklärung durch das Klassenmodell in den itemtypenspezifischen Lösungshäu-

figkeiten und Ratio-Scores.

Tabelle 9.7: Varianzaufklärung in den itemspezifischen Ratio-Scores und Lösungshäufigkeiten durch das Modell mit zwei Klassen.

	Ratio-Scores			Lösungshäufigkeiten		
	F _(1;203)	p	η ²	F _(1;203)	p	η ²
Typ-I	0,06	.806	0.00	31,99	<.001	0.14
Typ-II	8,73	.003	0.04	94,59	<.001	0.32
mit Okkl.	4,18	.042	0.02	49,45	<.001	0.20
ohne Okkl.	1,79	.182	0.01	76,30	<.001	0.27

Diese Differenzierungen der Häufigkeiten und Ratio-Scores sind in Tabelle 9.7 für das Modell mit zwei Klassen vorgenommen worden. Das Modell mit zwei Klassen differenzierte bezüglich der Lösungshäufigkeiten kaum zwischen beiden Klassen. In den Ratio-Scores unterschieden sich die Klassen insbesondere im Typ-II („Strukturitems“) und am wenigsten im Typ-I. Es handelte sich also hauptsächlich um eine Einteilung in Personen, die besonders gut bzw. besonders schlecht Typ-II-Items lösten, also um eine Klasseneinteilung, die für die Identifikation der Strategien anhand der Typ-I-Items eher ungünstig war.

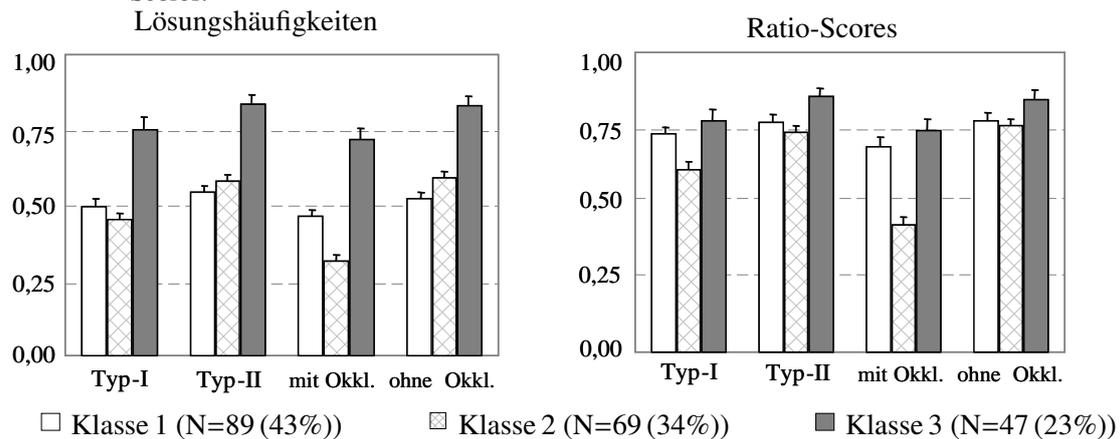
Tabelle 9.8: Varianzaufklärung in den itemspezifischen Ratio-Scores und Lösungshäufigkeiten durch das 3-Klassen-Modell.

	Ratio-Scores			Lösungshäufigkeiten		
	F _(2;202)	p	η ²	F _(2;202)	p	η ²
Typ-I	7,75	.001	0.07	31,41	<.001	0.24
Typ-II	4,64	.011	0.04	36,21	<.001	0.26
mit Okkl.	31,31	<.001	0.24	53,42	<.001	0.35
ohne Okkl.	2,85	.060	0.03	40,57	<.001	0.29

In Tabelle 9.8 sind die Ergebnisse der Analysen zur Varianzaufklärung des Modells mit drei Klassen dargestellt. Dieses Modell differenzierte bezüglich der Häufigkeiten und Ratio-Scores in allen Itemtypen. Insbesondere in den Lösungshäufigkeiten und Ratio-Scores der Items mit Okklusion unterschieden sich die drei identifizierten Klassen. Auch die Differenzierung bei den Häufigkeiten und Ratio-Scores der Typ-I-Items war hier signifikant, so dass sich im folgenden für das Modell mit drei Klassen entschieden wurde.

Zur Interpretation des Modells sind in Abbildung 9.7 die Mittelwerte der drei Klassen bezüglich der untersuchten Itemtypen dargestellt. Auf der linken Seite der Abbildung sind die Lösungshäufigkeiten zu sehen, auf der rechten die Ratio-Scores. Klasse 1, bildete mit 43 % aller Versuchspersonen die zahlenmäßig größte Klasse in der Stichprobe, gefolgt von Klasse 2 mit 34 % und Klasse 3 mit 23 %.

Abbildung 9.7: Interpretation des 3-Klassenmodells durch itemspezifische Lösungshäufigkeiten und Ratio-Scores.



Hinsichtlich der Lösungshäufigkeiten lösten Personen der Klasse 1 im Mittel zirka 50 % der Items von jedem Typ erfolgreich. Besonders hoch war die Häufigkeit für Typ-II-Items (54 %), besonders niedrig für Items mit Okklusion (47 %). Bei Personen der Klasse 2, schwankten die Lösungshäufigkeiten zwischen den Itemtypen stärker. Die mittleren Häufigkeiten in dieser Klasse waren für Items ohne Okklusion (59 %) und für Typ-II-Items (58 %) am höchsten. Die niedrigsten Lösungshäufigkeiten erzielten sie für Items mit Okklusion (32 %). Personen der Klasse 3 lösten mit Abstand die meisten Items jeden Typs. Ihre Lösungshäufigkeiten betragen zwischen 72 % und 84 %. Die niedrigste wurde auch in dieser Subgruppe für Items mit Okklusion ermittelt. Die Signifikanz der beschriebenen Gruppenunterschiede wurde mit Post-hoc-Paarvergleiche (Scheffé-Prozedur) überprüft. Die Mittelwerte der Klasse 3 lagen signifikant über den der Klasse 1 und 2 für alle itemspezifischen Häufigkeiten. Klasse 1 und 2 unterschieden sich nur bezüglich der Lösungshäufigkeiten in den Items mit Okklusion.

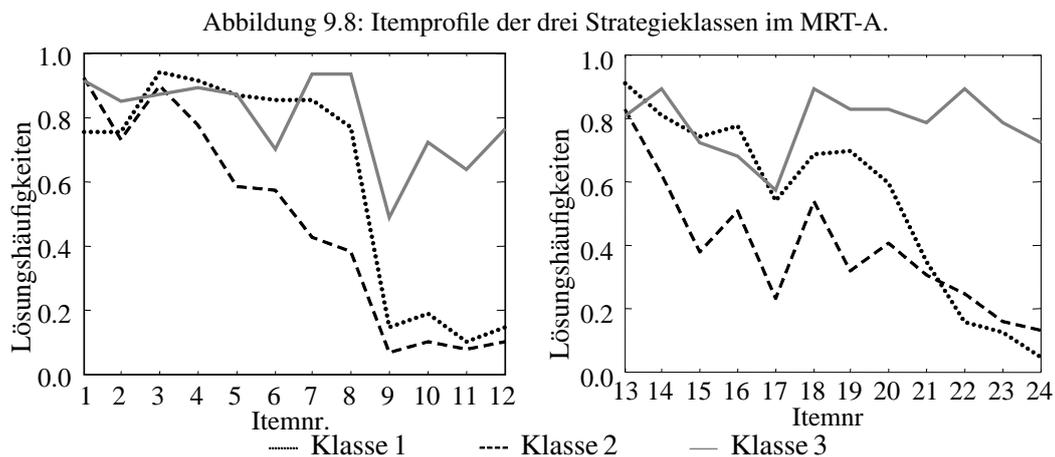
Zur Untersuchung der Lösungsgenauigkeit innerhalb der Klassen werden die in Abbildung 9.7 dargestellten Gruppenmittelwerte bezüglich der Ratio-Scores betrachtet, d. h. die Anzahl richtig gelöster Items im Verhältnis zu den tatsächlich bearbeiteten Items. Die Klasse 1 und 3 erzielten hier über alle Itemtypen die höchsten Mittelwerte. In Klasse 3 lagen diese zwischen 74 % und 85 %, in Klasse 1 ähnlich hoch zwischen 69 % und 77 %. Personen der Klasse 2 zeigten die größten Lösungsunsicherheiten für Items vom Typ-I (60 %) und für Items mit Okklusion (43 %). Post-hoc-Vergleiche ergaben nur für diese Itemtypen signifikant niedrigere Mittelwerte der Klasse 2, sowohl zur Klasse 1 als auch zur Klasse 3. Die Klassen 1 und 3 unterschieden sich hinsichtlich der Ratio-Scores in keinem der Itemtypen signifikant.

Aus den Betrachtungen zu den itemspezifischen Lösungshäufigkeiten und Ratio-Scores innerhalb der identifizierten Klassen ließen sich Schlussfolgerungen hinsichtlich der Bearbeitungsstrategien ableiten. Personen der Klasse 3 lösten offensichtlich unabhängig vom Typ die meisten Items sehr zuverlässig. Da sowohl Typ-I als auch Items mit Okklusion keine Schwierigkeiten bereiteten, waren Personen dieser Subgruppe wohl am ehesten in der Lage, schnell und exakt mentale Rotationen auszuführen. Klasse 1 zeichnet sich durch eine ähnliche Exaktheit bei der Lösung dieser Itemtypen aus

wie die Klasse 3. Jedoch bearbeiten Personen der Klasse 1 signifikant weniger Items. Vermutlich handelt es sich hier um Personen, die mentale Rotationen zur Bearbeitung der Aufgaben anwenden, jedoch viel langsamer als dies Personen der Klasse 3 taten.

In Abbildung 9.8 sind zur Untermauerung dieser Behauptung, die Itemprofile der Lösungshäufigkeiten über alle 24 Items für die 3 Klassen dargestellt. Personen der Klasse 3 (Volllinie) arbeiten in beiden Testhälften mit relativ konstanten Häufigkeiten die Items des Test ab. Auch das Profil der Klasse 1 (Punktlinie) folgt diesem Bearbeitungsmodus weitgehend, jedoch nur für die jeweils ersten acht Items jeder Testhälfte. Ab dem neunten bzw. 20. Item fallen die Lösungshäufigkeiten bis auf unter 20 % ab, was vermutlich darauf zurückzuführen war, dass durch das Erreichen des Zeitlimits die nachfolgenden Items nicht mehr bearbeitet werden konnten.

Klasse 2 fiel insgesamt durch stärkere Unsicherheit genau bei den Itemtypen auf, die mentale Rotation erfordern. Da sie aber Items, die leicht durch Detailvergleiche gelöst werden können, richtig bearbeiteten, waren diese Personen vermutlich diejenigen, die eher analytische Bearbeitungsstrategien im Test einsetzten. Im Itemprofil dieser Klasse (gestrichelte Linie) wird dieses Verhalten durch häufige Zacken zwischen zwei Items deutlich (z. B. Item 15 und Item 19).



Die Auswirkung unterschiedlicher Bearbeitungsstrategien auf die Gesamtleistung im MRT wurde anhand der Summenscores, der Anzahl der bearbeiteten Items und dem daraus gebildeten Ratio-Score für den gesamten Test untersucht. In Tabelle 9.9 sind die Gruppenmittelwerte der drei Leistungsmaße für die Strategieklassen dargestellt. Die höchste Gesamtpunktzahl von im Mittel 19 Punkten erzielten Personen der Klasse 3. Die Klassen 1 und 2 unterschieden sich trotz der wahrscheinlich grundverschiedenen Lösungsansätze nicht in den erzielten Summenscores. Bezüglich der Gruppenvarianzen bestanden geringfügige Unterschiede (Levene-Test: $F=2,00$; $p=.137$). Insgesamt konnten in der Stichprobe durch den Faktor „Strategie“ 32 % der Varianz aufgeklärt werden.

Hinsichtlich der Anzahl der bearbeiteten Items und demzufolge auch hinsichtlich der Ratio-Scores lagen signifikante Mittelwertsunterschiede zwischen allen drei Klassen vor. Personen der Klasse 1 bearbeiteten im Mittel 23 Items mit $SD=1,3$; das sind fast alle Items des Tests. Auf Grund der geringen Streuung dieses Merkmals innerhalb der Klasse 1 ergab der Levene-Test eine deutliche Abweichung von der Varianzhomogenität zwischen den Gruppen ($F=37,05$; $p<.001$), so dass die

Tabelle 9.9: Gruppenmittelwerte und Varianzaufklärung der Gesamtleistung im MRT-A durch die drei Strategieklassen.

	Klasse 1 (N=89; 43%)		Klasse 2 (N=69; 34%)		Klasse 3 (N=47; 23%)		$F_{(2;202)}$	p	η^2
	M	SD	M	SD	M	SD			
Summenscore	12,1	4,34	12,0	3,56	19,0	5,22	47,58	<.001	0.32
Anzahl bearbeiteter Items	16,4	4,33	18,7	4,06	23,4	1,24	53,89	<.001	0.35
Ratio-Score	0,74	0,20	0,65	0,17	0,81	0,21	9,93	<.001	0.09

Mittelwertsunterschiede in diesem Fall noch einmal mit einer robusteren Statistik überprüft wurden (Welch-Test: $F(2;123,2)=127,05; p<.001$). Die Varianzen in den Klassen bezüglich der Ratio-Scores unterschieden sich nicht signifikant ($F=1,15; p=.317$).

Zur weiteren Beurteilung der entdeckten Strategieklassen wurden korrelativ Zusammenhänge zwischen MRT-Leistung und weiteren zur Verfügung stehend Fähigkeitsmaßen innerhalb der einzelnen Klassen berechnet. Dazu wurden Korrelationskoeffizienten nach Pearson für den Summenscore aus dem MRT-A, den Aufgabengruppen *Reihen* („induktive Denkfähigkeit“) und der Aufgabe *Flächenzählen* (räumliche Vorstellung) berechnet. In Tabelle 9.10 sind die Koeffizienten getrennt für die drei Klassen dargestellt.

Tabelle 9.10: Zusammenhänge MRT-Leistung und induktives Schließen partialisiert nach Strategien.

	Klasse 1 (N=89)			Klasse 2 (N=69)			Klasse 3 (N=47)			Gesamt (N=205)		
	1.	2.	3.	1.	2.	3.	1.	2.	3.	1.	2.	3.
1. MRT-A	-	.195 ⁺	.275 ^{**}	-	.361 ^{**}	.219 ⁺	-	.208	-.02	-	.355 ^{**}	.278 ^{**}
2. Reihen		-	.326 ^{**}		-	.445 ^{**}		-	.366 [*]		-	.428 ^{**}
3. Flächenzählen			-			-			-			-

⁺ $p<0.1$, ^{*} $p<.05$, ^{**} $p<0.01$

Nur in der Klasse 2, die mit eher analytischen Strategien identifiziert wurde, korrelierte der MRT-Summenscore signifikant mit dem Summenscore aus der Aufgabengruppe *Reihen*. Die Korrelation zwischen diesen Fähigkeiten, die in der Gesamtstichprobe gefunden wurde, erklärte sich hauptsächlich durch diese Subgruppe. Insgesamt korrelierte in dieser Klasse im Vergleich zu den beiden anderen Klassen das Leistungsmaß für die Facette „induktive Denkfähigkeit“ am höchsten mit den Raumvorstellungsmaßen. Ein signifikanter Zusammenhang zwischen MRT und der Raumvorstellungsaufgabe aus dem PSB-R bestand nur in der Klasse 1. In der Klasse 3 korrelierte keines der Maße aus dem PSB-R mit einem signifikanten Koeffizienten, was auf die hohe Leistungsdichte innerhalb dieser Klasse und den damit verbunden Deckeneffekten in den Summenscores aller drei Tests zurückzuführen war.

Zur Bestimmung des Zusammenhangs zwischen der Strategiewahl im MRT und dem Geschlecht

wurde die Verteilung männlicher und weiblicher Versuchspersonen auf die drei Klassen untersucht. Tabelle 9.11 stellt die entsprechenden beobachteten Häufigkeiten auf den entsprechenden Faktorstufen in einer Kontingenztafel dar. Zur besseren Interpretierbarkeit wurden in der Tabelle die Klassen zusätzlich mit Bezeichnungen versehen, die das wesentliche Vorgehen der Personen beschreiben.

Tabelle 9.11: Zusammenhang Klasseneinteilung und Geschlechtszugehörigkeit.

	Klasse 1 Langsames Rotieren	Klasse 2 Detail- analysen	Klasse 3 Schnelles Rotieren	Gesamt
weiblich	31 (52%)	22 (37%)	7 (11%)	60 (100%)
männlich	58 (40%)	47 (32%)	40 (28%)	145 (100%)
korrigiert-standardisierte Residuen	1,53	0,59	-2,47	

In der Stichprobe wurden weibliche Versuchspersonen am häufigsten der Klasse der Personen zugeordnet, die vermutlich mentale Rotationen, aber in langsamer Weise, ausführten. Mehr als die Hälfte der Mädchen befanden sich in dieser Subgruppe. Der Rest befand sich zum größten Teil (37 %) in der Klasse der Personen, die die Aufgabe eher durch Detailvergleiche lösten. Am seltensten waren Mädchen in der im Test erfolgreichsten Klasse 3, deren Personen offensichtlich keine Schwierigkeiten in der Vorstellung schneller mentaler Rotation hatten.

Die Verteilung der Jungen auf die drei Strategieklassen wich nominal in allen Fällen etwas von der der Mädchen ab. Insgesamt waren sie prozentual weniger in den Klassen „Langsames Rotieren“ und „Detailanalyse“ und mehr in der Klasse „Schnelles Rotieren“ vertreten. Zur statistischen Absicherung der nominalen Unterschiede auf den einzelnen Faktorstufen wurden korrigiert-standardisierte Residuen berechnet. Werte, die im Betrag größer als 1.96 sind, zeigen eine auf dem 0.05%-Niveau signifikante Abweichung an. Statistisch signifikant war demzufolge nur die überproportionale Über- bzw. Unterbesetzung der Klasse „Schnelles Rotieren“ mit männlichen bzw. weiblichen Versuchspersonen.

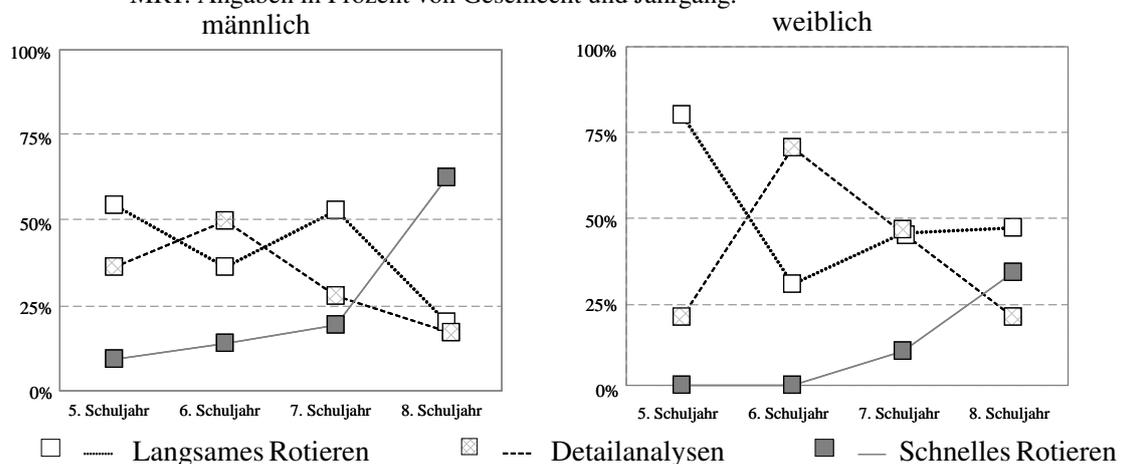
Insgesamt stellte sich der Zusammenhang zwischen Geschlecht und Strategie wegen der statistischen Gleichverteilung von Jungen und Mädchen auf die Klassen 1 und 2 eher moderat dar ($\chi^2=6,25$; $df=2$; $p=.044$). Der korrigierte Kontingenzkoeffizient betrug $C_{\text{kor}}=0.243$. Jedoch ist auch der eher moderate Zusammenhang ein Problem, wenn man mit varianzanalytischen Methoden den Einfluss beider Faktoren („Strategie“ und „Geschlecht“) auf die MRT-Leistung untersucht.

So konnten durch ein lineare Modell mit Strategie und Geschlecht als feste Faktoren in der ANOVA über 40 % der Varianz in den MRT-Summenscores aufgeklärt werden (korrigiertes $R^2=0.402$). Jeder der Haupteffekte war in diesem Modell signifikant (Geschlecht: $F(1,203)=32,35$; $p<.001$; $\eta^2=0.140$; Strategie: $F(2,203)=14,43$; $p<.001$; $\eta^2=0.127$). Bei der ebenfalls signifikanten Interaktion beider Faktor handelte es sich um eine ordinale Interaktion ($F(2,203)=3,91$; $p=.022$). Eine getrennte Interpretation der Haupteffekte war somit möglich. Männliche Versuchspersonen erzielten gegenüber weiblichen einen höheren MRT-Score, Personen welche die Strategie „Schnelles Rotieren“ anwendeten

erzielten gegenüber Personen der beiden anderen Strategieklassen die höchsten Summenscores. Die Interaktion beider Faktoren resultiert aus einer höheren Wirksamkeit der Strategie „Schnelles Rotieren“ in der Gruppe männlicher Versuchspersonen auf die unabhängige Variable.

Wie sich die im MRT identifizierten Strategien über die Mädchen und Jungen der einzelnen Schuljahrgänge verteilten, stellt Abbildung 9.9 dar. Dazu wurden prozentuale Häufigkeiten jeweils von den Schülerinnen bzw. Schülern einer Jahrgangsstufe gebildet. Aus der Abbildung wird deutlich, dass in den Schulklassen 5. bis 7. die wenigsten Schüler der Strategie „Schnelles Rotieren“ (grau) identifiziert wurden. Das galt sowohl für die Mädchen als auch für die Jungen. In den Schulklassen fünf und sechs waren überhaupt keine Mädchen vertreten, die diese Lösungsstrategie anwendeten. Analytische (schraffiert) oder die als „Langsames Rotieren“ (weiß) bezeichneten Strategien überwogen in diesem Alter. Letztere wurde bei fast allen Mädchen (80 %) der fünften Schulklasse festgestellt.

Abbildung 9.9: Verteilung der Schüler bzw. Schülerinnen einer Schuljahrgangsstufe auf die Strategien im MRT. Angaben in Prozent von Geschlecht und Jahrgang.



Ansonsten wechselten die Häufigkeiten für beide Arten von Strategien zwischen den Schuljahrgangsstufen und Geschlechtern sehr stark. Für die Strategiegruppe „Schnelles Rotieren“ konnte dagegen über die Jahrgangsstufen ein stetiger Anstieg der Häufigkeiten beobachtet werden. Insbesondere Jungen des achten Schuljahrs zeigten im MRT Lösungsmuster, die auf die Anwendung dieser Strategie schließen ließen. Deren Anzahl nahm im Vergleich zu den vorangegangenen Jahrgangsstufen von unter 20 % auf 63 % zu. Auch für die Mädchen war dieser Sprung in den Häufigkeiten zu beobachten, wenn auch in geringerem Umfang: 33 % der Mädchen des achten Schuljahrs wendeten diese Strategie im MRT an. Am häufigsten aber waren diese mit den Strategien „Langsames Rotieren“ vertreten (47 %).

Insgesamt waren die statistischen Zusammenhänge zwischen Bearbeitungsstrategie im MRT und dem Alter der Schüler, ausgedrückt durch Jahrgangsstufen signifikant. Für die Schülerinnen betrug der korrigierte Kontingenzkoeffizient $C_{\text{kor}}=0.682$ ($\chi^2=18,17$; $df=6$; $p=.006$), für die Schüler $C_{\text{kor}}=0.651$ ($N=145$; $\chi^2=39,01$; $df=6$; $p<.001$). Die stärksten Unterschiede zwischen Jungen und Mädchen wurden in den Jahrgangsstufen fünfte und achte in der Besetzung der Strategieklassen 1

und 3 entdeckt. Die korrigierten Residuen zeigten hier signifikant mehr Mädchen in der Kategorie „Langsames Rotieren“ und mehr Jungen in der Kategorie „Schnelles Rotieren“ an.

9.4.2 Würfelaufgaben

Als weitere Strategieberichtung durch Mustererkennung wurden die Items aus dem Test *Würfelaufgaben* ausgewertet. Nach Dichotomisierung der 20 Itemantworten jeder Person wurden diese zu einer $M \times N$ Datenmatrix ($M=20$ Items und $N=205$ Versuchspersonen) zusammengestellt. Diese Datenmatrix war Eingabe für das Consensus-Clustering. Parameter für die im Clustering verwendete NMF-Heuristik waren $N=800$ Iteration und $\lambda=0.1$. Es wurden Modellparameter für $k=2$ bis $k=5$ in jeweils 20 Durchläufen während der Cluster-Prozedur erprobt.

In Tabelle 9.12 sind die zur Modellselektion herangezogenen Kriterien für die berechneten Modelle zusammengefasst. Die kophenetischen Korrelationskoeffizienten und relativen Residuen dienten zunächst zur Auswahl günstiger Faktorisierungen der Datenmatrix und wurden aus dem Consensus-Clustering gewonnen. Bezüglich dieser Maße wurde ein Modell mit zwei und ein Modell mit drei Klassen für günstig befunden. Für diese beiden Modelle sind in der Tabelle die Kriterien zur Differenzierung anhand der im Test identifizierten Itemtypen mit angegeben. Es wurden die zu Mittelwerten zusammengefassten itemspezifischen Lösungshäufigkeiten verwendet. In der Tabelle werden die Anteile der Varianzaufklärung durch die Klassifizierung anhand des Effektstärkemaß η^2 und des Signifikanzniveaus der Mittelwertdifferenzen dargestellt. Für die Entscheidung zwischen beiden Modellen war aus theoretischen Überlegungen eine ausreichende Differenzierung der Modelle anhand der Lösungshäufigkeiten für „Raumwürfel“ entscheidend. Das Modell mit drei Klassen war hier mit $\eta^2=0.17$ dem 2-Klassen-Modell überlegen.

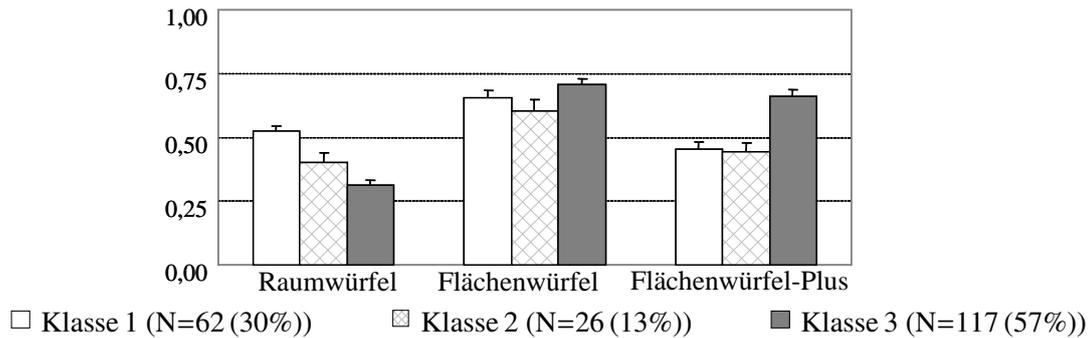
Tabelle 9.12: Modellparameter k für die Klassifizierung der Itemantwortmuster aus den *Würfelaufgaben*. φ_k : kophenet. Korr.koeff., rel. Res.: relative Residuen der SNMF und Varianzaufklärung des Modells bezüglich der itemspezifischen Lösungshäufigkeiten.

k	φ_k	rel. Res.	Raumwürfel		Flächenwürfel		Flächenwürfel-Plus	
			η^2	p	η^2	p	η^2	p
2	0,975	0,22	0.05	.001	0.09	<.001	0.08	<.001
3	0,980	0,19	0.17	<.001	0.02	.095	0.14	<.001
4	0,953	0,19	-	-	-	-	-	-
5	0,950	0,18	-	-	-	-	-	-

In Abbildung 9.10 sind die Lösungshäufigkeiten für die drei Itemtypen in den identifizierten Klassen grafisch dargestellt. Die Ratio-Scores werden hier nicht berichtet, da sie sich von den Lösungshäufigkeiten kaum unterscheiden und deshalb keine zusätzlichen Informationen zur Interpretation der Klassen beitragen konnten. In Klasse 1 waren 30 % aller Versuchspersonen, Klasse 2 bildete mit nur 13 % die kleinste Klasse und Klasse 3 mit 57 % die zahlenmäßig größte Subgruppe.

Insgesamt wurden „Raumwürfeln“ im Mittel in den drei Klassen seltener richtig bearbeitet als

Abbildung 9.10: Itemspezifische Lösungshäufigkeiten für das 3-Klassen-Modell der Würfelaufgaben.



andere Aufgabentypen. Nur Klasse 1 erreichte eine Häufigkeit von über 50 %. In Klasse 2 und 3 waren es nur 40 % bzw. sogar nur 31 %. Damit schnitten Personen der Klasse 3 bei diesem Itemtyp am schlechtesten ab. „Flächenwürfel“ wurden dagegen deutlich häufiger richtig beantwortet (60 % – 71 %). Bei diesem Itemtyp schnitten Personen der Klasse 3 mit 71 % am besten ab; genauso bei Aufgaben mit „Flächenwürfel-Plus“. Hier lösten Personen der Klasse 3 mit einer Häufigkeit von 66 % die meisten Items richtig. In den Klassen 1 und 2 lagen die Häufigkeiten nur um 45 %.

Mittelwertsunterschiede zwischen den einzelnen Klassen wurden durch Post-hoc-Paarvergleiche (Scheffé-Prozedur) auf Signifikanz geprüft. Personen der Klasse 1 lösten signifikant mehr „Raumwürfel“ als die Personen der Klasse 3. Aber in der Klasse 3 wurden signifikant mehr „Flächenwürfel“ und „Flächenwürfel-Plus“ gelöst als in den Klassen 1 und 2. Die in der Abbildung 9.10 ersichtlichen Mittelwertsunterschiede zwischen Klasse 1 und 2 bei „Raumwürfeln“ waren dagegen nur tendenziell ($p=.067$).

Zur Ableitung der Bearbeitungsstrategien aus diesen Resultaten lässt sich also feststellen, dass insbesondere zwei Subgruppen identifiziert wurden, die konträre Antwortmuster für „Raumwürfel“ und „Flächenwürfel“ aufweisen. Während in Klasse 3 vorwiegend „Flächenwürfel“ und „Flächenwürfel-Plus“, also Items, die eher analytische Fähigkeiten erfordern, gelöst wurden, dominierte Klasse 1 in den Lösungen für „Raumwürfel“, also für Items, die ohne mentales Rotieren nicht gelöst werden können. Personen einer weiteren Klasse, die zahlenmäßig kleinste, zeigten ein ähnliches Verhaltensmuster wie Personen der Klasse 1, jedoch nicht so stark ausgeprägt für die „Raumwürfel“.

Zur Auswirkung der Bearbeitungsstrategien auf die Gesamtleistung in den *Würfelaufgaben* wurden Summenscores und die Anzahl der bearbeiteten Items zwischen den drei Klassen verglichen. In Tabelle 9.9 sind Gruppenmittelwerte und Prüfgrößen dargestellt. Die in den Klassen gemittelten Summenscores unterschieden sich nicht und lagen jeweils zwischen zehn und elf Punkten. Es bestanden keine Varianzunterschiede zwischen den Klassen in diesem Leistungsmaß (Levene-Test: $F=1,45$; $p=.237$).

In der Anzahl der bearbeiteter Items unterschieden sich die Gruppen geringfügig sowohl in den Mittelwerten als auch in den Varianzen (Levene-Test: $F=29,57$; $p<.001$). Zur Absicherung der statistischen Signifikanz der Mittelwertdifferenzen wurde deshalb noch der robustere Welche-Test durch-

Tabelle 9.13: Mittelwertsunterschiede zwischen den Strategieklassen bezüglich der Gesamtleistung in den *Würfelaufgaben*.

	Klasse 1 (N=62; 30%)		Klasse 2 (N=26; 13%)		Klasse 3 (N=117; 57%)		F	p	η^2
	M	SD	M	SD	M	SD			
Summenscore	11,2	3,58	10,1	3,44	11,3	4,09	1,07	.345	0.01
Anzahl bearbeiteter Items	19,7	0,70	19,8	0,46	18,4	3,42	7,32	.001	0.07

geführt ($F(2;113,6)=10,12; p<.001$).

Zur Bestimmung des Zusammenhangs zwischen den im Test angewendeten Lösungsstrategien und der Geschlechtszugehörigkeit sind in Tabelle 9.14 die beobachteten Häufigkeiten in den entsprechenden Kategorien beider Faktoren zusammengefasst. Nominal ergaben sich nur geringfügige Unterschiede in den Besetzungen der Klassen. Mädchen waren prozentual etwas häufiger in der Klasse 2 und seltener in Klasse 3. Auf die Klasse 1, also die Klasse mit besseren Leistungen für „Raumwürfel“, verteilten sich beide Geschlechter prozentual gleich stark. Anhand der Gegenüberstellung der Kategorien Geschlecht und Strategie konnte hier also kein signifikanter Zusammenhang festgestellt werden ($\chi^2=1,55; df=2; p=.460$).

Tabelle 9.14: Zusammenhang Geschlechtszugehörigkeit und Strategien in den *Würfelaufgaben*.

	Klasse 1 - Raumwürfel	Klasse 2	Klasse 3 - Flächen- Plus	Gesamt
weiblich	19 (32%)	10 (17%)	31 (51%)	60 (100%)
männlich	43 (30%)	16 (11%)	86 (59%)	145 (100%)
korrigiert-standardisierte Residuen	0,29	1,10	-1,01	

9.4.3 Zusammenhänge zwischen den Klasseneinteilungen

Zur Klärung der Fragestellung, ob die Strategiewahl in einem Raumvorstellungstest materialspezifische oder eine invariante Personeneigenschaft ist, wurden korrelativ Zusammenhänge zwischen den Klassifikationen aus MRT und *Würfelaufgaben* untersucht. Da auf Grund der durch die Kreuztabellierung entstanden hohen Anzahl von Kategorien im χ^2 -Test mehr als 20 % der Felder eine erwartete Häufigkeit kleiner als 5 aufwiesen, wurden Klassen aus dem MRT zusammengefasst, um so die Kategorienzahl wieder etwas zu verringern. Sinnvoll schien hier eine Zusammenfassung der Klassen 1 und 3 zu einer Klasse, die eher holistische Strategien im Test einsetzte. Die Klasse 2 wurde, auf Grund der im Abschnitt 9.4.1 identifizierten Eigenschaften, mit eher analytischen Bearbeitungsstrategien in Verbindung gebracht. Die durchgeführte dichotome Klassifizierung entspricht so in direkter Weise dem vorgeschlagenen Konstrukt der räumlichen Präferenz.

Tabelle 9.15 zeigt die Verteilung der holistisch bzw. analytisch im MRT arbeitenden Versuchspersonen auf die Klasseneinteilung aus den *Würfelauflagen*. Nominal ergab sich ein unerwartetes Ergebnis, indem prozentual etwas mehr analytisch arbeitende Person in einer Klasse vertreten waren, die besonders gut „Raumwürfel“ lösen konnten. Holistisch arbeitende Personen waren etwas häufiger in der Klassen 2, die zwar mit einem ähnlichen Profil wie Klasse 1 identifiziert wurden, jedoch bei „Raumwürfeln“ etwas schlechter abschnitten. Insgesamt konnte jedoch auch hier keine signifikant überproportionalen Häufigkeiten in einzelnen Klassen beobachtet werden ($\chi^2=2,61$; $df=2$; $p=.271$). Mit dem korrigierten Kontingenzkoeffizienten von $C_{\text{kor}}=0.294$ kann höchstens von einem geringen Zusammenhang ausgegangen werden.

Tabelle 9.15: Verteilung der im MRT-A eher holistisch bzw. analytisch arbeitenden Versuchspersonen auf die Strategieklassen in den *Würfelauflagen*.

	Klasse 1 - Raumwürfel	Klasse 2	Klasse 3 - Flächen- Plus	Gesamt
analytisch	25 (36%)	6 (9%)	38 (55%)	69 (100%)
holistisch	37 (27%)	20 (15%)	79 (58%)	136 (100%)
korrigiert-standardisierte Residuen	1,33	-1,22	-0,41	

An dieser Stelle wurde noch die Möglichkeit des Einflusses der Bearbeitungsstrategien auf das Prä-Post-Test-Design zum Vergleich der Papier-Bleistift- und der Computerversion des MRT-B untersucht. Adaptiv wurde im Laufe der Untersuchung die Frage aufgeworfen, dass sich möglicherweise die parallelisierten Gruppen in den im Vortest angewandten Strategien unterscheiden und sich dies auf die Leistungen im Computer bzw. Papier-Bleistift-Test auswirkte.

Getrennt nach Geschlecht wurden deshalb die Verteilungen der Personen des Computer- bzw. Papier-Bleistift-Tests auf die Strategiekategorie, die aus den Itemmustern des Vortests (MRT-A) ermittelt wurden, bestimmt. Sowohl für die Stichprobe der Schülerinnen als auch der Schüler ergaben sich keine signifikanten Abweichungen von den Erwartungswerten (weiblich: $N=60$; $\chi^2=1,81$; $df=2$; $p=.404$; männlich: $N=145$; $\chi^2=0,30$; $df=2$; $p=.861$). D. h. durch die für Alter und Geschlecht getrennte Parallelisierung nach dem MRT-Summenscore, wurde weitgehend auch eine Gleichverteilung der Strategien auf die beiden Gruppen Computertest und Papier-Bleistift-Test erreicht.

9.5 Zusammenhänge zu Strategiefragebögen

Aus dem Fragebogen zur Selbsteinschätzung der im MRT angewandten Bearbeitungsstrategien wurden die ersten zwei Fragen auf Itemebene ausgewertet. Daraus ging hervor, dass 73 % der Schüler, 75 % der Studenten und 85 % Senioren Spaß an der Bearbeitung des MRT hatten. Nur 8 % der Schüler gaben an, dass ihnen die Aufgaben schwer fielen. Bereits 33 % der Studenten und der überwiegende Teil der älteren Versuchspersonen (65 %) beantwortet diese Frage mit „Ja“.

Die folgenden 22 Fragen zu den Strategien wurden in fünf Kategorien zusammengefasst. Die erste Kategorie unterscheidet zwischen den Bearbeitungsmodi innerhalb eines MRT-Items (Frage 3 bis 8). Es konnte unterschieden werden in Personen, die strikt in Reihenfolge die Ausgangsfiguren mit den vier Alternativfiguren verglichen, Personen, die zwar systematisch vorgingen, aber auch Vergleiche zwischen Alternativfiguren selbst durchführten, und Personen, die angaben, bei der Auswahl keine Systematik angewandt zu haben.

Die zweite Kategorie unterteilte die Angaben der Versuchspersonen in eher holistische oder eher analytischen Strategien (Frage 9 bis 14). Die dritte Kategorie erfasst, ob Hilfsmittel oder die Hände zur Simulation der Rotation eingesetzt wurden (Frage 15 und 17). Die vierte Kategorie bestimmt, ob während des Tests eine feste Strategie konsequent eingesetzt wurde (Frage 18 bis 20). Die letzte Kategorie ermittelt, ob Personen sich der Richtigkeit ihrer Antworten versicherten oder eher rieten, um das Zeitlimit einzuhalten (Frage 21 bis 24).

Die deskriptiven Häufigkeiten über die Altersgruppen bezüglich dieser fünf Kategorien sind in Tabelle 9.16 dargestellt. In der Kategorie Auswahl der Alternativfiguren gaben in jeder Altersgruppe mehr als die Hälfte der Versuchspersonen an, systematisch vorgegangen zu sein. Die beiden Extreme unsystematisches bzw. striktes Abarbeiten wurde von Schülern und Studenten mit gleicher Häufigkeit berichtet. In der Gruppe der Senioren gab es allerdings nur eine Versuchsperson, die, laut eigenen Angaben, die ganz strikte Variante des Bearbeitungsmodus einsetzte.

Zur Differenzierung der holistischen bzw. analytischen Strategien gaben Schüler in fast der Hälfte der Fälle an holistische Strategien benutzt zu haben. Studenten und Senioren gaben häufiger analytische als holistische Strategien an. Hier spielt die Verteilung von Jungen und Mädchen in der Schülerstichprobe eine entscheidende Rolle und wird später noch mit berücksichtigt. Nichtsdestotrotz konnten durch die Antworten im Fragebogen in jeder der Altersgruppen ca. 30 % der Versuchspersonen keiner der beiden Ausprägungen einer Strategie zugeordnet werden.

Aus der Kategorie 3 wurde deutlich, dass die meisten Versuchspersonen laut eigenen Angaben keine Hilfsmittel oder die Hände zur Simulation der Lösung benutzten. Bei den Angaben zur Kategorie 4 ergab sich, dass die meisten Personen Strategien im MRT wechselnden oder ihr Vorgehen nicht genau angeben konnten. Durch die Fragen der Kategorie 5 wurde ermittelt, dass nur sehr wenige Personen bewusst einige Antworten geraten hatten; über die Hälfte der Personen legten den Schwerpunkt auf Richtigkeit der Lösungen.

Zusammenhänge zwischen Selbsteinschätzung und den Faktoren „Geschlecht“, „Testversion“ und den automatisch ermittelten Strategieklassen wurden gleichermaßen mit kategorialen Zusammenhangskoeffizienten untersucht. Dazu wurde die Altersstichprobe der Schüler ausgewählt, da nur hier die unabhängigen Variablen in allen zu untersuchenden Ausprägungen vorlagen.

Bezüglich des Geschlechts konnten in den Kategorien 1 bis 4 keine signifikanten Abweichungen von den Erwartungswerten gefunden werden. Insbesondere differenzierte die Kategorie der holistischen vs. analytischen Strategien nicht zwischen Jungen und Mädchen ($\chi^2=0,60$; $df=2$; $p=.742$). Nur bezüglich der Angaben zur Genauigkeit, mit der der MRT bearbeitet wurde, bestand ein Unter-

Tabelle 9.16: Häufigkeiten in den fünf Kategorien des Strategiefragebogens zum MRT.

	Schüler N=205	Studenten N=36	Senioren N=20
<i>1. Auswahl der Alternativfiguren</i>			
strikt	42 (20%)	9 (25%)	1 (5%)
systematisch	118 (58%)	19 (53%)	13 (65%)
unsystematisch	45 (22%)	8 (22%)	6 (30%)
<i>2. Holistisch vs. analytisch</i>			
holistisch	94 (47%)	9 (25%)	4 (20%)
analytisch	53 (26%)	16 (44%)	9 (45%)
unbekannt	54 (27%)	11 (31%)	7 (35%)
<i>3. Hilfsmittel oder Hände zur Simulation</i>			
mit	6 (3%)	6 (17%)	2 (10%)
ohne	173 (85%)	28 (78%)	10 (50%)
unbekannt	25 (12%)	2 (6%)	8 (40%)
<i>4. Feste Strategie</i>			
fest	62 (31%)	12 (33%)	1 (5%)
wechselnd	59 (29%)	15 (42%)	10 (50%)
unbekannt	80 (40%)	9 (25%)	9 (45%)
<i>5. Richtigkeit der Antwort</i>			
richtig	113 (57%)	24 (67%)	12 (63%)
geraten	18 (9%)	-	2 (11%)
unbekannt	68 (34%)	12 (33%)	5 (26%)

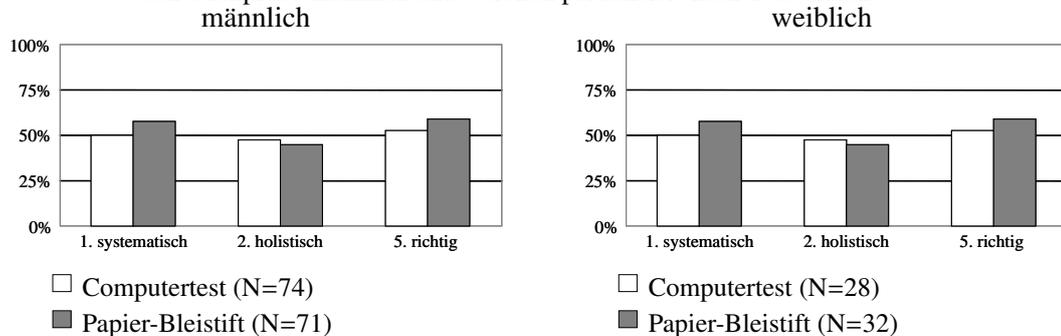
schied in den Häufigkeiten ($\chi^2=6,32$; $df=2$; $p=.042$). 17 % der Mädchen gaben hier gegenüber nur 6 % der Jungen an, im Test geraten zu haben.

Statistisch signifikante Zusammenhänge zwischen den durch die automatische Mustererkennung gewonnen Klasseneinteilung und den Kategorien des Fragebogens bestanden ebenfalls nicht. Hauptaugenmerk galt hier der Kategorie 2, in der sich aber die Personen nahe zu gleich auf die Klassen der automatischen Analyse verteilten ($\chi^2=1,77$; $df=4$; $p=.777$).

Mit dem Faktor „Testversion“ wurde untersucht, ob sich Personen, die den Computertest des MRT absolvierten, sich in ihren Angaben zu den verwendeten Strategien von den Personen unterschieden, die den Papier-Bleistift-Test durchführten. Tendenziell mehr Personen aus dem Computertest konnten nicht genau angeben, ob sie eine feste oder mehrere sich abwechselnde Strategien bei der Bearbeitung einsetzten ($\chi^2=5,38$; $df=2$; $p=.068$). Nominal waren das 48 % im Computertest zu 32 % im Papier-Bleistift-Test. In allen anderen Fragebogenkategorien wurden keine signifikanten Unterschiede gefunden.

Da sich in der Durchführung des MRT als Computertest oder Papier-Bleistift-Test Interaktionen mit dem Faktor „Geschlecht“ zeigten, wurde diese auch in der Betrachtung der Strategien berücksichtigt. Abbildung 9.11 zeigt drei der Kategorien des Fragebogens, in denen zumindest tendenziell signifikant unterschiedliche Aussagen in Abhängigkeit von der bearbeiteten Testversionen gegeben wurden. Es sind nur die Stufen der Kategorien angegeben, auf denen Unterscheide bestanden.

Abbildung 9.11: Häufigkeiten für ausgewählte Stufen der Kategorien 1, 2 und 5 des Strategiefragebogens in den Stichprobe männlich und weiblich partialisiert nach Testversion.



In der Stichprobe der Jungen gab es keine Unterschiede in den Antworten des Strategiefragebogens, egal, ob sie ihn nach dem Computer- oder dem Papier-Bleistift-Test bearbeiteten. Mädchen unterschieden sich jedoch in den von ihnen beschriebenen Vorgehensweisen je nach Testversion. Die Computertestgruppe gab häufiger an, nur systematisch vorgegangen zu sein, statt in einer strikten Reihenfolge (nicht in der Abbildung angegeben). Der deutlichster Unterschied bestand in der Angabe zur Anwendung holistischer bzw. analytischer Strategien. Während 57 % der Mädchen im Papier-Bleistift-Test holistische Strategien angaben, waren es im Computertest nur 34 %. Außerdem legten die Mädchen im Computertest seltener Wert auf die Richtigkeit der Lösung und neigten stärker zum Raten.

9.6 Zusammenfassung

Ein anhand von Gruppenmittelwerten durchgeführtes Screening der Stichproben ergab bezüglich Schulleistung bzw. Schulabschluss, der „induktiven Denkfähigkeit“ und einem einfachen Test zur Raumvorstellung eine weitgehende Vergleichbarkeit der getesteten männlichen und weiblichen Versuchspersonen. In der Vorerfahrung im Umgang mit Computerprogrammen und -spielen differierten Alters- und Geschlechtsgruppen. Das wurde besonders deutlich bei Nutzung von Computerspielen. Die überwiegende Zahl der Senioren und der weiblichen Versuchspersonen (80% – 90%) spielen so gut wie nie am Computer, jüngere männliche (Schüler und Studenten) dagegen laut eignen Angaben fast täglich.

Für den MRT zeigten sich die für diesen Test typischen großen Geschlechterdifferenzen in allen drei Altersgruppen. Männliche Versuchspersonen schnitten im Mitteln in allen Gruppen besser ab als weibliche. Der Effekt lag in einer Größenordnung zwischen $d=0.8$ und $d=2.0$ (Studenten). Für Schüler wurden Geschlechterdifferenzen in Abhängigkeit von der jeweiligen Schuljahresstufen be-

trachtet. Zwischen dem fünften bis siebten Schuljahr blieben die Differenzen relativ konstant, mit einem mittleren Effekt um $d=0.7$. In der Jahrgangsstufe acht wuchs der Geschlechtseffekt auf $d=1.5$.

Für die *Würfelaufgaben* wurden keine signifikanten Geschlechtsdifferenzen in den drei Altersstichproben nachgewiesen. Dennoch bestanden mittlere Zusammenhänge zwischen den Leistungen in diesem Test und dem MRT. Diese variierten geschlechtsspezifisch. In einer linearen Regression mit dem MRT-Summscore als Prädiktorvariable konnten ca. 22 % der Varianz für männliche und 11 % für weibliche Versuchspersonen in den Daten der *Würfelaufgaben* aufgeklärt werden.

In einem Prä-Post-Test wurde die Vergleichbarkeit des Computertest des MRT mit der ursprünglichen Papier-Bleistift-Version auf dem Niveau von Gruppenmittelwerten (Summscore) geprüft. Zwischen den Post-Test-Gruppen Papier-Bleistift und Computertest bestanden insgesamt kein Mittelwertesunterschied.

In einer getrennten Betrachtung nach Geschlecht und Altersgruppe wurden tendenzielle Differenzen zwischen beiden Versionen in Interaktion mit dem Geschlecht nachgewiesen. Schülerinnen der Computertestgruppe unterschieden sich von den Schülerinnen der Papier-Bleistift-Gruppe in den mittleren Summscores um ca. 2,6 Punkte. Für den Geschlechtseffekt bedeutet dies, dass im Computertest dieser bei ungefähr 0.7 verblieb, während er in der Wiederholung mit der Papier-Bleistift-Version auf 0.38 schrumpfte.

Mit linearen Regressionsmodellen und Kovarianzanalysen wurde dieser Trend genauer untersucht. Am stärksten profitierten Mädchen, die die Papier-Bleistift-Version im Post-Test absolvierten, von der Testwiederholung, am wenigsten die Mädchen im Computertest. Für die Jungen wirkte die Vortestleistung als Kovariate in der Papier-Bleistift- und der Computertest-Gruppe gleich. Die Varianzaufklärung durch die Vortestleistung betrug hier fast 40 %. Für die Mädchen war diese abhängig von der durchgeführten Post-Test-Version.

In den qualitativen Analysen wurde die Methode SNMF und Consensus-Clustering zur automatischen Mustererkennung und Klassifizierung der im MRT-A und den *Würfelaufgaben* eingesetzten Lösungsstrategien erprobt. Es konnten auf den Datensätzen der ca. 200 Schüler für den MRT und die *Würfelaufgaben* jeweils Modelle mit drei Klassen identifiziert werden.

Die drei Klassen im MRT unterschieden sich in der Gesamtestleistung und hinsichtlich der verwendeten Lösungsstrategien. Es bestanden Zusammenhänge zwischen der Geschlechtszugehörigkeit und der Art der verwendeten Strategie. Insbesondere in der leistungsstärksten Klasse, deren Personen weder Schwierigkeiten mit den Items des Tests noch mit der Zeitbegrenzung hatten, waren signifikant mehr Jungen. Mädchen waren dagegen häufiger in einer Klasse, die zwar mentale Rotation ausführten, jedoch wesentlich langsamer. Die Diskrepanz zwischen den Geschlechtern in der Verteilung der effizientesten Strategien war am stärksten im achten Schuljahr. Insgesamt nahm die Anwendung dieser effizienten Strategie über das Alter zu. Wobei auch in den Jahrgängen davor Jungen diese Strategiekategorie dominierten.

Personen einer weiteren Klasse wiesen ebenfalls Schwierigkeiten mit dem Zeitlimit auf. Typ-I-Items und Items mit Okklusion lösten Personen dieser Klasse seltener als Personen der beiden

anderen Klassen. Aber sie erzielten im Vergleich zu den anderen Klassen ähnliche Resultate für Typ-II-Items. Deshalb wurde geschlussfolgert, dass Personen dieser Klasse vorzugsweise Items lösten, die sich mit Detailvergleichen lösen lassen. Diese Strategie wurde von Mädchen und Jungen gleichermaßen oft gewählt.

Die Validität dieser Klasseneinteilung wurde durch klassenspezifische Korrelationen mit Außenkriterien geprüft. „Induktive Denkfähigkeit“ korrelierte mit dem MRT-Score nur in der Klasse mit Detailvergleichen signifikant und in der Höhe eines mittleren Zusammenhangs.

In den *Würfelaufgaben* wurden spezifische Antwortmuster vor allem für „Raumwürfel“ und „Flächenwürfel-Plus“ ausgemacht. In der zahlenmäßig größten Klasse wurden besonders häufig „Flächenwürfel-Plus“ gelöst. Eine weitere Klasse löste vor allem „Raumwürfel“. Dazwischen lag noch eine sehr kleine Klasse, die aber im Lösungsverhalten der letzteren ähnelte. Für diese Klasseneinteilung aus den *Würfelaufgaben* wurden kein signifikanten Zusammenhänge zwischen Strategie und Geschlecht deutlich.

Zusammenhänge zwischen einer Präferenz für analytische bzw. holistische Lösungsstrategien über die Raumvorstellungstests MRT und *Würfelaufgaben* hinweg konnten ebenfalls nicht gezeigt werden. Auch aus den Strategiefragebögen konnten diesbezüglich keine weiteren Erkenntnisse gewonnen werden.

10 Experiment 2 - Pfadintegration

10.1 Allgemeine kognitive Maße

Tabelle 10.1 stellt Gruppenmittelwerte aus den Test zur allgemeinen kognitiven Leistungsfähigkeit für die Altersgruppen jüngere, mittlere und ältere Erwachsene getrennt nach Geschlecht dar.

Tabelle 10.1: Mittelwerte allgemeiner kognitiver Maße in den Alters- und Geschlechtsgruppen.

		männlich		weiblich		F _(5;194)	p
		M	SD	M	SD		
<i>Verbale Intelligenz:</i>							
MWT-B	jüngere	110,4	14,55	111,6	12,53	1,00	.419
	mittlere	117,7	11,91	110,1	14,69		
	ältere	112,2	14,38	116,0	11,01		
<i>Visuo-Motorik und kognitive Flexibilität:</i>							
TMT-A	jüngere	27,9	7,27	24,3	6,69	9,76**	<.001
	mittlere	28,7	4,94	29,0	8,52		
	ältere	37,0	12,11	35,2	9,78		
TMT-B	jüngere	59,2	15,44	49,7	13,61	10,52**	<.001
	mittlere	66,4	10,99	61,9	15,44		
	ältere	79,1	25,00	68,0	19,10		
<i>Befindlichkeit:</i>							
Bf-S	jüngere	10,8	8,13	9,2	6,21	0,52	.759
	mittlere	8,9	5,56	11,4	7,87		
	ältere	10,5	5,96	9,1	6,20		
<i>Visuo-konstruktive Fähigkeiten:</i>							
RCF-copy	jüngere	35,7	0,56	34,7	2,12	1,97	.143
	mittlere	35,2	1,78	34,3	2,12		
	ältere	35,4	0,80	35,4	1,62		

** p<.01

Bezüglich der in den Interaktionsgruppen im Mittel erzielten Rohwerte im Test zur verbalen Intelligenz (MWT-B) ergaben sich keine signifikanten Unterschiede. Auch in der durch die Befindlichkeitsskala (Bf-S) erhobenen aktuellen Stimmungslage bestanden keine signifikanten Mittelwertsunterschiede zwischen den Gruppen. In visuo-konstruktiven Fähigkeiten (RCF-copy) lagen in allen Gruppen fast identische Mittelwerte und geringe Streuungen der Rohwerte vor, woraus auf ein ähnliches Niveau bei allen Versuchspersonen geschlossen werden konnte.

Hinsichtlich Visuo-Motorik und kognitiver Umstellfähigkeit, in der Tabelle als gemittelte Rohwerte des TMT-A bzw. TMT-B dargestellt, wurden signifikante Unterschiede zwischen den Gruppen nachgewiesen. Diese gehen auf die zu erwartenden Alterseffekte in diesen Leistungsmaßen zurück. Mittelwerte und Standardabweichungen in den Altersgruppen lagen aber auf den für die jeweiligen Altersbereiche typischen Niveaus. Auch die nominalen Differenzen zwischen den Gruppen befanden sich innerhalb der zu erwartenden altersbedingten Leistungsdifferenzen.

10.2 Raumkognitive Maße

10.2.1 Raumvorstellungstests

In Tabelle 10.2 sind die Summenscores als Gruppenmittelwerte der drei erhobenen klassischen Raumvorstellungstests *Figurenauswahl*, Untertest 10 des LPS und der RCF in der Recall-Bedingung dargestellt.

Tabelle 10.2: Gruppenmittelwerte der Leistungsmaße aus den Raumvorstellungstests.

		männlich		weiblich		F _(5;194)	p
		M	SD	M	SD		
<i>Mentale Rotation</i>							
FA	jüngere	12,2	4,13	10,6	3,57	4,22**	.001
	mittlere	13,5	2,58	9,0	3,22		
	ältere	9,7	4,66	8,6	4,04		
<i>Feldabhängigkeit / Räumliche Wahrnehmung</i>							
LPS10	jüngere	31,5	5,93	31,7	4,90	9,97**	<.001
	mittlere	30,5	5,41	26,7	5,36		
	ältere	24,3	6,05	25,4	5,79		
<i>Visuell-räumliches Gedächtnis</i>							
RCF-recall	jüngere	25,1	5,06	22,8	6,03	7,16**	<.001
	mittlere	21,6	4,53	18,8	4,09		
	ältere	22,5	5,23	17,6	4,91		

** p<.01

Bezüglich jedes der hier erhobenen Leistungsmaße lagen signifikante Unterschiede zwischen den Interaktionsgruppen vor. Im Test *Figurenauswahl* lag dieses Merkmal in der Stichprobe normalverteilt und varianzhomogen in den Gruppen vor (K-S-Test: $Z=0,96$; $p=.311$; Levene-Test: $F=1,34$; $p=.252$). In der Varianzanalyse (ANOVA) waren Haupteffekte für den Faktor „Altersgruppe“ und für den Faktor „Geschlecht“ signifikant ($F(2,194)=5,48$; $p=.005$ bzw. $F(3,194)=11,54$; $p=.001$). Zur Abschätzung der Größe des Effekts in der Stichprobe ergab sich für beide Faktoren $\eta^2=0.07$. Ein Interaktionseffekt zwischen Alter und Geschlecht lag nicht vor ($F(2,194)=1,79$; $p=.171$). Die erzielten mittleren Punktwerte auf den Faktorstufen betragen für jüngere Erwachsene 11,2 (SD=3,83), für die Gruppe im mittleren Erwachsenenalter $M=10,8$ (SD=3,68) und für ältere Erwachsene 9,2 (SD=4,35) Punkte. Männliche Versuchspersonen erreichten im Mittel 11,4 (SD=4,36) Punkte und weibliche Versuchspersonen 9,7 (SD=3,75).

Auch der Summenscore für den LPS10 lag in der Stichprobe normalverteilt und in den Gruppen varianzhomogen vor. (K-S-Test: $Z=0,97$; $p=.299$; Levene-Test: $F=0,79$; $p=.559$). In der Varianzanalyse ergab sich für diese Leistungsmaß ein starker Effekt für die Altersgruppen ($F(2,194)=22,13$; $p<.001$) mit über 20 % Varianzaufklärung in der Stichprobe durch diesen Faktor ($\eta^2=.233$). Es wurde kein Geschlechtseffekt ($F(2,194)=0,71$; $p=.403$) und kein signifikante Interaktionseffekt nachgewiesen ($F(2,194)=1,74$; $p=.180$). Die nominale Leistung in diesem Test hing also deutlich mit der Zugehörigkeit zu einer Altersgruppe zusammen. Die Gruppe der jüngsten Versuchspersonen erzielten mit $M=31,7$ (SD=5,26) die höchste Punktzahl vor den mittleren ($M=28,2$; SD=5,60) und älteren ($M=24,9$; SD=5,89) Erwachsenen.

Aus dem Raumvorstellungstest zum visuell-räumlichen Gedächtnis (RCF-recall) wurden die Rohwerte verwendet, die ebenfalls in der Stichprobe normalverteilt und varianzhomogen in den Gruppen vorlagen (K-S-Test: $Z=0,94$; $p=.341$ Levene-Test: $F=0,80$; $p=.549$). Neben einem deutlichen Effekt der Altersgruppenzugehörigkeit ($F(2,194)=14,50$; $p<.001$) über den 16 % der Varianzaufgeklärt werden konnte, bestand auch ein geringerer Effekt für den Faktor „Geschlecht“ ($F(2,194)=7,73$; $p=.006$; $\eta^2=0.05$). Signifikante Interaktionen zwischen beiden Faktoren konnten nicht gezeigt werden ($F(2,194)=0,03$; $p=.968$). Die Gruppe der jüngeren Versuchspersonen erzielten mit Abstand die höchste Punktzahl ($M=23,6$; SD=5,77). Die Gruppen der mittleren und älteren Erwachsene unterscheiden sich im Mittel um nur einen Punkt mit $M=19,9$ (SD=4,41) bzw. $M=18,9$ (SD=5,23). Männliche Versuchspersonen erzielten durchschnittlich zwei Punkte mehr als weibliche Versuchspersonen (männlich: $M=22,5$; SD= 5,47; weiblich: $M=20,5$; SD=5,87).

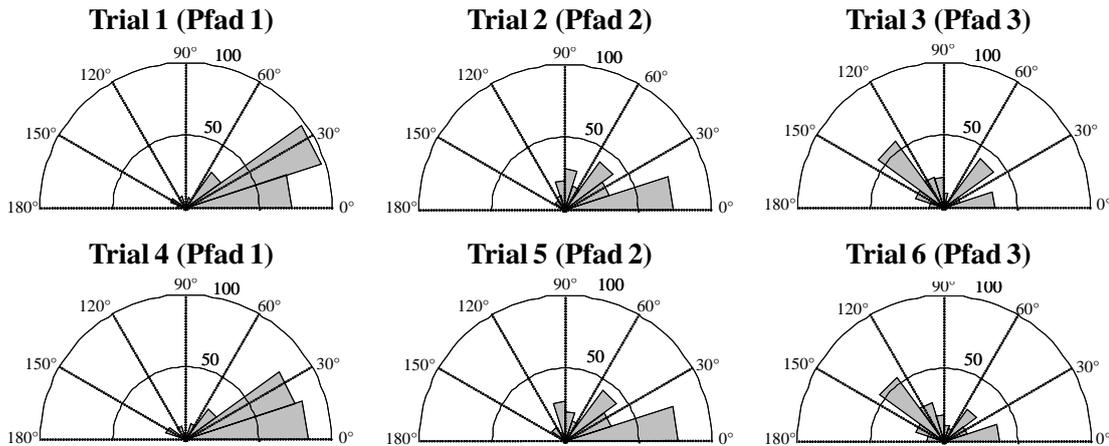
10.2.2 Kompass-Zeigeaufgabe

Pfad und Geschwindigkeit

In der Darstellung der Ergebnisse zur Kompass-Zeigeaufgabe wurden zunächst die Wirkung der Innersubjektfaktoren Art des gelaufenen Pfads (Pfad) und Darbietungsgeschwindigkeit (Geschwindigkeit) auf die Testleistung untersucht. Operationalisiert wurde die Testleistung aus den insgesamt sechs Durchgängen (Trials) über die absoluten Richtungsschätzfehler (σ). Abbildung 10.1 zeigt die

Verteilungen der Schätzfehler als absolute Häufigkeiten in jedem der Trials.

Abbildung 10.1: Verteilungen der Schätzfehler aus der Kompass-Zeigeaufgabe.



Auf dem Halbkreis sind die Schätzfehler in Grad von 0° bis 180° abgetragen. Die grauen Dreiecke sind die Fehlerkategorien. Die Breite einer Kategorie beträgt jeweils 18° . Die Höhe eines Dreiecks gibt die absolute Häufigkeit innerhalb der Kategorie an. Auf dem inneren Halbkreis beträgt diese 50 Personen, auf dem äußeren 100 Personen.

Aus den Verteilungen können erste Aussagen über die Schwierigkeit der einzelnen Trials getroffen werden. Unterschiedliche Darbeitungsgeschwindigkeiten führten offensichtlich für keinen der drei Pfad zu veränderten Verteilungen der Schätzfehler. Trials mit dem Pfad 1 haben offensichtlich zu den genauesten Schätzungen der Heimkehrrichtung geführt. Fast alle Versuchspersonen lagen mit ihren absoluten Schätzfehlern unter 36° . Für den Pfad 2 konnten immer noch recht viele Versuchspersonen (ca. 40 %) sehr genaue Angaben ($\sigma < 18^\circ$) machen. Jedoch verteilen sich die übrigen bereits über den gesamten Wertebereich. In den Trials 4 und 6 wurde der Pfad 3 dargeboten, für den die Schätzfehler bis auf einige Häufungen über den Wertebereich gleichverteilt sind.

Tabelle 10.3 fasst einige Kennwerte der Verteilungen aus Abbildung 10.1 und die Prüfgrößen der Anpassungstests auf Normal- bzw. Gleichverteilung zusammen. In keinem Trial lag mit dem absoluten Schätzfehler ein normalverteiltes Merkmal vor. Auch die Annahme einer Gleichverteilung in Trails mit dem Pfad 3 musste anhand der Prüfgrößen verworfen werden. Zur Durchführung statistischer Unterschieds- und Zusammenhangsüberprüfungen wurden im folgenden nicht-parametrische Verfahren eingesetzt.

Mittelwertsunterschiede zwischen den einzelnen Trials bestanden, wie bereits aus den Verteilungsformen ersichtlich, nur zwischen Durchgängen mit unterschiedlichen Pfaden. Eine paarweise Überprüfung der Trials mit gleichen Pfaden mit dem Wilcoxon-Test für abhängige Stichproben ergab sowohl für die Paarungen Trial 1–4 ($Z=0,31$; $p=.755$) und Trial 2–5 ($Z=0,08$; $p=.938$) als auch Trial 3–6 ($Z=-0,40$; $p=.687$) keine signifikanten Mittelwertsunterschiede. Die geringsten Schätzfehler der Heimkehrrichtung wurden in Trial 1 und 3, gefolgt von Trial 2 und 5 gemessen. Der größte Schätzfehler mit im Mittel fast 90° Abweichung von der korrekten Richtung lag in den Trials 3 und

Tabelle 10.3: Kennwerte der Verteilungen der Schätzfehler aus den einzelnen Versuchsbedingungen der Zeigeaufgabe.

	M	SD	Min-Max	Normalvert.		Gleichvert.	
				Z [†]	p	Z [†]	p
Trial 1	35°	41,7°	0-174°	4,29	<.001	8,53	<.001
Trial 2	45°	38,7°	0°-168°	2,27	<.001	4,94	<.001
Trial 3	87°	54,5°	0°-180°	1,99	<.001	1,39	.042
Trial 4	38°	43,1°	0°-174°	3,74	<.001	7,68	<.001
Trial 5	46°	40,8°	0°-162°	1,99	<.001	5,14	<.001
Trial 6	89°	54,4°	0°-180°	1,83	<.001	1,51	.021

[†]Prüfgröße des Kolmogorov-Smirnov-Anpassungstest

6 vor.

Die Zusammenhänge der Schätzgenauigkeit in Durchgängen mit gleichen Pfaden wurden durch Korrelationen überprüft. Dazu wurden Rangkorrelationskoeffizienten (Spearman's-Rho) zwischen allen Paaren von Trials berechnet (s. Tab. 10.4).

Tabelle 10.4: Korrelationen (Spearman's-Rho) der Versuchsbedingungen untereinander bei der Kompass-Zeigeaufgabe.

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6
Trial 1	-	.184**	-.04	.482**	.146*	-.03
Trial 2		-	.092	.240**	.605**	.180*
Trial 3			-	-.05	.111	.690**
Trial 4				-	.256**	-.06
Trial 5					-	.171*
Trial 6						-

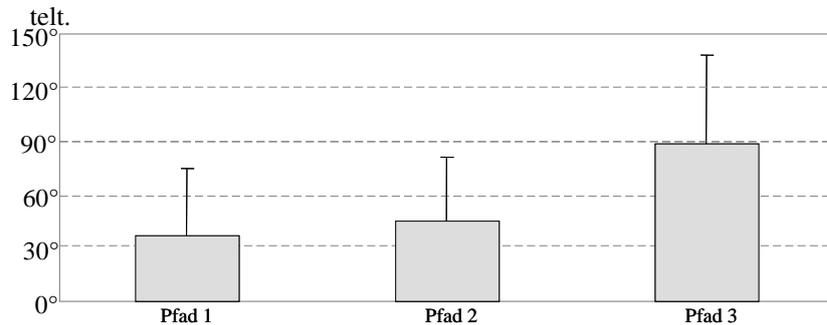
* p<.05, ** p<.01

Die stärksten Korrelation bestanden, wie erwartet, zwischen Bedingungen mit gleich Pfaden (3. Nebendiagonale der Korrelationsmatrix). Die Koeffizienten zeigen für alle Pfade einen starken Zusammenhang an. Schwache bis mittlere Zusammenhänge konnten zwischen Trials mit dem Pfad 1 und dem Pfad 2 nachgewiesen werden. Sehr schwach korrelierten die Trials mit Pfad 2 und Pfad 3 untereinander. Keinen statistischen Zusammenhang gab es zwischen den Schätzfehlern nach Trials des Pfad 1 mit den Schätzfehlern nach Trials mit Pfad 3.

In der Aggregation über Trials mit gleichen Pfaden ergab sich ein mittlerer Schätzfehler für Pfad 1 von 36° (SD=37,7), für Pfad 2 von 45° (SD=35,4) und für Pfad 3 von 88°(SD=50,3). Wie in Abbildung 10.2 zu erkennen, unterschieden sich die absoluten Schätzfehler je nach der Art des ver-

wendeten Pfades insgesamt sehr deutlich. Der Friedman-Test für mehrere unabhängige Stichprobe bestätigte diesen als signifikant ($\chi^2=71,28$; $df=2$; $p<.001$).

Abbildung 10.2: Absolute Schätzfehler und Standardabweichungen über Trials mit gleichen Pfaden gemittelt.



In Paarvergleichen der Schätzfehler ergaben sich für Pfad 1 und 2 eine signifikante Differenz der Mittelwerte von $9,0^\circ$ (Wilcoxon-Test: $Z=-2,38$; $p=.017$). Zwischen Pfad 1 und 3 bestand sogar eine Differenz der Mittelwerte von $51,6^\circ$ ($Z=-9,03$; $p<.001$). Zwischen den mittleren Schätzfehlern aus Pfad 2 und 3 betrug diese $42,6^\circ$ ($Z=-8,56$; $p<.001$).

In der Aggregation über Trials mit gleichen Geschwindigkeiten ergab sich für die einfache Darstellungsgeschwindigkeit (Trial 1 bis 3) ein mittlerer Schätzfehler von 56° ($SD=28,8$) und für die doppelte $M=57^\circ$ ($SD=30,2$). Die Verteilung der Schätzfehler nach Aggregation entsprach in beiden Merkmalen einer Normalverteilung (Geschw. 1: $Z=0,85$; $p=.469$; Geschw. 2: $Z=0,81$; $p=.522$). Die Prüfung der Mittelwerte mit dem t-Test für abhängige Stichproben ergab keine signifikanten Unterschiede ($t(199)=-0,74$; $p=.463$). Die Korrelation zwischen beiden Messwerten betrug $r=0.700$ ($p<.001$).

Alter und Geschlecht

In der Untersuchung des Einflusses der Zwischensubjektfaktoren „Altersgruppe“ und „Geschlecht“ auf die Schätzleistung wurden als unabhängige Variablen die über Trials mit gleichen Pfaden aggregierten, mittleren absoluten Fehler betrachtet. Tabelle 10.5 zeigt Mittelwerte und Standardabweichungen der Interaktionsgruppen und Prüfgrößen des Kruskal-Wallis-Test für mehrere unabhängige Stichproben.

Für jedes der erhobenen Leistungsmaße wurden signifikante Unterschiede der Gruppenmittelwerte ermittelt. Die genauesten Richtungsschätzungen wurden von der Gruppe jüngerer männlicher Erwachsener für Trials mit Pfad 1 gegeben. Sowohl Mittelwert als auch Varianz der Daten waren hier im Vergleich zu anderen Gruppen und Versuchsbedingungen am geringsten. Insgesamt variierten die Gruppenmittelwerte der Schätzfehler für Pfad 1 am stärksten, Gruppenunterschiede der Schätzfehler für Pfad 2 waren am geringsten. Die Schätzfehler für Pfad 3 lagen für alle Alters- und Geschlechtsgruppen im Mittel über den der anderen Versuchsbedingungen. Die Mittelwerte verdeutlichen aber auch Interaktionen zwischen Versuchsbedingungen und Zwischensubjektfaktoren.

Tabelle 10.5: Mittlere Schätzfehler und Reaktionszeit der Zeigeaufgabe in den Alters- und Geschlechtsgruppen.

		männlich		weiblich		χ^2	df	p
		M	SD	M	SD			
<i>mittlere absolute Schätzfehler:</i>								
Pfad 1	jüngere	20°	16,5°	35°	33,5°			
	mittlere	33°	37,8°	65°	38,5°	22,96**	5	<.001
	ältere	44°	44,1°	48°	50,1°			
Pfad 2	jüngere	41°	35,6°	35°	30,2°			
	mittlere	35°	26,1°	52°	36,1°	11,62*	5	.040
	ältere	58°	37,1°	51°	37,7°			
Pfad 3	jüngere	76°	57,3°	70°	48,1°			
	mittlere	102°	43,5°	100°	36,9°	17,80**	5	.003
	ältere	110°	43,4°	95°	42,4°			
<i>mittlere Reaktionszeit:</i>								
	jüngere	27,5	15,69	14,3	4,18			
[Sek]	mittlere	15,8	3,06	16,2	3,40	27,95**	5	<.001
	ältere	25,2	17,40	18,5	9,91			

* p<.05, ** p<.01

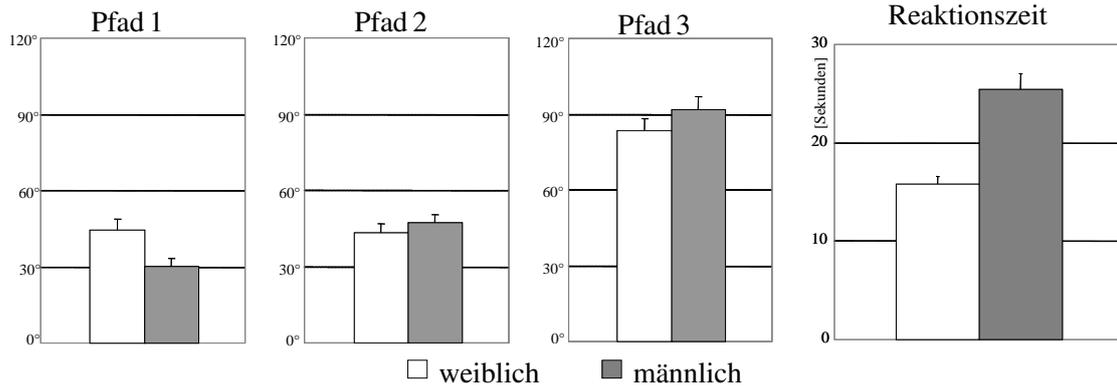
Die Varianz der Schätzfehler lag über alle Versuchsbedingungen und alle Interaktionsgruppen mit 30° bis 60° Standardabweichung sehr hoch.

Zusätzlich wurde als Verhaltensmaß aus dem VPT noch die über alle Trials gemittelte Reaktionszeit in die Tabelle 10.5 aufgenommen. Auch bestanden signifikante Unterschiede zwischen den Interaktionsgruppen. Im Mittel wurden für die Abgabe einer Richtungsschätzung 16 Sekunden benötigt mit wenigen Sekunden Variation. Nur die Gruppen der jüngeren und älteren Männer wichen von diesen mittleren Verhaltensparametern deutlich ab.

Die Zusammenfassung der Schätzfehler über die Stufen des Faktors „Geschlecht“ zeigt Abbildung 10.3. Deutlich zuerkennen ist ein signifikanter Geschlechtsunterschied für die absoluten Schätzfehler in Trials mit Pfad 1 (Mann-Whitney U-Test: $Z=-3,07$; $p=.002$). Weibliche Versuchspersonen erzielten für diesen Pfad mittlere Schätzfehler von 44° (SD=41,2) und männliche von 30° (SD=33,6). Im Verhältnis zur Standardabweichung ausgedrückt ist der Mittelwertsunterschied aber eher als geringer Effekt zu beurteilen ($d=0,38$). Als Zusammenhangsmaß wurde der Rangkorrelationskoeffizient Spearman's-Rho mit $-0,218$ ($p=.002$) bestimmt.

Für den Pfad 2 unterschieden sich die mittleren Schätzleistungen der Männern und Frauen nicht (U-Test: $Z=-0,61$ $p=.543$). Beide Geschlechter erzielten Schätzfehler von zirka 45° (SD=36,3). Im

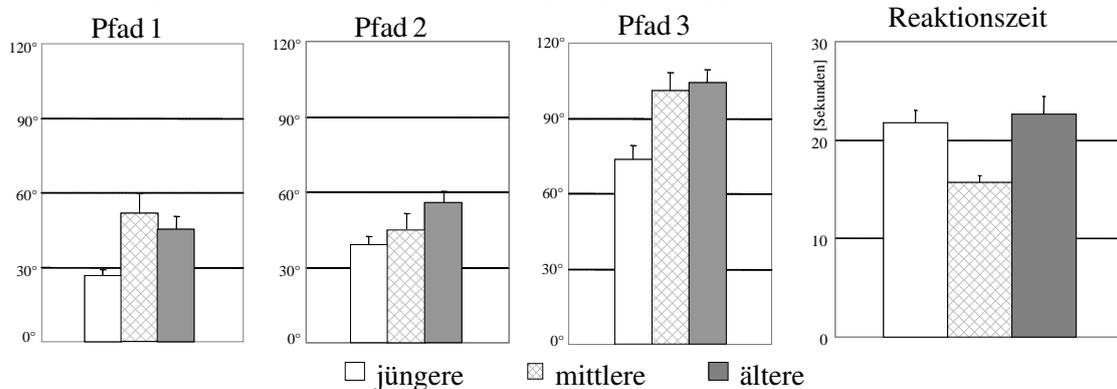
Abbildung 10.3: Geschlechtsdifferenzen in der Zeigeaufgabe. Gruppenmittelwerte und Standardfehler.



Pfad 3 kehrte sich der Mittelwertsunterschied tendenziell um ($Z=-1,60$; $p=.109$). Die gemittelten Schätzfehler der weiblichen Versuchspersonen betragen für diese Trials $M=83^\circ$ ($SD=46,0$) und für die männlich $M=91^\circ$ ($SD=53,3$). In den Reaktionszeit wurde eine signifikante Differenz der Gruppenmittelwerte von ca. zehn Sekunden ermittelt ($Z=-4,61$; $p<.001$). Die mittleren Schätzzeiten pro Trial betragen 15,8 Sekunden ($SD=6,61$) für die Gruppe der Frauen und 25,5 Sekunden ($SD=15,91$) für die Gruppe der männlichen Versuchspersonen.

Zu den Gruppenunterschieden zwischen den Altersgruppen sind in Abbildung 10.4 Mittelwerte und Standardfehler dargestellt. Kruskal-Wallis-Test ergaben signifikante Gruppenunterschiede der mittleren Schätzfehler für alle drei Pfade (Pfad 1: $\chi^2=10,53$; $df=2$; $p=.005$; Pfad 2: $\chi^2=9,22$; $df=2$; $p=.010$; Pfad 3: $\chi^2=14,84$; $df=2$; $p=.001$).

Abbildung 10.4: Altersunterschiede in der Zeigeaufgabe. Gruppenmittelwerte und Standardfehler.



Paarvergleiche zwischen den Altersgruppen identifizierten die Mittelwertsunterschiede zwischen jüngeren und älteren Erwachsenen als signifikant. Dies galt für die mittleren Schätzfehler aller drei Pfade (Pfad 1: $Z=1,93$; $p=.054$; Pfad 2: $Z=3,00$; $p=.003$; Pfad 3: $Z=3,62$; $p<.001$). Jüngere Erwachsene und Probanden der mittleren Altersgruppe unterschieden sich in ihren Mittelwerten nur für Pfad 1 und Pfad 3 signifikant ($Z=-3,15$; $p=.002$ bzw. $Z=2,27$; $p=.023$). Bezüglich der Reaktionszeiten bestanden keine statistisch signifikanten Unterschiede (H-Test: $\chi^2=2,23$; $df=2$; $p=.328$).

Nach der konservativen und für innere und äußere Faktoren getrennten inferenzstatistischen Beur-

teilung durch nicht-parametrische Verfahren wurde trotz der Verletzung der Normalverteilung der Schätzfehler das Design durch eine Varianzanalyse mit Messwiederholung (RM ANOVA) analysiert. Der messwiederholte Faktor war die Art des verwendeten Pfades. Unabhängige Variablen waren die über Trails mit gleichen Pfaden gemittelten absoluten Schätzfehler.

Die Überprüfung der Varianzhomogenität der Schätzfehler zwischen den Stufen der Faktoren „Altersgruppe“ und „Geschlecht“ ergaben auch signifikante Abweichungen von dieser Voraussetzung (Pfad 1: $F=8,84$; $p<.001$; Pfad 2: $F=1,49$; $p=.196$; Pfad 3: $F=5,11$; $p<.001$). Durch die Korrelationen zwischen den Schätzfehlern aus Pfad 1 und 2 bzw. 2 und 3 konnte keine Sphärität angenommen werden. Der Mauchly-Test ergab hier $W=0,815$ ($p<.001$), so dass für die Bestimmung der Prüfgrößen eine Korrektur der Freiheitsgrade nach Greenhouse-Geisser vorgenommen wurde ($\epsilon = 0,844$).

In der Varianzanalyse ergab sich, wie durch die vorangegangenen Betrachtungen erwartet, ein Haupteffekt für den inneren Faktor „Pfad“ ($F(1,69;327,3)=78,68$; $p<.001$). Die Größe des Effekts betrug auf Stichprobeniveau $\eta^2=0.29$. Interaktionen dieses Faktors mit dem Faktor „Altersgruppe“ waren nicht signifikant ($F(3,37;327,3)=1,39$; $p=.243$), Interaktionen mit dem Faktor „Geschlecht“ signifikant ($F(1,69;327,3)=3,78$; $p=.030$). Letztere waren sehr gering ($\eta^2=0.02$) und gehen auf die Umkehrung der Mittelwertsunterschiede von Pfad 1 zum Pfad 3 zurück. Interaktionen aller drei Faktoren konnten nicht nachgewiesen werden ($F(3,37;327,3)=0,24$; $p=.890$).

Abschließend wurden varianzanalytisch die Effekte der Zwischensubjektfaktoren auf den Gesamtfehler untersucht. Der mittlere Gesamtfehler lag in der Stichprobe um den Mittelwert $M=57^\circ$ mit einer Standardabweichung von $SD=27,2^\circ$ normalverteilt und in den Interaktionsgruppen varianzhomogen vor (K-S-Test: $Z=0,861$; $p=.449$; Levene-Test: $F=0,58$; $p=.714$). Für den Faktor „Altersgruppe“ ergab sich bezüglich des Gesamtfehlers ein signifikante Haupteffekt ($F(2,194)=16,02$; $p<.001$). Mehr als 14 % der Varianzaufklärung konnten dem Faktor zugeschrieben werden. Ein signifikanter Geschlechtseffekt bezüglich des Gesamtfehlers ($F(1,194)=0,72$; $p=.398$) oder ein Interaktionseffekt mit dem Faktor „Altersgruppe“ bestand nicht ($F(1,194)=1,72$; $p=.183$).

Die Differenz der Mittelwerte zwischen der Gruppe der jüngeren und der mittleren Erwachsenen betrug nominal $19,5^\circ$ und ist statistisch signifikant. Jüngere und ältere unterschieden sich signifikant um $22,0^\circ$. Mittelwertsdifferenzen zwischen mittleren und älteren Erwachsenen waren mit $2,5^\circ$ nur marginal und nicht signifikant.

10.2.3 Analyse zum Desorientierungseffekt

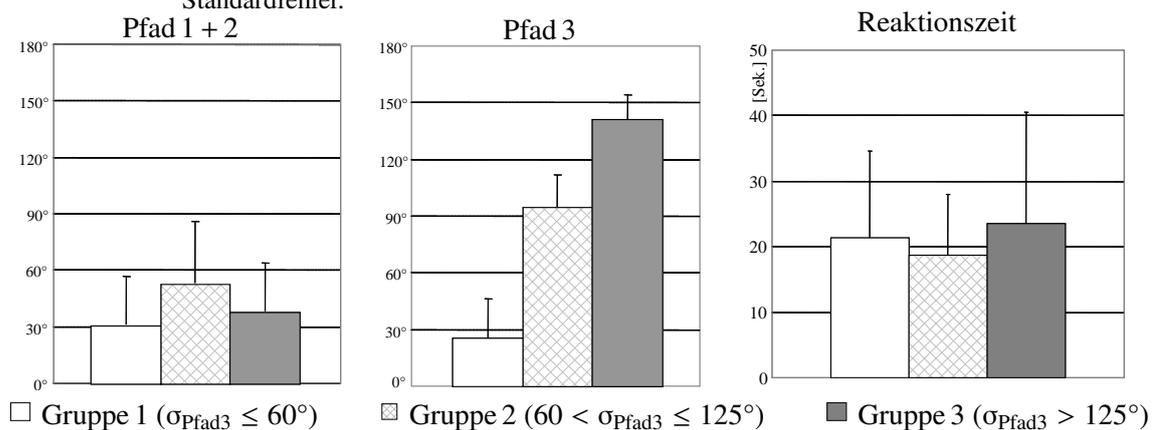
Versuchsbedingungen mit dem Pfad 3 wurden in die Untersuchung aufgenommen, um den Desorientierungseffekt bei Drehung um mehr als 90° zu Verhaltensdifferenzierung einzusetzen. Wären alle Versuchspersonen in diesen Bedingungen desorientiert, würden die Schätzfehler sich über den gesamten Wertebereich gleich verteilen. Stattdessen wurde bei den Häufigkeiten der Schätzfehler aus Trials mit Pfad 3 eine annähernde Gleichverteilung beobachtet (Kolmogorov-Smirnov-Test: $Z=1,22$; $p=.102$). Beobachtet wurden aber auch Häufungen der Antworten bei 126° , bei 54° und für Schätzfehler nahe 0° . Offensichtlich gaben einige Versuchspersonen über die erwarteten Häufigkei-

ten hinaus systematische Antworten. Darüber hinaus konnten einige die Richtung trotz der in den Trials implementierten Schwierigkeiten die Richtung sehr gut einschätzen.

Für die weitere Analyse der Schätzfehler aus diesen Trials wurde der über Trials mit Pfad 3 aggregierte Mittelwert betrachtet. Weiterhin wurde angenommen, dass sich Versuchspersonen anhand der Performanz in diesen Trials in drei Gruppen unterscheiden lassen. Sehr gute Schätzer, die offensichtlich über zusätzliches Raumwissen verfügen um den Desorientierungseffekt zu kompensieren, sollten durch sehr kleine Schätzfehler auffallen. Personen, die durch Pfadintegration die Richtung abschätzten, sollten eher gleichverteilt über den gesamten Wertebereich liegen. Von letzteren sollten aber jene unterschieden werden, die sehr große Schätzfehler begingen, jedoch mit einer Systematik.

Zur weiteren Beurteilung dieser Schätzfehler wurde eine Aufteilung der Versuchspersonen in drei gleich große Gruppen anhand des mittleren Schätzfehlers für Trails mit Pfad 3 vorgenommen. Dazu wurde ein Perzentil-Split, ähnlich dem Median-Splitverfahren, angewandt. Die erste Teilung erfolgte beim 33%-Perzentilwert von $\sigma < 60^\circ$, die zweite Teilung beim 66%-Perzentilwert von $\sigma < 125^\circ$. In Abbildung 10.5 sind die mittleren Schätzfehler und Reaktionszeiten in diesen drei Gruppen dargestellt.

Abbildung 10.5: Mittlere Schätzfehler für Bedingungen mit Desorientierungseffekt (Pfad 3) und ohne (Pfad 1 + 2) der drei Subgruppen. Mittlerer Reaktionszeiten in den Subgruppen. Mittelwerte und Standardfehler.



Hinsichtlich der Reaktionszeiten gab es keine signifikanten Unterschiede zwischen den Gruppen ($\chi^2=0,59$; $df=2$; $p=.746$). In den mittleren Schätzfehlern zum Pfad 3 unterschieden sich die Gruppen, bedingt durch die Art der Klassifizierung, maximal. Der mittlere Schätzfehler der Gruppe 1 betrug lediglich 26° ($SD=21,1^\circ$), der Gruppe 2 95° ($SD=17,7^\circ$) und der Gruppe 3 141° ($SD=13,3^\circ$).

Die Schätzfehler aus Trials mit Pfad 1 und Pfad 2 wurden für die Gegenüberstellung aggregiert zu einem Schätzfehler aus Trials mit Pfaden ohne Desorientierung. Paarvergleiche ergaben signifikante Mittelwertsunterschiede zwischen Gruppe 2 und 3 ($Z=2,85$; $p=.004$) und Gruppe 1 und Gruppe 2 ($Z=-4,29$; $p<.001$). Tendenziell signifikant war der Unterschied zwischen Gruppe 1 und Gruppe 3 ($Z=-1,87$; $p=.061$). Die absoluten Schätzfehler betragen für letztere 31° ($SD=25,4$) bzw. 38° ($SD=25,5$). Die Gruppe 2 unterschied sich im mittleren Fehler sehr deutlich von den anderen Gruppen ($M=53^\circ$; $SD=33,1^\circ$).

Zur Beschreibung der Gruppen wurden im Folgenden Bezeichnungen vergeben. Personen der Gruppe 1 wurden mit „Orientiert“ bezeichnet, da sie trotz der Schwierigkeit des Pfad 3 den Desorientierungseffekt kompensieren konnten. Personen der Gruppe 2 waren in diesen Trials demzufolge „Desorientiert“. Bei Personen der Gruppe 3 wurde vermutet, dass diese die von ihnen häufig eingestellte Heimkehrrichtung von 180° eher geraten hatten.

Neben den Mittelwertsvergleichen wurden Korrelationen zwischen den mittleren Schätzfehlern für die verschiedenen Pfade innerhalb der drei Gruppen bestimmt. Tabelle 10.6 zeigt die entsprechenden Korrelationsmatrizen. Es wurden Rangkorrelationskoeffizienten Spearmans-Rho berechnet.

Tabelle 10.6: Korrelationen (Spearman's-Rho) der Schätzfehler in den drei Gruppen.

	Orientiert (N=64)			Desorientiert (N=69)			Geraten (N=67)		
	1	2	3	1	2	3	1	2	3
Pfad 1	-	.379**	.009	-	.265*	-0.09	-	.213 ⁺	-0.14
Pfad 2		-	.256*		-	-0.11		-	-0.12
Pfad 3			-			-			-

⁺ p<.10, * p<.05, ** p<.01

In der Gruppe „Orientiert“ bestand lediglich ein tendenzieller und schwächerer Zusammenhang zwischen den Schätzfehlern nach Pfad 1- und Pfad 2-Bedingungen. In Gruppe „Desorientiert“ war diese Zusammenhang nur marginal stärker. Nicht signifikant waren in beiden Gruppen Zusammenhänge zwischen den Schätzfehlern für Pfad 1 bzw. 2 mit dem Schätzfehler für Pfad 3. In Gruppe „Geraten“ bestand hier ein schwacher bis mittlerer Zusammenhang zwischen Pfad 2 und Pfad 3 Schätzfehlern. In dieser Gruppe korrelierten Schätzfehler für Pfad 1 und 2 außerdem am stärksten.

Zur Bestimmung von Zusammenhängen zwischen Desorientierung in Bedingungen mit Pfad 3 und der Zugehörigkeit zu einer Alters- bzw. Geschlechtsgruppe wurden Häufigkeitsverteilungen der Merkmale in Kontingenztabellen überprüft. Die Verteilung weiblicher bzw. männlicher Versuchspersonen auf die drei Gruppen sind in Tabelle 10.7 dargestellt.

Tabelle 10.7: Verteilung männlicher und weiblicher Versuchspersonen auf die Gruppen nach Schätzgenauigkeit im Pfad 3.

	Orientiert	Desorientiert	Geraten	Gesamt
männlich	36 (32 %)	29 (26 %)	47 (42 %)	112 (100 %)
weiblich	28 (32 %)	40 (46 %)	20 (28 %)	88 (100 %)
korrigiert-standardisierte Residuen	0,05	-2,89	2,86	

Insgesamt unterschieden sich in den Gruppen die beobachteten von den erwarteten Häufigkeiten signifikant ($\chi^2=10,91$; $df=2$; $p=.004$). Mit dem korrigierten Kontingenzkoeffizienten von $C_{\text{kor}}=0.321$ konnte ein mittlerer Zusammenhang zwischen beiden Merkmalen ausgemacht werden. In der Betrachtung der Häufigkeiten auf den einzelnen Faktorstufen wurden in der Gruppe „Desorientiert“

signifikant mehr weibliche Versuchspersonen und in der Gruppe „Geraten“ signifikant mehr männliche beobachtet. In Tabelle 10.7 ist diese durch die korrigierten Residuen mit Werten größer 1.96 ablesbar. In der Gruppe „Orientiert“ waren Männer und Frauen prozentual gleich stark vertreten.

Bezüglich der Altersgruppen sind die entsprechenden Häufigkeitsverteilung in Tabelle 10.8 zusammengefasst. Auch hier wurden deutliche Abweichungen der Häufigkeiten von den Erwartungswerten beobachtet ($\chi^2=25,01$; $df=4$; $p=.000$). Insgesamt konnte zwischen den Faktoren „Altersgruppe“ und Gruppenzugehörigkeit nach Schätzgenauigkeit im Pfad 3 ein mittlerer Zusammenhang festgestellt werden ($C_{\text{kor}}=0.408$).

Tabelle 10.8: Verteilung der Altersgruppen auf die drei Gruppen nach Schätzgenauigkeit im Pfad 3.

	Orientiert	Desorientiert	Geraten	Gesamt
jüngere	48 (47 %) **	26 (25 %)	29 (28 %)	103 (100 %)
mittlere	3 (11 %)	16 (59 %) **	8 (30 %)	27 (100 %)
ältere	13 (19 %)	27 (39 %) *	30 (42 %) *	70 (100 %)

** korrigiertes Residuum > 2.58 * korrigiertes Residuum > 1.96

Signifikante Abweichungen auf den einzelnen Faktorstufen wurden in der Gruppe „Orientiert“ für die Häufigkeit der jüngeren Versuchspersonen beobachtet. 47 % der jüngeren Erwachsenen waren in dieser Gruppe. Im Vergleich dazu waren aus den anderen Altersgruppen jeweils weniger als 20 % hier vertreten. In der Gruppe „Desorientiert“ waren überproportional viele Erwachsenen mittleren Alters. In der dritten Gruppe, die mit dem oben beschriebenen Rateverhalten in Verbindung gebracht wurde, waren dagegen häufiger ältere Erwachsenen vertreten.

10.2.4 Zusammenhänge zwischen den Maßen

Abschließend wurden Zusammenhänge zwischen Schätzfehlern mit den anderen im Experiment erhobene raumkognitiven Leistungsmaßen korrelativ untersucht. Tabelle 10.9 zeigt die Korrelationsmatrix der Spearmans-Rho-Koeffizienten, welche berechnet wurden zwischen den Summenscores der Raumvorstellungstest und den mittleren absoluten Schätzfehlern für Pfad 1 bis 3. Gesamtfehler ist der über alle Trails gemittelte absolute Schätzfehler.

Die Interkorrelationen der Summenscores der Raumvorstellungstests lagen in dieser Stichprobe alle in einem Bereich mittlerer Zusammenhänge. Ebenfalls mittlere oder nur schwache Interkorrelationen bestanden zwischen den Schätzfehlern in der Zeigeaufgabe. Der Gesamtfehler korrelierte entsprechend hoch mit dem ihm zugrunde liegenden Maßen. Dominiert wurde dieser durch die hohen Schätzfehler für Pfad 2 und 3, was sich in sehr hohen Koeffizienten bis zu $\rho=0.7$ äußerte.

Zusammenhänge zwischen Summenscores und Schätzfehlern für einzelne Pfade waren eher schwach, jedoch in fast allen Fällen signifikant. Die Richtung der Korrelation stimmte in allen Fällen überein. D.h. bessere Leistung im Raumvorstellungstest bedeutete tendenziell geringere Schätzfehler in der Zeigeaufgabe. Für den Gesamtfehler ergaben sich mittlere Korrelationen mit den Sum-

Tabelle 10.9: Korrelationen (Spearman's-Rho) zwischen Schätzfehlern in der Zeigeaufgabe und den weiteren erhobenen raumkognitiven Leistungsmaßen.

	1. FA	2. LPS10	3. RCF-delay	4. Pfad 1	5. Pfad 2	6. Pfad 3	7. Gesamtfehler
1.	-	.392**	.271**	-.03	-.145 ⁺	-.139 ⁺	-.154 ⁺
2.		-	.378**	-.161*	-.223**	-.210*	-.347**
3.			-	-.186*	-.256**	-.181*	-.315**
4.				-	.303**	-.05	.476**
5.					-	.165*	.704**
6.						-	.659**
7.							-

⁺ p<.10, * p<.05, ** p<.01

menscores der Raumvorstellungstests zur räumlichen Wahrnehmung (Feldabhängigkeit) und dem visuell-räumlichen Gedächtnis.

Da in allen drei Leistungsmaßen Altersdifferenzen vorlagen, wurden Partial-Korrelation zwischen dem Gesamtfehler und den Raumvorstellungstests mit Altersgruppe als Kontrollvariable berechnet. Lediglich eine schwächere Korrelation zwischen LPS10 und Gesamtfehler mit $r=-0.173$ ($p=.034$) konnte noch nachgewiesen werden.

10.3 Zusammenfassung

Im Screening allgemeiner kognitiver Fähigkeiten innerhalb der Stichprobe ergab sich zwischen den Alters- und Geschlechtsgruppen eine weitgehende Vergleichbarkeit. Unterschiede bestanden in Visuo-Motorik und kognitiver Flexibilität gemessen über die Verarbeitungsgeschwindigkeiten im TMT.

In den Leistungsmaßen aus klassischen Raumvorstellungstests zeigten sich vorwiegend Altersunterschiede. In den Tests zur Feldabhängigkeit und zum visuell-räumlichen Gedächtnis konnten ca. 20 % der Varianz durch die Zuweisung zu den Altersgruppen aufgeklärt werden. Ein geringer Alterseffekte bestanden im Test *Figurenauswahl* (9 % Varianzaufklärung). Geschlechtseffekte wurden ebenfalls, wenn auch in geringerem Maße, im Test zum visuell-räumlichen Gedächtnis und dem Test *Figurenauswahl* nachgewiesen. In beiden Tests lagen die Gruppenmittelwerte männlicher über denen weiblicher Versuchspersonen.

Zur Auswertung der Kompass-Zeigeaufgabe wurden allgemeine und differentielle Effekte der Aufgabe auf die absoluten Schätzfehler betrachtet. Durch diese Transformation lagen die Richtungs-schätzfehler in keinem der Trials normalverteilt vor. Schätzfehler aus Bedingungen mit dem Pfad 3 näherten sich aber einer Gleichverteilung.

Mittlere Schätzfehler für Bedingungen mit Pfad 1 und 2 lagen bei 35° bis 45°. Bedingungen mit dem Pfad 3 führten zu den mit Abstand größten mittleren Schätzfehlern. Sie lagen mit ca. 90° in

der Mitte des Wertebereichs. Somit betrug der Unterschied der mittleren Schätzfehler zwischen Pfaden mit 90°-Turns und 135°-Turns 40° bis 50°. Im Vergleich dazu stieg der mittlere Fehler um weniger als 10°, wenn drei statt zwei 90°-Turns gelaufen wurden. In der Aggregation von Bedingungen mit einfacher bzw. doppelter Darbietungsgeschwindigkeit zeigten sich keine Differenzen der Mittelwerte. Bedingungen mit gleichen Pfaden korrelierten allerdings hoch miteinander.

In der Untersuchung der Zwischensubjektfaktoren zeigte sich eine deutliche Altersdifferenzierung der Aufgabe. Jüngere Erwachsene erzielten im Mittel den geringsten Schätzfehler gegenüber der mittleren und älteren Altersgruppe. Die Differenz der Gruppenmittelwerte des Gesamtfehlers betrug ca. 20°. Von der mittleren zur älteren Gruppe wurde keine weitere Verschlechterung beobachtet. Geschlechtseffekte wurden in Interaktion mit dem Faktor „Pfad“ beobachtet. Nur in Bedingungen mit Pfad 1 erzielten männliche Versuchspersonen um ca. 14° bessere Schätzungen. Der Effekt war allerdings gering.

In der Auswertung der Reaktionszeiten ergab sich ein sehr heterogenes Bild. Gruppenmittelwerte schwankten mit deutlichen Varianzunterschieden zwischen 15 und ca. 30 Sekunden pro Schätzung. Weibliche Versuchspersonen benötigten zehn Sekunden weniger als männliche Versuchspersonen. Zwischen den Altersgruppen lagen aber keine signifikanten Unterschiede vor.

Abschließend wurde der Desorientierungseffekt in Bedingungen mit 135°-Turns zur Differenzierung der Strategien in der Zeigeaufgabe eingesetzt. Aus der Verteilung der Schätzfehler für die Pfad 3-Bedingungen wurde auf drei mögliche Gruppen geschlossen. Die Klassifizierung erfolgte durch ein Perzentil-Splitverfahren in drei ungefähr gleich große Gruppen.

Bei der Gruppe von Versuchspersonen mit dem geringsten absoluten Schätzfehler für Pfad 3 wurde vermutet, dass sie neben Pfadintegrationsprozessen auch Überblickwissen zur Bestimmung der Heimkehrrichtung nutzten. Die mittlere Gruppe versuchte vermutlich durch Pfadintegration die Richtung zu bestimmen. Deren Schätzfehler verteilten sich gleichmäßig um einen Mittelwert von 95° und wies somit einen typische Desorientierungseffekt aus. Für Personen der dritten Gruppe wurde wiederum eine Systematik der Schätzfehler mit einer Häufung bei einem bestimmten Wert festgestellt. Daraus wurde abgeleitet, dass diese Personen die Richtung eher geraten hatten.

In Bedingungen mit ausschließlich 90°-Turns unterschieden sich die Gruppen mit dem orientierten und dem Rateverhalten nicht. Personen der Gruppe mit desorientiertem Verhalten wiesen allerdings schon in diesen Bedingungen größere Schätzfehler auf. Signifikante Korrelationen zwischen den Schätzfehlern aus Bedingungen mit Pfad 1 und Bedingungen mit Pfad 3 bestanden in keiner der drei der Gruppe. Für Personen mit orientiertem Verhalten bestand eine geringe statistisch bedeutsame Korrelation zwischen Pfad 2 und 3.

In der Verteilung der Alters- und Geschlechtsgruppen auf die gewonnene Klassifizierung bestanden auf einzelnen Faktorstufen Abweichungen von den erwarteten Werten. Weibliche bzw. älterer Versuchspersonen wurden öfter der Gruppe, die einen Desorientierungseffekt zeigten, zugewiesen. Männliche Versuchspersonen waren öfter in der Gruppe mit Rateverhalten.

11 Experiment 3 - Place-Learning

11.1 Allgemeine kognitive Maße

In Tabelle 11.1 sind die Mittelwerte bezüglich aller Faktorstufen der unabhängigen Variablen Altersgruppe und Geschlecht dargestellt und auf Gruppenunterschiede varianzanalytische getestet worden.

Tabelle 11.1: Mittelwerte und Standardabweichungen jüngerer und älterer Erwachsener für beide Geschlechter der Tests zum Screening allgemeiner kognitiver Fähigkeiten.

		weiblich		männlich		F _(3;54)	p
		M	SD	M	SD		
<i>Verbale Intelligenz:</i>							
MWT-B	jüngere	104,4	10,41	110,6	16,93	0,89	0.454
	ältere	111,8	9,98	108,3	14,08		
<i>Visuo-Motorik und kognitive Flexibilität:</i>							
TMT-A	jüngere	29,4	7,05	27,4	7,60	3,46*	0.022
	ältere	35,8	10,65	36,3	10,58		
TMT-B	jüngere	54,2	12,18	60,9	17,96	4,27**	0.009
	ältere	70,9	21,47	80,9	29,83		
<i>Befindlichkeit:</i>							
Bf-S	jüngere	8,3	6,52	10,6	8,88	1,18	0.328
	ältere	7,7	5,91	12,1	7,07		
<i>Visuo-konstruktive Fähigkeiten:</i>							
RCF-copy	jüngere	35,9	0,54	35,7	0,61	0,48	0,697
	ältere	35,9	0,26	35,8	0,56		

** p<.01, * p<0.05

Hinsichtlich der Rohwerte aus der Schätzung der kristallinen Intelligenz (MWT-B) zeigten sich keine Gruppenunterschiede in der Stichprobe. Auch bei visuo-konstruktiven Fähigkeiten ergaben sich bei einem Vergleich der Rohwerte aus dem RCF-copy-Test keine statistisch signifikanten Gruppenunterschiede. Bei der Einschätzung der aktuellen Stimmungslage zeigten sich außer einer Tendenz männlicher Versuchspersonen zu höheren Scores in der Bf-Skala ($t(56)=1,81$; $p=.076$) keine

bedeutsamen Differenzen und insbesondere keine Auffälligkeiten in Richtung pathologischer Verstimmungen einzelner Versuchspersonen.

Für Visuo-Motorik und kognitiven Flexibilität liefern die Verarbeitungszeiten aus dem TMT-A bzw. TMT-B deutliche Hinweise auf bestehende Gruppenunterschiede hinsichtlich dieser Fähigkeiten in der untersuchten Stichprobe. Die Mittelwerte und Varianzen in der Tabelle legen nahe, dass diese Differenzen hauptsächlich einem Effekt der Zuordnung zu den beiden Altersgruppen zuzuschreiben sind.

Sowohl für den TMT-A als auch für den TMT-B lagen varianzhomogene, normalverteilte Daten vor, so dass zweifaktorielle Varianzanalysen durchgeführt werden konnten (TMT-A: $F=1,01$; $p=.396$; $Z=0,86$; $p=.454$; TMT-B: $F=1,81$; $p=.157$; $Z=1,05$; $p=.216$). Es ergaben sich für visuo-motorische Fähigkeiten (TMT-A) ein signifikanter Haupteffekt für den Faktor „Altersgruppe“ ($F(1,54)=10,03$; $p=.003$), kein signifikanter Effekt für den Faktor „Geschlecht“ ($F(1,54)=0,09$; $p=.762$) und kein Interaktionseffekt ($F(1,54)=0,28$; $p=.602$). Auch hinsichtlich der kognitiven Flexibilität (TMT-B) zeigte sich ein deutlicher Alterseffekt ($F(1,54)=10,54$; $p=.002$), ebenfalls kein statistisch signifikanter Einfluss der Geschlechtszugehörigkeit ($F(1,54)=2,15$; $p=.149$) und keine Interaktion beider Faktoren ($F(1,54)=0,08$; $p=.772$).

Abschließend zur Betrachtung der allgemeinen Leistungsmaße lässt sich feststellen, dass sich die getesteten älteren Erwachsenen gegenüber den jüngeren im Bereich Visuo-Motorik und kognitive Flexibilität durch höhere Verarbeitungszeiten unterscheiden. Konkret betragen die mittleren Differenz zwischen jüngeren und älteren Probanden im TMT-A 7,6 Sekunden ($SD= 2,38$) und im TMT-B 18,36 Sekunden ($SD=5,70$).

11.2 Vorerfahrung

In Tabelle 11.2 sind Summenscores der Interaktionsgruppen zur Abschätzung von Unterschieden in der Vorerfahrung mit Computern und Raumvorstellungstest abgebildet. Signifikante Unterschiede in diesem Maß traten insbesondere zwischen jüngeren und älteren Männern zu tage. Die Varianzaufklärung durch den Faktor „Gruppenzugehörigkeit“ betrug $\eta^2=0.231$.

Tabelle 11.2: Summenscores in den Interaktionsgruppen des Fragebogens Vorerfahrung.

	weiblich		männlich		$F_{(3;54)}$	p
	M	SD	M	SD		
jüngere	5,64	2,37	7,00	2,75	5,39	.003
ältere	5,16	1,67	4,20	2,40		

11.3 Raumkognitive Maße

11.3.1 Raumvorstellungstests

Tests RCF-recall und LPS10

Auch in den Tests zu Erinnerung visuell-räumlicher Informationen (RCF-recall) und zur feldabhängigen Entscheidungsfindung (LPS10) ergaben sich deutliche Unterschiede zwischen den Altersgruppen, aber keine signifikanten Geschlechtsdifferenzen. Tabelle 11.3 verdeutlicht dies anhand der Gruppenmittelwerte und der F-Statistiken aus der einfaktoriellen Varianzanalyse mit Faktor „Gruppenzugehörigkeit“.

Tabelle 11.3: Gruppenspezifische Testleistungen aus dem LPS10 und dem RCF-recall.

		männlich		weiblich		F _(3;54)	p
		M	SD	M	SD		
<i>Feldabhängigkeit:</i>							
LPS10	jünger	31,1	5,05	29,5	5,35	8,81**	<.001
	ältere	22,6	5,60	23,1	6,23		
<i>Visuell-räumliches Gedächtnis:</i>							
RCF-recall	jüngere	25,4	6,06	23,5	5,27	6,50**	<.001
	ältere	19,8	5,26	16,9	5,99		

**p<.01

Levene-Tests bestätigen die aus der Tabelle ersichtlichen Varianzhomogenitäten zwischen den Gruppen (LPS10: $F=0,25$; $p=.862$; RCF-recall: $F=0,09$; $p=.963$). Zweifaktorielle Varianzanalysen weisen die deutlichen Gruppenunterschiede als Alterseffekte aus. Im LPS können 32 % der Varianz über diesen Faktor aufgeklärt werden ($F(1,54)=25,77$; $p<.001$; $\eta^2=.32$) und im RCF-recall immerhin noch 24 % ($F(1,54)=16,71$; $p<.001$; $\eta^2=.24$).

Reine Geschlechtseffekte oder Interaktionen mit dem Faktor „Altersgruppe“ wurden in den vorliegenden Daten nicht gefunden. Der LPS10 zeigte sich hier als besonders unauffällig: Für den Faktor „Geschlecht“ betrug die Prüfgröße $F(1,54)=0,14$ ($p=.707$) und für die Interaktion $F(1,54)=0,55$ ($p=.462$). Auch im RCF-recall konnte kein Interaktionseffekt nachgewiesen werden ($F(1,54)=0,12$; $p=.728$), aber eine geringe, wenn auch statistische nicht signifikante Differenz, zwischen männlichen und weiblichen Versuchspersonen ($F(1,54)=2,64$; $p=.110$).

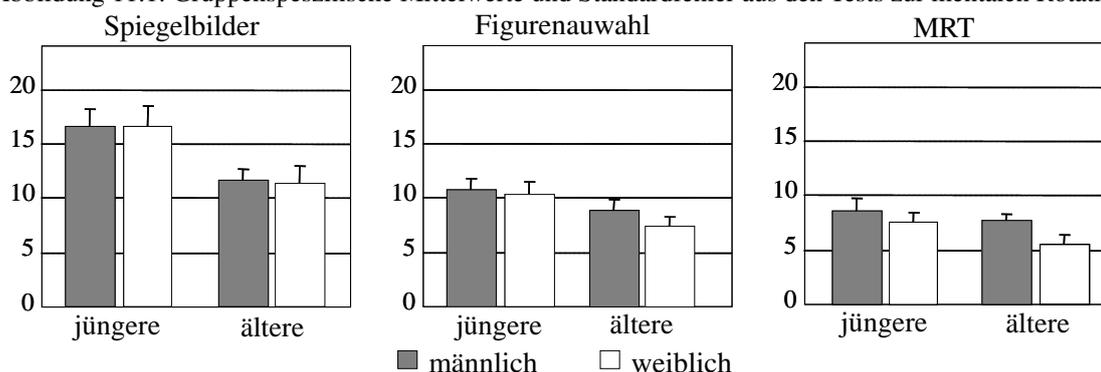
Die Ergebnisse des LPS10 verdeutlichten eine mittlere Differenz von 7,5 Punkten zwischen jüngeren und älteren Erwachsenen. Bei jüngeren Erwachsenen waren die über den Test operationalisierten Fähigkeiten der räumlichen Wahrnehmung höher ausgeprägt. Bezüglich des visuell-räumlichen Gedächtnisses ergab sich ein ähnlicher Unterschied mit einer mittleren Differenz von 6,1 Punkten im RCF-recall. Bei diesem Test ist möglicherweise die Testleistung in geringem Umfang auch

von der Geschlechtszugehörigkeit abhängig, so dass männliche Versuchspersonen etwas bessere Ergebnisse erzielten; die beobachtete mittlere Differenz betrug 2,4 Punkte.

Tests zur mentalen Rotation

Diese Raumvorstellungsfacette spielte bei der Lösung der Tests *Spiegelbilder* (SP), *Figurenauswahl* (FA) und MRT eine Rolle. Abbildung 11.1 stellt Gruppenmittelwerte der Summenscores aus den drei Tests dar. Es sind deutliche Differenzen zwischen älteren und jüngeren Erwachsenen und geringere Geschlechterunterschiede zu erkennen.

Abbildung 11.1: Gruppenspezifische Mittelwerte und Standardfehler aus den Tests zur mentalen Rotation.



Varianzhomogenität zwischen den Gruppen lag für alle drei Test vor, so dass Varianzanalysen durchgeführt wurden (SP: $F=1,69$; $p=.179$; FA: $F=0,43$; $p=.729$; MRT: $F=2,10$; $p=.111$). Zwischen den Faktoren „Altersgruppe“ und „Geschlecht“ ergaben sich in keiner der Analysen signifikante Interaktionseffekte, so dass im folgenden Mittelwertsunterschiede jeweils nur anhand eines Faktors dargestellt werden (SP: ($F(1,54)=0,012$; $p=.912$; FA: ($F(1,54)=0,22$; $p=.639$; MRT: ($F(1,54)=0,42$; $p=.519$)).

Alterseffekte zeigten sich insbesondere im Test *Spiegelbilder* und etwas geringer im Test *Figurenauswahl* (s. Tab. 11.4). Im Test *Spiegelbilder* lösten jüngere Erwachsene im Mittel 5,1 Items mehr als ältere Erwachsene. Dieser Unterschied ist signifikant und mit mehr als 0.8 Standardabweichungen als großer Effekt einzustufen ($\eta^2=0.17$). Mithin konnten durch die Zuordnung zu den Altersgruppen 17 % der Varianz aufgeklärt werden. Etwas geringer aber dennoch nachweisbar, war der Alterseffekt im Test *Figurenauswahl*. Hier betrug die Gruppenmittelwertsdifferenz mehr als 0.5 Standardabweichungen und gilt noch als Effekt mittlerer Größe. Im MRT war dieser Effekt zwar sichtbar, konnte aber als systematischer Einfluss des Faktors „Altersgruppe“ statistisch nicht abgesichert werden.

Wie Abbildung 11.1 bereits angedeutet, ergaben sich keine Geschlechterdifferenzen in den Test *Spiegelbilder* und *Figurenauswahl*. Männliche und weibliche Versuchspersonen erzielten im Mittel nahezu identische Scores. Tabelle 11.5 fasst die diesbezüglichen deskriptiven Maße und inferenzstatistischen Prüfgrößen zusammen. Nur im MRT zeigte sich eine Tendenz zu besseren Testleistungen für männliche Versuchspersonen. Der Unterschied ist nur mit einem Alpha-Fehler von 10 % stati-

Tabelle 11.4: Alterseffekte in Tests zur mentalen Rotation. Mittelwertesunterschiede zwischen jüngeren und älteren Erwachsenen.

		M	SD	F_(1;54)	p	Cohens d	η²
Spiegelbilder	jüngere	16,6	6,35	11,1**	.002	0.89	.17
	ältere	11,5	5,00				
Figurenauswahl	jüngere	10,5	3,90	5,38*	.024	0.62	.09
	ältere	8,1	3,90				
MRT	jüngere	8,0	4,01	2,43	.125	0.39	.04
	ältere	6,6	3,08				

** p<.01, * p<.05

stisch abgesichert und fiel mit $d < 0.5$ geringer aus als die erwarteten 0.7 Standardabweichungen. Da die Größe des Effekts von wenigstens $d = 0.7$ als empirisch gut abgesichert gelten kann, ist es möglich eine *post hoc*-Schätzung des hier in Kauf genommenen Beta-Fehlers durchzuführen. Bei einer Gruppenstärke von $N_{\text{männlich}} = N_{\text{weiblich}} = 29$ ergibt sich auf dem 95%-Signifikanzniveau bei einem erwarteten Effekt der Größe $d = 0.7$ eine Teststärke von 0.745. Mit einer 25%-igen Wahrscheinlichkeit konnte der Effekt also nicht gefunden werden.

Tabelle 11.5: Geschlechtseffekte in Tests zur mentalen Rotation. Mittelwertesunterschiede zwischen männlichen und weiblichen Versuchspersonen.

		M	SD	F_(1;54)	p	Cohens d	η²
Spiegelbilder	männlich	14,0	5,45	0,00	.949	0.02	.00
	weiblich	13,9	6,96				
Figurenauswahl	männlich	9,7	4,18	0,67	.418	0.20	.01
	weiblich	8,9	3,94				
MRT	männlich	8,1	3,43	3,28 ⁺	.076	0.48	.06
	weiblich	6,4	3,63				

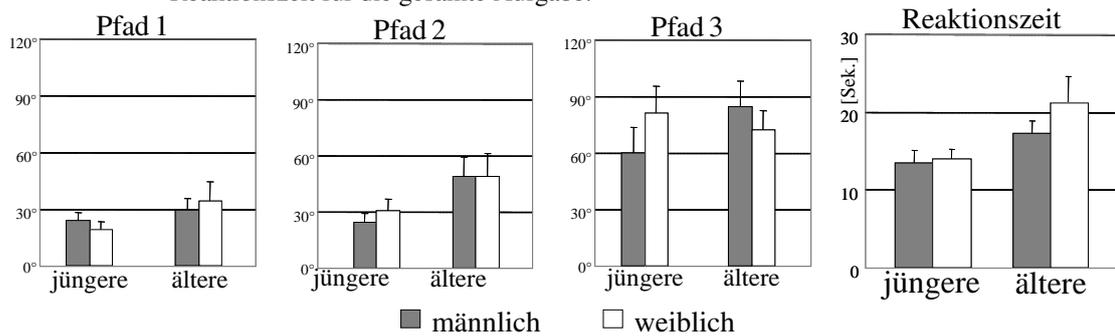
⁺ p<.10

11.3.2 Kompass-Zeigeaufgabe

Entsprechend den statistischen Auswertungen des Experiments 2 (s. Absch.10.2.2) werden die Richtungsschätzungen in der Kompass-Zeigeaufgabe als absolute Schätzfehler über Trials mit gleichen Pfaden gemittelt. Abbildung 11.2 zeigt die gruppenspezifischen mittleren Schätzfehler in Grad und die Reaktionszeit. Die Reaktionszeit wurden über alle Trials der Aufgabe gemittelt.

Wie bereits in Experiment 2 demonstriert, ergaben sich auch in dieser kleineren Stichprobe deutliche Unterschiede in den mittleren Schätzfehlern bezüglich der verwendeten drei Pfade. Für den

Abbildung 11.2: Mittlerer absolute Schätzfehler der Heimkehrrichtung für die Pfade 1 bis 3 und die mittlere Reaktionszeit für die gesamte Aufgabe.



Pfad 1 ergaben sich 27° ($SD=26^\circ$), für den Pfad 2 39° ($SD=36^\circ$) und für den Pfad 3 75° ($SD=49^\circ$) mittlere Abweichungen von der tatsächlichen Heimkehrrichtung. In einer Varianzanalyse mit Messwiederholung mit „Pfad“ als Innersubjektfaktor und „Altersgruppe“ und „Geschlecht“ als Zwischensubjektfaktoren wurden die Differenzen auf statistische Signifikanz geprüft. Die Überprüfung auf unkorrelierte Wiederholungsmessung ergab mit Mauchlys-Sphärizitätstest eine Verletzung dieser Voraussetzung ($W=0,84$; $p=.009$). Der Korrekturfaktor nach Greenhouse-Geisser wurde verwendet ($\epsilon=0,86$).

Die Prüfung der Voraussetzungen der Varianzanalyse ergab für den Gesamtfehler aus allen drei Pfaden im Levene-Test mit $F=0,98$ ($p=.411$) und im Kolmogorov-Smirnov-Test mit $Z=0,62$ ($p=.837$) keine Verletzung der Varianzhomogenität bzw. Normalverteilung. Lediglich auf den Faktorstufen „Pfad 1“ (Levene-Tests: $F=3,86$; $p=.014$; K-S-Tests: $Z=1,41$; $p=.037$) und „Pfad 2“ ($F=5,37$; $p=.003$) ergaben sich signifikante Abweichungen von diesen Voraussetzungen.

Für den Innersubjektfaktor „Pfad“ wurde ein großer Einfluss auf die Genauigkeit der Richtungs-schätzung festgestellt ($F(1,72;92,88)=25,54$; $p<.001$; $\eta^2=0.32$). Interaktionen mit den Zwischen-subjektfaktoren bestanden nicht (Altersgruppe: $F=0,62$; $p=.837$; Geschlecht: $F=0,05$; $p=.933$; Interaktion: $F=1,23$; $p=.293$). Ein Haupteffekt für den Faktor „Altersgruppe“ konnte ebenfalls als statistisch signifikant identifiziert ($F(1,54)=4,96$; $p=.030$), jedoch in seiner Bedeutsamkeit nur als geringe ($\eta^2=0.09$) eingeschätzt werden. Die absoluten Schätzfehler über alle Pfade gemittelt betragen für jüngere Erwachsenen 40° ($SD=21,3$) und für ältere 53° ($SD=23,4$). Geschlechtsdifferenzen und Interaktionseffekte zwischen den Faktoren konnten nicht nachgewiesen werden (Geschlecht: $F(1,54)=0,25$; $p=.620$; Interaktion: $F(1,54)=0,70$; $p=.407$).

Die Abbildung 11.2 verdeutlicht auch Unterschiede in den Reaktionszeiten, die sich auch in der Varianzanalyse als bedeutsame Differenzen zwischen den Gruppen herausstellten ($F(3,54)=2,92$; $p=.042$). Eine zweifaktorielle Varianzanalyse ergab auch hier einen Haupteffekt für den Faktor „Altersgruppe“ ($F(1,54)=6,85$; $p=.011$; $\eta^2=0.12$). Signifikante Mittelwertsunterschiede zwischen männlichen und weiblichen Versuchspersonen konnten weder als Haupteffekt noch in der Interaktion mit dem Alter gefunden werden (Geschlecht: $F(1,54)=1,15$; $p=.289$; Interaktion: $F(1,54)=0,69$; $p=.410$).

Eine Aufteilung der Stichprobe in jüngere und ältere Probanden offenbarte, dass durch den Al-

terseffekt bedingt insgesamt eine bimodale Verteilung resultiert. Da aber die Verteilungen dieser beiden Gruppen deutliche Varianzunterschiede aufwiesen, wurde die Signifikanz des Alterseffekts noch einmal mit dem für Varianzhomogenität korrigierten t-Test durchgeführt. Dieser ergab mit $t(43,58)=2,68$ ($p=.010$) ebenfalls einen statistisch bedeutsamen Unterschied, der mit $d=0.68$ als mittlerer Effekt eingestuft werden konnte. Praktisch betrug die mittlere Differenz aber nur wenige Sekunden. Während jüngerer Versuchspersonen im Mittel 13,8 Sekunden ($SD=5,04$ Sekunden) für das Einstellen der Kompassnadel benötigten, waren es bei älteren Probanden im Mittel 19,2 Sekunden ($SD=9,95$ Sekunden).

11.3.3 Water-Maze-Aufgabe

Training

Aus der Trainingsphase lagen von jedem Probanden als Verlaufsdaten Pfadlängen und Navigationszeiten vor. Zunächst wurden daraus Summenscores der Gesamttrainingstrecke und Trainingsdauer gebildet (s. Tab. 11.6). Der Quotient aus Trainingsstrecke und Dauer ergaben als Durchschnittsgeschwindigkeit ein Maß für die Effizienz der Navigation.

Tabelle 11.6: Gruppenunterschiede für die im virtuellen Training zurückgelegte Gesamtstrecke und Navigationszeit. Der Quotient aus beiden Maßen ergibt eine Durchschnittsgeschwindigkeit

		männlich		weiblich		F _(3;54)	p
		M	SD	M	SD		
<i>Pfadlänge:</i>							
[Pixel]	jüngere	56992	29878,4	54490	22500,2	0,80	.500
	ältere	62748	25883,9	70308	38738,8		
<i>Navigationszeit:</i>							
[Sekunden]	jüngere	217	125,6	296	75,6	14,08**	<.001
	ältere	414	94,5	517	199,3		
<i>Durchschnittsgeschwindigkeit:</i>							
[Pixel pro Sekunde]	jüngere	259	44,5	186	67,6	15,16**	<.001
	ältere	153	41,1	137	54,8		

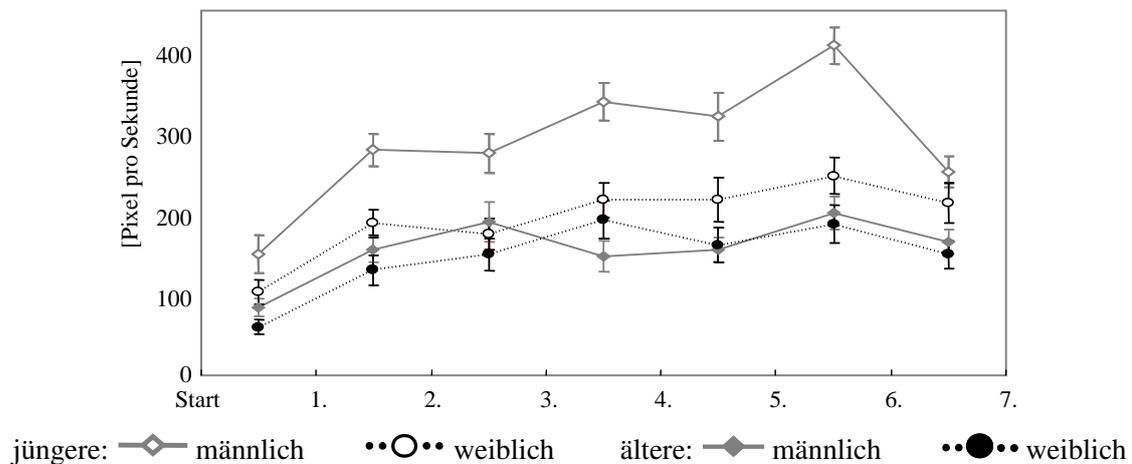
**p<.01

In Tabelle 11.6 ist für die abhängige Variable „Pfadlänge“ zu erkennen, dass keine alters- und geschlechtsspezifischen Unterschiede vorlagen. Aber hinsichtlich der Gesamttrainingszeit waren deutliche Alters- ($F(1,54)=35,32$; $p<.001$; $\eta^2=0.40$) und Geschlechtsdifferenzen ($F(1,54)=6,75$; $p=.012$; $\eta^2=0.11$) zu beobachten. Die Mittelwerte in der Tabelle zeigen, dass jüngere Männer die mit Abstand geringste Zeit in der Trainingsaufgabe verbringen (mittlere Zeit ca. 3,5 Minuten), gefolgt von jüngeren Frauen (ca. 5 Minuten), älteren Männern (ca. 7 Minuten) und älteren Frauen mit ca.

8,5 Minuten. Diese Angaben sind allerdings nur sehr grobe Abschätzungen für die individuellen Trainingszeiten, zumal sich die Varianzen zwischen den Gruppen z. T. beträchtlich unterscheiden (Levene-Test: $F=6,38$; $p=.001$).

In der Betrachtung der Gruppendifferenzen aus den Durchschnittsgeschwindigkeiten verdeutlichte sich der Vorteil der jüngeren Männer während des Trainings. Signifikante Unterschiede mit um 40 % bis 90 % höheren Geschwindigkeiten ergaben sich in Post-hoc-Paarvergleichen (Scheffé-Prozedur) ausschließlich zwischen der Gruppe der jüngeren Männer und allen drei anderen Gruppen. Die Voraussetzung der Varianzhomogenität der Gruppen war in diesem Fall erfüllt (Levene-Test: $F=1,29$; $p=.285$).

Abbildung 11.3: Gruppenspezifische Trends in der Navigationsgeschwindigkeit während des virtuellen Trainings. Gruppenmittelwerte und Standardfehler.



Zur genaueren Aufklärung dieses Unterschiedes wurden neben dem Summenscore die Veränderung der Durchschnittsgeschwindigkeit während des Trainings anhand der Verlaufsdaten betrachtet. Dazu sind die Geschwindigkeiten für alle Teilstrecken vom Startpunkt bis zum Öffnen der letzten Kiste berechnet, als Gruppenmittelwerte zusammengefasst und in Abhängigkeit von der zugehörigen Teilstrecke in einem Diagramm dargestellt worden (s. Abb. 11.3).

Tabelle 11.7: Parameter der linearen Regressionen mit Zeit als Prädiktor (indirekt über die Nummer der geöffneten Kiste).

		R^2	$F_{(1;5)}$	p	Konstante	Anstieg
jüngere	männlich	0.34	2,61	.167	201,0	21,7
	weiblich	0.66	9,84	.026	124,5	17,4
ältere	männlich	0.38	3,02	.143	113,8	10,7
	weiblich	0.47	4,44	.089	90,6	14,0

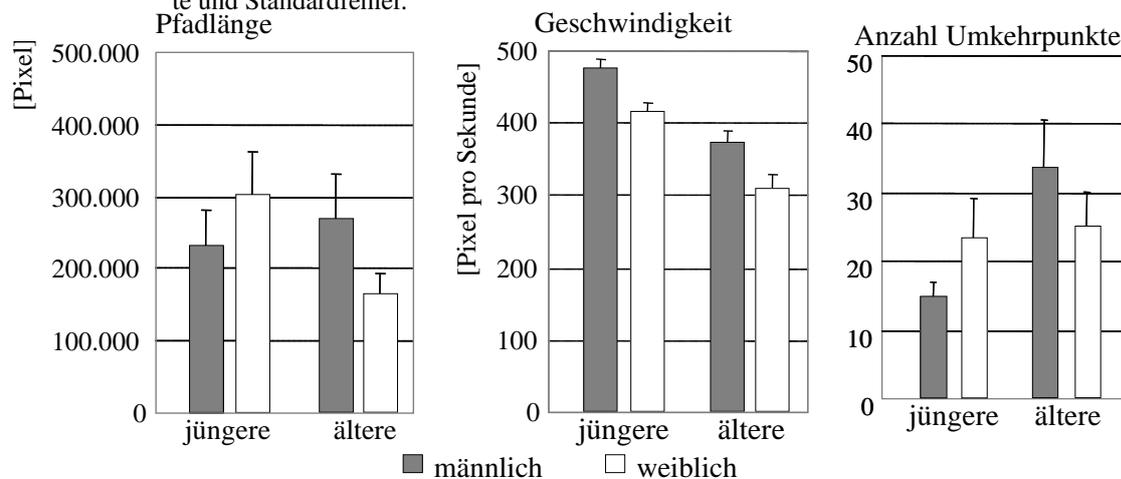
In jeder der Gruppen ist aus den Kurven der Abbildung 11.3 eine positive Entwicklung der Navigationsgeschwindigkeit während des Trainings zu erkennen. Lineare Kurvenanpassungen bestätigen teilweise diesen Trend in allen vier Gruppen (s. Tab. 11.7). Die positiven und von Null verschie-

denen Anstiege zeigen, dass im Mittel alle Gruppen während des Trainings ihre Effizienz steigern konnten. Eine Abschätzung des absoluten Zuwachses der Geschwindigkeiten ergab für weibliche Versuchspersonen unabhängig von der Altersgruppe eine 100-prozentige Steigerung in Bezug auf das mittlere Ausgangsniveau; ältere Männer lagen hier bei 66 % und jüngere bei 76 %.

Explorationsdurchgang

Abbildung 11.4 fasst die Verhaltensmaße durchschnittliche Geschwindigkeit, Pfadlänge und die Anzahl der Umkehrpunkte aus dem Explorationsdurchgang des VWM zusammen.

Abbildung 11.4: Verhaltensmaße aus dem Explorationsdurchgang des VWM. Gruppenspezifische Mittelwerte und Standardfehler.



Bei den Explorationsgeschwindigkeiten wiederholte sich das Bild, welches schon während des Trainings beobachtet wurde. Zur Durchführung einer Varianzanalyse wurden Varianzhomogenität (Levene-Test: $F=1,39$; $p=.256$) und Normalverteilung (K-S-Test: $Z=0,683$; $p=.739$) geprüft. Es gab deutliche Gruppenunterschiede für die mittlere Geschwindigkeit, mit der die Insel erkundet wurde ($F(3,54)=18,63$; $p<.001$). Jüngere Versuchspersonen bewegen sich am schnellsten durch die virtuelle Umgebung der Insel und liegen über dem Mittelwert der Gesamtstichprobe ($F(1,54)=40,83$; $p<.001$; $\eta^2=0.431$). Neben dem sehr großen Alterseffekt fand sich auch ein Haupteffekt für den Faktor „Geschlecht“ ($F(1,54)=15,04$; $p<.001$; $\eta^2=0.218$). Interaktionen zwischen beiden Faktoren bestanden für diese abhängige Variable nicht ($F(1,54)=0,00$; $p=.981$).

Die Analyse der Pfadlängen ergab keine signifikanten Unterschiede zwischen den Gruppen ($F=1,24$; $p=.304$). Im Mittel legten die Versuchspersonen 210.000 Pixel Wegstrecke während der Exploration zurück; das entspricht in etwa einer „realen“ Strecke von 700 Metern. Auffällig sind die Mittelwert- und Varianzunterschiede zwischen jüngeren und älteren Frauen ($t(18,66)=2,04$; $p=.056$; Levene-Test: $F=4,24$; $p=.049$).

Aus der Anzahl der Umkehrpunkte lässt sich ablesen, wie oft eine Versuchsperson während der Exploration die eingeschlagenen Richtung abrupt änderte. Auch in diesem Verhaltensmaß lagen deutliche Unterschiede in den Varianzen der Gruppen vor (Levene-Test: $F=5,90$; $p=.001$). Bei ge-

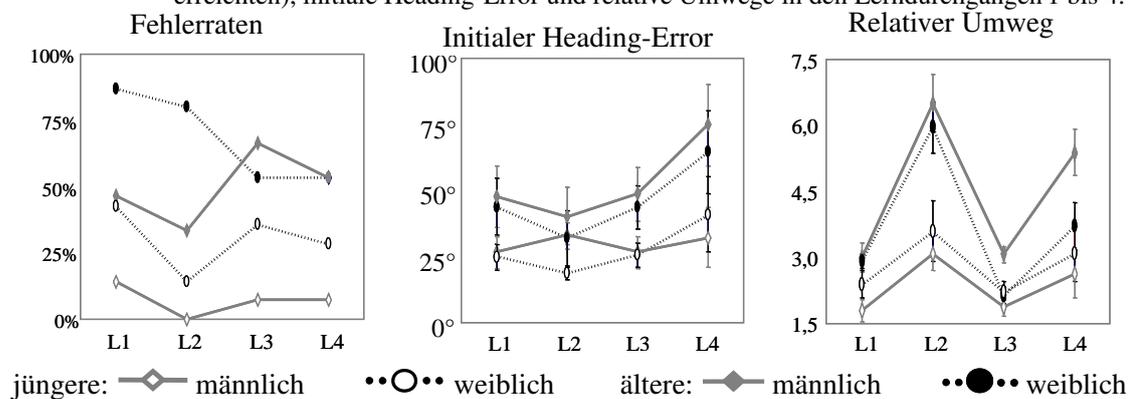
nauerer Betrachtung erkennt man, dass jüngere Männer in Mittelwert und Varianz sich von allen anderen Gruppen stark unterscheiden. Im Mittel hatten sie 14,9 Umkehrpunkte (SD=8,58) auf ihrem Explorationsweg. Alle anderen Gruppen hatten eine mittlere Zahl der Umkehrpunkte von 27,4 (SD=23,78). Der t-Test für heterogene Varianzen weist diesen Unterschied als signifikant aus ($t(54,95)=2,93; p=.005$).

Lerndurchgänge

In Abbildung 11.5 sind Ergebnisse der Lernphase für jeden der vier Durchgänge getrennt dargestellt. Unter Fehlerraten wurde in Abhängigkeit von den Faktoren „Altersgruppe“ und „Geschlecht“ die Anzahl der Personen prozentual zur Gesamtpersonenzahl einer Interaktionsgruppe bestimmt, die das Ziel im jeweiligen Durchgang nicht erreichten.

Aus der Abbildung wird deutlich, dass je nach Alter und Geschlecht die Lerndurchgänge unterschiedlich häufig erfolgreich bearbeitet wurden. Von den jüngeren männlichen Versuchspersonen erreichten in jedem Durchgang fast alle das Ziel. Die Fehlerkurve liegt hier nahe Null. Jüngere weibliche Versuchspersonen fanden im Vergleich häufiger das Ziel nicht. Jedoch waren in jedem Durchgang mehr als 50 % dieser Gruppe erfolgreich. Den älteren Versuchspersonen bereiten die Lerndurchgänge deutlich größere Schwierigkeiten. Besonders ältere Frauen konnten häufig das Ziel nicht erfolgreich ansteuern. Zu Beginn der Lernphase waren das ca. 90 % und am Ende noch über 50 % der Personen dieser Gruppe. Bei den älteren Männern schwankten die Fehlerraten zwischen dem ersten und vierten Durchgang um die 50 %.

Abbildung 11.5: Fehlerraten (Prozentzahl der Versuchspersonen innerhalb einer Gruppe, die das Ziel nicht erreichten), initiale Heading-Error und relative Umwege in den Lerndurchgängen 1 bis 4.



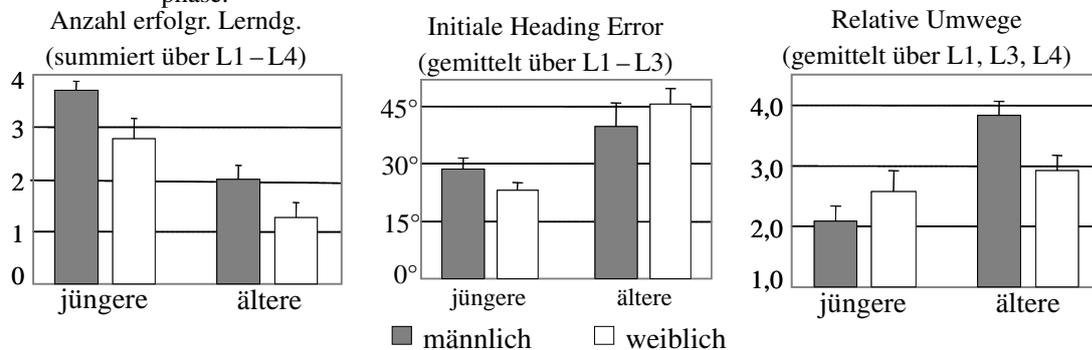
Neben Fehlerraten wurde die Performanz in den Lerndurchgängen durch die Verhaltensmaße des initialen Heading-Errors und des relativen Umwegs spezifiziert. Für den initialen Heading-Error, als Maß für die Exaktheit der zu Beginn eines Trials erinnerten Zielposition, ergab sich ein eher gleichbleibender Verlauf über alle Trials und Gruppen. Außer für Bedingung L4 schwankten diese Werte zwischen 30° bis 50°. Die Gruppenmittelwerte jüngere Versuchspersonen lagen hier in allen Durchgängen unterhalb derer der älteren Erwachsenen. Insbesondere jüngere Frauen konnten die initiale Zielrichtung im Mittel sehr genau einschätzen.

Die starke Abweichung im Durchgang L4 lag darin begründet, da der initiale Heading-Error nur Gültigkeit für Startpositionen vom Rand hat (s. Abschn. 6.3.3). Da aber in L4 aus der Mitte gestartet wurde, misst der initiale Heading-Error in dieser Bedingung nicht die initiale Zielreaktion. In den folgenden Betrachtungen wurden dann nur die initialen Heading-Error über die Bedingungen L1 bis L3 berücksichtigt.

Relative Umwege pro Trial wurden durch die Normierung der zurückgelegten Pfadlängen mit dem kürzesten Abstand zwischen Start- und Zielposition (Luftlinie) berechnet. Im entsprechenden Diagramm der Abbildung 11.5 ist zu erkennen, wie für die Durchgänge L2 und L4 dieses Maß sprunghaft anstieg. Besonders auffällig ist die hohe Alterssensitivität dieses Verhaltensmaßes in der Bedingung L2. Ältere Versuchspersonen legten hier signifikant längere, im Mittel doppelt so lange Wege, wie jüngere Versuchspersonen zurück ($F(1,54)=23,65; p<.001$). Offensichtlich führten besondere Startpositionen (L2 Start im Zielquadranten, L4 Start in der Mitte) zu einer stärkeren Erhöhung des Schwierigkeitsgrads als es in andere Versuchsbedingungen der Fall war.

Aus den Verhaltensmaßen der Lernphase wurden zur vereinfachten Darstellung der Alters- und Geschlechtsunterschiede Mittelwerte über die Durchgänge aggregiert. Die erfolgreichen Lerndurchgänge jeder Versuchsperson wurden zu einem Summenscore zusammengefasst, die relativen Umwege wurden über die Durchgänge L1, L2 und L4 gemittelt und die initialen Heading-Error über die Durchgänge L1 bis L3. Gruppenmittelwerte für die Interaktion „Altersgruppe“ und „Geschlecht“ der so aggregierten Variablen sind in Abbildung 11.6 dargestellt.

Abbildung 11.6: Summenscore, Mittelwerte und Standardfehler der Verhaltensmaße aus der gesamten Lernphase.



In den Gruppenmittelwerten des Summenscores der erfolgreich absolvierten Lerndurchgänge zeigten sich die Alters- und Geschlechtsunterschiede deutlicher. Abbildung 11.6 zeigt die signifikanten Mittelwertsunterschiede für diese abhängige Variable ($F(3,54)=13,45; p<.001$). Die Varianzen unterschieden sich ebenfalls zwischen den Gruppen (Levene-Test: 6,69; $df=3; p=.001$). Aufgrund der wenigen Kategorien der abhängigen Variable wurden Prüfgrößen und Signifikanztests mit nicht-parametrische Verfahren abgesichert. Die Prüfgröße über die Ränge des Kruskal-Wallis H-Test ($\chi^2=24,30; p<.001$) bestätigt die beobachteten Gruppenunterschiede als statistisch signifikant.

Auch die Geschlechtsdifferenzen können über den t-Test als bedeutsam identifiziert werden ($t(56)=2,33; p=.023$); bzw. über den U-Test ($Z=-2,28; p=.023$). Im Mittel absolvierten männliche Versuchspersonen 2,8 ($SD=1,20$) und weibliche 2,0 ($SD=1,49$) der vier Lerndurchgänge erfolgreich.

In Standardabweichungen ausgedrückt, konnte diese Differenz mit $d=0.62$ als mittlerer Effekt eingestuft werden.

Deutlicher fällt der Unterschied zwischen jüngeren und älteren Probanden aus. Während ältere Erwachsene im Mittel nur 1,6 (SD=1,10) der Trials mit dem Auffinden der Ziellokation abschließen konnten, waren es bei den jüngeren Versuchspersonen $M=3,3$ (SD=1,21). Die statistischen Prüfgrößen lauten $t(56)=5,34$ ($p<.001$) bzw. $Z=-4,36$ ($p<.001$). Da Varianzhomogenität zwischen den Altersgruppen ($F=0,54$; $p=0.466$) und Anpassung auf Normalverteilung ($Z=1,03$; $p=.24$) gewährleistet war, konnte mit Hilfe der Varianzanalyse der stichprobenspezifische Effekt für den Faktor „Altersgruppe“ bestimmt werden. Insgesamt konnten 37,1 % der Varianz aufgeklärt werden.

Der mittlere initiale Heading-Error differenzierte ebenfalls sehr stark zwischen den Altersgruppen ($F(1,54)=F=16,13$; $p<.001$; $\eta^2=0.23$). Dabei ergab sich ein signifikant niedrigeres Gruppenmittel für jüngere ($M=26^\circ$; $SD=9,9^\circ$) gegenüber älteren Versuchspersonen ($M=43^\circ$; $SD=19,9$). Die Werte verdeutlichen die im Mittel genauere Einschätzung der initialen Zielrichtung bei jüngeren, aber auch eine größere Variabilität dieser Fähigkeit bei älteren Erwachsenen. Es fanden sich dem zufolge charakteristische Unterschiede in den Streuungen für beide Gruppen für dieses Verhaltensmaß (Levene-Test: $F=6,64$; $p=.001$). Wegen der Varianzhomogenität wurde die statistische Bedeutsamkeit des Mittelwertsunterschieds noch einmal mit dem in den Freiheitsgraden korrigierten t-Test durchgeführt ($t(43,17)=-4,11$; $p<.001$). Geschlechts- und Interaktionseffekte wurden nicht gefunden ($F(1,54)=1,91$; $p=.172$ bzw. $F(1,54)=0,00$; $p=.99$).

Die gemittelten relativen Umwege lagen als normalverteiltes und varianzhomogenes Merkmal in der Stichprobe vor (K-S-Test: $Z=0,79$; $p=.557$; Levene-Test: $F=1,56$; $p=.210$). Der Alterseffekt konnte mit $\eta^2=0.21$ ($F=14,26$; $p<.001$) nachgewiesen werden. Der Faktor „Geschlecht“ konnte nicht interpretiert werden, da, wie in Abbildung 11.6 zu erkennen, ein disordinaler Interaktionseffekt vorlag ($F=6,10$; $p=.017$).

Testdurchgang 1

Zunächst wurde von dieser Testbedingung ausgezählt, wie viele der Versuchspersonen den Durchgang erfolgreich absolviert hatten und deren Anzahl getrennt nach Gruppenzugehörigkeit betrachtet. Tabelle 11.8 zeigt die entsprechende Kontingenztafel für die Faktoren „Altersgruppe“ und „Geschlecht“. Insgesamt fanden 78 % aller Versuchspersonen die versteckte Ziellokation in diesem Test. Diese verteilten sich auf die vier Gruppen mit einer Häufigkeit von 60 % bis 93 %. Der geringste Anteil erfolgreicher Versuchspersonen wurde innerhalb der Gruppe älterer männliche Erwachsener gefunden. Korrigiert-standardisierte Residuen wiesen allerdings nur eine tendenzielle Abweichung des beobachteten vom erwarteten Wert auf dieser Faktorstufe aus. Insgesamt bestand kein statistisch signifikanter Zusammenhang zwischen Gruppenzugehörigkeit und Lösungserfolg ($\chi^2=5,23$; $df=3$; $p=.156$).

Tabelle 11.8 verdeutlicht auch, dass zehn der insgesamt 13 nicht erfolgreichen Versuchspersonen ältere Erwachsene waren und somit die Altersgruppenzugehörigkeit hier statistisch bedeutsam war. Für

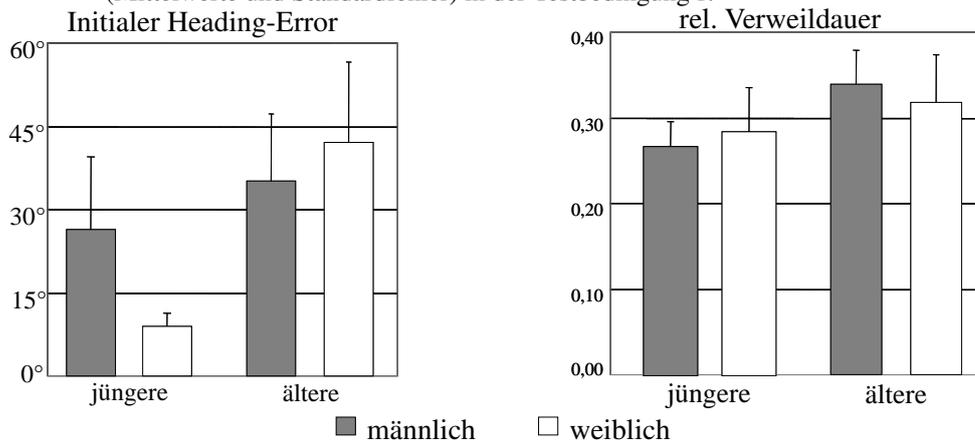
Tabelle 11.8: Kontingenztafel der erfolgreichen bzw. nicht erfolgreichen Personen im Testdurchgang 1.

	jüngerer		ältere		Gesamt
	männlich	weiblich	männlich	weiblich	
erfolgreich	13 (93%)	12 (86%)	9 (60%)	11 (73%)	45 (78%)
nicht erfolgreich	1 (7%)	2 (14%)	6 (40%)	4 (27%)	13 (22%)
korrigiert-standardisierte Residuen	1,6	0,8	-1,9	-0,5	

eine entsprechende Zusammenfassung der Zellen ergab sich folgende Prüfgröße $\chi^2=4,26$ ($p=.039$; $df=1$). Der korrigierte Kontingenzkoeffizient betrug $C_{\text{kor}}=0.368$.

Die initialen Heading-Error lagen in dieser Testbedingung im Mittel bei 29° ($SD=45,1$). Auffällig waren bei diesem Maß die großen Streuungen, die allerdings nicht in allen Gruppe auftraten (s. Abb. 11.7). Jüngere Frauen hatten im Vergleich zu allen anderen Gruppen den geringsten mittleren Fehlerwert von 9° und die geringste Standardabweichung von $9,0^\circ$. Außerdem wichen die Häufigkeiten der initialen Heading-Error in der Stichprobe deutlich von der Normalverteilung ab (K-S-Test: $Z=2,09$; $p<.001$).

Abbildung 11.7: Gruppenspezifische initiale Heading-Error und relative Verweildauern im Zielquadrant (Mittelwerte und Standardfehler) in der Testbedingung 1.



Der Gruppenunterschied war insgesamt statistisch signifikant ($\chi^2=10,80$; $df=3$; $p=.013$). Ebenso der Unterschied zwischen den Altersgruppen ($Z=3,06$; $p=.002$). Jüngere Versuchspersonen erzielten im Mittel einen Fehler von 18° ($SD=35,7$), ältere 39° ($SD=51,0$). Geschlechtsdifferenzen konnten wegen der Interaktion nicht nachgewiesen werden ($Z=0,18$; $p=.858$). Neben den Faktoren „Alter“ und „Geschlecht“ wurde bei der Auswertung der „Testerfolg“ als unabhängige Variable berücksichtigt. Bezüglich des initialen Heading-Errors ergab sich aber kein signifikanter Unterschied ($Z=0,51$; $p=.608$).

Die Verweildauer im Zielquadranten betrug im Mittel ca. ein Drittel der Gesamtnavigationszeit in diesem Durcjang ($M=0,3$; $SD=0,17$). Das Merkmal lag in der Stichprobe normalverteilt aber varianzinhomogen vor (K-S-Test: $Z=0,50$; $p=.968$; Levene-Test: ($F=3,19$; $p=.031$)). Es ergaben sich

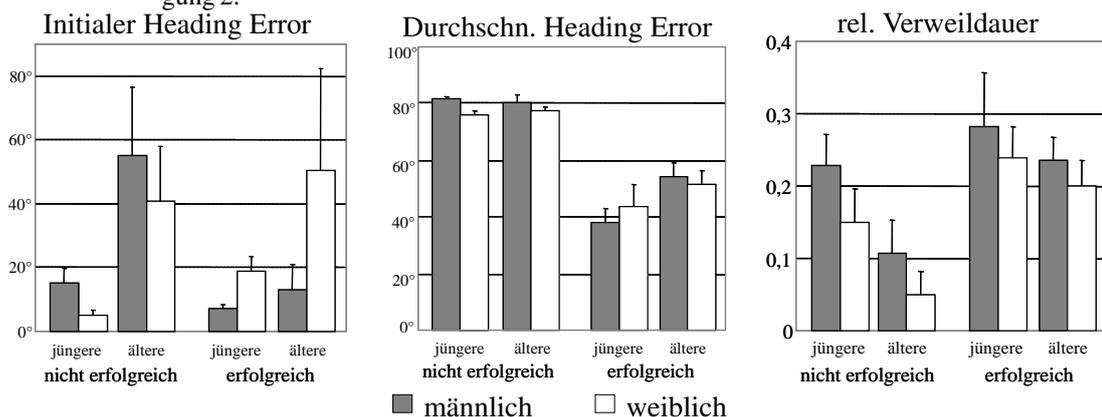
keine gruppenspezifischen Mittelwertsunterschiede ($F(3,54)=0,49; p=.694$). Tendenzielle Unterschiede bestanden bezüglich des Faktors „Testerfolg“ ($t(56)=-1,99; p.052$). Personen, die das Ziel nicht erreichten, verbrachten ca. 20 % der Navigationszeit im Zielquadranten ($M=0,22; SD=0,19$). Für erfolgreiche Probanden betrug dieser Wert mehr als 30 % ($M=0,33; SD=0,16$).

Testdurchgang 2

In der zweiten Testbedingung ergab sich folgendes Bild: 36 (62%) aller Versuchspersonen konnten das Ziel ohne Landmarken nicht auffinden. Lediglich 22 (38%) Probanden waren in der Lage, in diesem Test erfolgreich zu navigieren. Die Zusammenfassung der Erfolgshäufigkeiten für alle vier Gruppen ergab keinerlei Zusammenhang zwischen den Faktoren ($\chi^2=0,68; df=3; p=.878$).

Auch ein Zusammenhang zum Lösungserfolg der vorangegangenen Testbedingungen konnte nicht hergestellt werden ($\chi^2=0,48; p=.488; df=1$). Von den 48 zuvor erfolgreichen Probanden fanden nur 16 (36%) das Ziel im Testdurchgang 2. Von den 13 Versuchspersonen, die zuvor nicht erfolgreich waren, fanden sieben (54%) das Ziel.

Abbildung 11.8: Gruppenspezifische Verhaltensmaße (Mittelwerte und Standardfehler) in der Testbedingung 2.



In Abbildung 11.8 sind Gruppenmittelwerte und Standardfehler als weiteres Maß zur Quantifizierung der Navigationsleistung im Testdurchgang 2 dargestellt. „Testerfolg“ wurde hier als dritter Gruppenfaktor in die Darstellung mit aufgenommen.

Die initialen Heading-Error lagen auch in diesem Durchgang nicht normalverteilung vor (K-S-Test: $Z=2,02; p=.001$). Signifikante Gruppenunterschiede konnten hier nur zwischen den Altersgruppen festgestellt werden ($Z=-2,06; p=.039$). Jüngere Versuchspersonen starteten den Durchgang mit einem mittleren Richtungsfehler von 11° ($SD=10,9^\circ$). Ältere Versuchspersonen überschätzen sich dagegen um 40° ($SD=53,1^\circ$) im Mittel. Mittelwertsunterschiede zwischen erfolgreichen und nicht erfolgreichen Probanden bestanden nicht ($Z=0,04; p=.968$). Auch Geschlechtsdifferenzen bestanden nicht ($Z=0,09; p=.932$). Mögliche Interaktionen der drei Faktoren wurden mit den nicht-parametrischen Verfahren nicht weiter untersucht.

Bezüglich der durchschnittlichen Heading-Error unterschieden sich insbesondere erfolgreiche von

nicht erfolgreichen Versuchspersonen. Probanden, die das Ziel nicht erreichten, hatten im Mittel einen durchschnittlichen Heading-Error von 79° ($SD=5,8$). Erfolgreiche Personen erzielten Werte um 48° ($SD=13,5$). Die Aufteilung nach Testerfolg ermittelte somit zwei Gruppen extremer Mittelwertsunterschiede. Der Kolmogorov-Smirnov-Tests ergaben eine gute Anpassung an Normalverteilungen, innerhalb jeder dieser extremen Gruppen (nicht erfolgreich: $Z=0,90$; $p=.392$; erfolgreich: $Z=0,81$; $p=.523$). Die Größe des Effekts des Testerfolgs betrug $\eta^2=0,77$ ($F(1,50)=171,56$; $p<.001$).

Innerhalb der Gruppe der nicht erfolgreichen Versuchspersonen erzielten alle Untergruppen einen ähnlichen Mittelwert von 79° . In Bezug auf erfolgreiche Personen bestand nur ein geringer Haupteffekt für den Faktor „Altersgruppe“ ($F(1,50)=5,61$; $p=.022$; $\eta^2=0,10$). Jüngere erfolgreiche Versuchspersonen wichen im Mittel um 41° ($SD=13,7^\circ$) und ältere um 53° ($SD=11,2^\circ$) von der optimalen Route ab.

Die Messwerte der relativen Verweildauer aus der Testbedingung 2 lagen insgesamt normalverteilt und varianzhomogen vor, so dass Gruppenunterschiede in einer mehrfaktoriellen Varianzanalyse untersucht wurden (K-S-Test: $Z=1,12$; $p=.160$ Levene-Test: $F=0,21$; $p=.891$). Insgesamt zeigte sich der stärkste Unterschied zwischen erfolgreichen und nicht erfolgreichen Versuchspersonen ($F(1,50)=10,85$; $p=.002$; $\eta^2=0,18$). Die 36 Personen, die das Ziel nicht erreichten, verbrachten im Mittel nur 13 % der gesamten Zeit im Zielquadranten ($M=0,13$; $SD=0,14$). Erfolgreiche Probanden verbrachten dagegen 24 % der Gesamtzeit ($M=0,24$; $SD=0,10$) in der Nähe des Ziels.

Unabhängig davon, ob das Ziel erreicht wurde, unterschieden sich jüngere und ältere Versuchspersonen signifikant ($F(1,50)=5,58$; $p=.022$; $\eta^2=0,10$) männliche und weibliche wenigstens tendenziell ($F(1,50)=2,66$; $p=.109$). Signifikante Interaktionen bestanden zwischen keinen der Faktoren, obwohl in der Gruppe der nicht erfolgreichen Versuchspersonen Alters- und Geschlechtsdifferenzen nominal stärker vorlagen. Die Interpretation der Interaktionen von drei Faktoren war hier nur bedingt möglich, da die Anzahl der Versuchspersonen auf den jeweiligen Faktorstufen in drei Fällen lediglich fünf Personen betrug.

11.3.4 Zusammenhänge zwischen den Maßen

Abschließend zu den quantitativen Auswertungen der Water-Maze-Aufgabe wurden Produkt-Moment-Korrelationen zwischen ausgewählten Testwerten und den Leistungsmaßen des selbst entwickelten Verfahrens bestimmt. Dies waren die Summenscores aus dem Test *Spiegelbilder* (SP), dem MRT, dem Recall des *Rey-Oesterrieth-Complexfigur-Test* (RCF-recall) und dem Test zur Feldabhängigkeit (LPS10). Für das Konstrukt Pfadintegration wurde der Mittelwerte der Schätzfehler aus Bedingungen mit den Pfaden 1 und 2 und der mittlere Schätzfehler für Pfad 3 verwendet. Die umfangreichen Maße aus der Water-Maze-Aufgabe wurden auf das Aggregat der erfolgreichen Lerndurchgänge (L1 bis L4) und die relativen Verweildauern in den Testbedingungen 1 und 2 reduziert.

Die stärksten Zusammenhänge ergaben sich zwischen den „traditionellen“ Raumvorstellungstests (s. Tab. 11.9). Diese lagen mit Koeffizienten zwischen $r=0,40$ und $r=0,57$ alle im Bereich mittlerer

Tabelle 11.9: Korrelationen ausgewählter raumkognitiver Maße.

	Mentale Rotation		visuell-räumliches Gedächtnis	Feldabhängigkeit	Pfadintegration		Place learning		
	SP	MRT	RCF-recall	LPS10	Pfad 1+2	Pfad 3	L1 - L4	T1	T2
SP	–	.57**	.48**	.56**	-.27*	-.12	.30*	-.20	.25+
MRT		–	.41**	.40**	-.13	-.30*	.26+	.07	.19
RCF-recall			–	.51**	-.22+	.02	.30*	.08	.40**
LPS10				–	-.25+	-.13	.39**	-.20	.17
Pfad 1+2					–	.06	-.32*	.05	-.09
Pfad 3						–	-.11	.14	-.17
L1 – L4							–	.22+	.28*
T1								–	-.06
T2									–

**p<.01, *p<.05, +p<.1

und großer Korrelationen. Der deutlichste Zusammenhang besteht mit $r=0.57$ zwischen den beiden Tests zur mentalen Rotation. Ähnlich hohe Werte ergaben sich in den Korrelation mit der Leistung im LPS10.

Die Korrelationen der Summenscores aus den Raumvorstellungstests mit den Schätzfehler der Zeigeaufgabe sind in Tabelle 11.9 unter den Spalten Pfadintegration zusammengefasst. Negative Koeffizienten zeigen hier einen positiven Zusammenhang zwischen der Leistung in der Zeigeaufgabe und den Leistungen in den Raumvorstellungstests an. Für die Genauigkeit der Schätzungen in Testbedingungen mit einfachen Pfaden (Pfad 1 und 2) ergaben sich insgesamt geringe Zusammenhänge.

Der Test *Spiegelbilder* korrelierte mit $r=-0.27$ noch am stärksten mit dem mittleren Schätzfehler. Die geringste Korrelation ergab sich mit dem MRT. Das gilt allerdings nicht für Bedingungen mit dem Pfad 3. Die einzige statistisch signifikante Korrelation bestand hier zum schwierigsten im Experiment verwendeten Raumvorstellungstest, dem MRT. Mit $r=-0.30$ lag ein mittlerer Zusammenhang zwischen beiden Testleistungen vor.

Bezüglich der Anzahl erfolgreich absolvierter Lerndurchgänge (L1 bis L4) in der Water-Maze-Aufgabe ergaben sich insgesamt zu den Raumvorstellungstests und zur Kompass-Zeigeaufgabe mittlere Zusammenhänge. Die Leistung im Testdurchgang 1 korrelierte nur gering mit der Anzahl der zuvor erfolgreich absolvierten Lerndurchgänge.

In der Testbedingung 2 ergaben sich Zusammenhänge zwischen der Verweildauer im Zielquadranten und dem Summenscore im Test zum visuelle-räumlichen Gedächtnis. Der Koeffizient lag mit $r=0.40$ in einer Größenordnung, wie sie innerhalb dieser Stichprobe nur zwischen den Raumvorstellungstests selbst gefunden wurden. Kein signifikanter Zusammenhang bestand zwischen der relativen Verweildauern und der durch die Kompass-Zeigeaufgabe (Pfad 1 und 2) gemessene Pfadintegrationsleistung.

Ergänzend wurden dazu Korrelationen in den nach Testerfolg aufgeteilten Subgruppe berechnet. Nur für Versuchspersonen, die Testdurchgang 2 nicht erfolgreich absolviert hatten, korrelierten der initiale Heading-Error mit dem über Pfad 1 und 2 gemitteltem Schätzfehler signifikant ($r=0.32$; $p=.058$; $N=36$). Bei erfolgreichen Versuchspersonen bestanden keine systematischen Zusammenhänge zwischen den Schätzfehlern aus beiden Aufgaben. Auch die anhand der Schätzfehler aus Pfad 3 getroffene Klassifizierung der Probanden stand hier in keinem statistischen Zusammenhang mit dem Erfolg in Testbedingung 2 der Water-Maze-Aufgabe ($\chi^2=0,76$; $p=.684$; $df=2$).

Ebenfalls ergänzend wurden die größeren Korrelationen aus Tabelle 11.9 hinsichtlich der Vermittelten Eigenschaft der Faktoren „Altersgruppe“ und „Geschlecht“ überprüft. Partial-Korrelationen mit Geschlecht als Kontrollvariable blieben im Wesentlichen zu denen in der Tabelle angegebenen Koeffizienten unverändert. Mit der Kontrollvariable Alter blieb nur eine der betrachteten Koeffizienten statistisch bedeutsam, die Partial-Korrelation zwischen dem Summenscore des RCF-delay und der relativen Verweildauer im Zielquadranten in Testbedingung 2 betrug dann $r=0.304$ ($p=.022$).

11.4 Muster- und Klassenanalysen

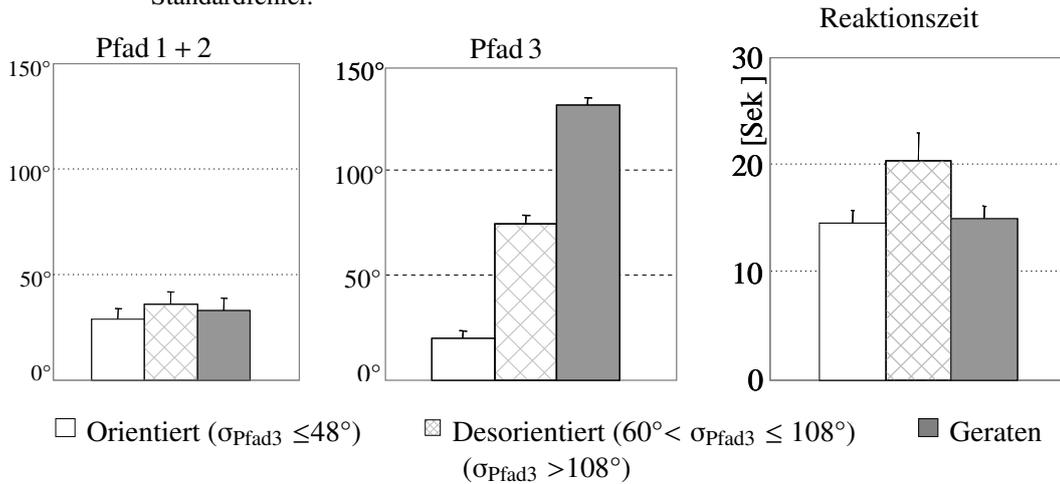
11.4.1 Desorientierungseffekt der Kompass-Zeigeaufgabe

Der im Experiment 2 beschriebene Desorientierungseffekt in Versuchsbedingungen mit Pfad 3 wurde hier ebenfalls zu einer Klassifizierung der Versuchspersonen eingesetzt. Dazu wurden, bezugnehmend auf die Ergebnisse in Abschnitt 10.2.3, drei gleichgroße Subgruppen anhand der 33%- und 66%-Perzentilwerte des Schätzfehlers aus Pfad 3 ($\sigma_{\text{Pfad 3}}$) gebildet. Abbildung 11.9 zeigt Mittelwerte der Schätzfehler und mittlere Reaktionszeit in den so bestimmten Subgruppen. Die Cut-Off-Werte zur Klassifizierung waren $\sigma_{\text{Pfad 3}}=48^\circ$ bzw. $\sigma_{\text{Pfad 3}}=108^\circ$. Versuchspersonen in der Gruppe „Orientiert“ erzielten in Testbedingungen mit dem Pfad 3 die besten Ergebnisse mit einem mittleren Schätzfehler von lediglich 20° ($SD=16,1$). In der Gruppe „Desorientiert“ lagen die mittleren Fehler bei 75° ($SD=19,3$) und in der Gruppe „Geraten“ bei 132° ($SD=16,6$).

Obwohl die Unterschiede in Schätzfehlern des Pfad 3 zwischen den Klassen sehr große waren, unterschieden sie sich nicht in den Schätzfehlern aus Bedingungen, in denen keine Abzweigungen größer als 90° verwendet wurden ($F(2,55)=0,41$; $p=.665$). Abbildung 11.9 zeigt dazu die über Pfad 1 und Pfad 2 gemittelten Schätzfehler. Dieser betrug für alle drei Subgruppen im Mittel 33° ($SD=23,4$). Der Levene-Test zeigte einen signifikanten Varianzunterschied an ($F=3,43$; $p=.040$). Unterschiede in den Reaktionszeiten sind zwar in der rechten Abbildung 11.9 zuerkennen, waren aber statistisch nicht signifikant. Die Prüfgröße des Levene-Tests deutete mit $F=5,13$ ($p=.009$) auf inhomogene Varianzen zwischen den drei Klassen. Die bezüglich dieser Verletzung robuste F-Statistik des Welch-Tests ergab für die Mittelwertsunterschiede keine signifikantes Resultat ($F(2,34,21)=1,93$; $p=.161$).

Zusammenhänge zwischen dieser Klasseneinteilung und der Altersgruppen- bzw. Geschlechtszugehörigkeit stellen die Kontingenztafeln der Tabellen 11.10 und 11.11 dar. Für den Faktor „Altersgruppe“ wurden signifikante Abweichungen der beobachteten Häufigkeiten von den Erwartungs-

Abbildung 11.9: Mittlere Schätzfehler für Bedingungen mit Desorientierungseffekt (Pfad 3) und ohne (Pfad 1+2). Mittlere Reaktionszeiten über alle Bedingungen. Dargestellt sind Mittelwerte und Standardfehler.



werten gefunden ($\chi^2=8,58$; $p=.014$; $df=2$). Der Kontingenzkoeffizient betrug $C=0.360$. Aus der Tabelle 11.10 wird deutlich, dass sich insbesondere die Besetzungen der Gruppen „Orientiert“ und „Desorientiert“ deutlich unterschieden. Die korrigiert-standardisierten Residuen wiesen eine signifikant überproportionale Besetzung der Gruppe „Orientiert“ mit älteren Versuchspersonen aus. 50 % der älteren Erwachsenen gehören also zu dem Drittel der Stichprobe, für das ein mittlerer Schätzfehler von 75° im Pfad 3 ermittelt wurde. Jüngere Erwachsene befanden sich dagegen zu 46 % in der Gruppe „Orientiert“, deren Versuchspersonen sich durch einen besonders geringen Schätzfehler in dieser Versuchsbedingung auszeichneten.

Tabelle 11.10: Verteilung jüngerer und älterer Versuchspersonen auf die nach dem Schätzfehler für Pfad 3 festgelegten Subgruppen.

	Orientiert (N=20)	Desorientiert (N=19)	Geraten (N=19)	Gesamt (N=58)
jüngere	13 (46,4%)	4 (14,3%)	11 (39,3%)	28 (100%)
ältere	7 (23,3%)	15 (50,0%)	8 (26,7%)	30 (100%)
korrigiert-standardisierte Residuen	1,9	-2,9	1,0	

In Tabelle 11.11 ist eine ähnliche Verteilung der Zellhäufigkeiten für den Faktor „Geschlecht“ zu erkennen. Die Prüfgröße der χ^2 -Statistik wies signifikante Unterschiede zwischen den tatsächlichen und den erwarteten Belegungen der Klassen mit weiblichen und männlichen Personen aus ($\chi^2=6,54$; $p=.038$; $df=2$; $C=0.32$). Eine Zuweisung von 48 % aller weiblichen Probanden zur Gruppe „Desorientiert“ und ca. 45% aller männlichen zur Gruppe „Orientiert“ wurde durch die Residuen als bedeutsame Verschiebung der Häufigkeiten auf diesen Faktorstufen ausgewiesen.

Bezüglich der Gruppe „Geraten“ ergaben sich sowohl für den Faktor „Geschlecht“ als auch für den Faktor „Altersgruppe“ keine bedeutsamen Differenzen zu den erwarteten Häufigkeiten. Eine Kombination beider Merkmale unter Berücksichtigung der Klasseneinteilung konnte nicht sinnvoll

Tabelle 11.11: Verteilung männlicher und weiblicher Versuchspersonen in der Klasseneinteilung.

	Klasse 1	Klasse 2	Klasse 3	Gesamt
männlich	13 (44,8%)	5 (17,2%)	11 (37,9%)	29 (100%)
weiblich	7 (24,1%)	14 (48,3%)	8 (27,6%)	29 (100%)
korrigiert-standardisierte Residuen	1,7	-2,5	0,8	

für eine χ^2 -Statistik verwendet werden, da weit mehr als 20 % der Zellen der Kreuztabelle Nullwerte enthielten.

11.4.2 Lösungsstrategien des Mental Rotations Test

Zum MRT wurden ebenfalls, wie bereits im Experiment 1 dargestellt, Musteranalysen auf der Ebene dichotomer Itemantworten durchgeführt. In der Stichprobe dieses Experiments wurde die NMF-Klassifizierungsmethode mit den Parametern $N=800$, $\lambda = 0.1$ und 20 Läufen pro erprobte Clustergröße k eingesetzt. Nach den vorgeschlagenen Modellselektionskriterien ergaben sich Lösungen für $k = 2$ bzw. $k=4$ als optimal (s. Abb. 11.10).

Abbildung 11.10: Modellselektion für die Musteranalyse beim MRT. (a) φ -k-Diagramm für erprobte Modellparameter $k=2$ bis $k=5$ (b) Consensus-Plots für $k=2$ und $k=4$.

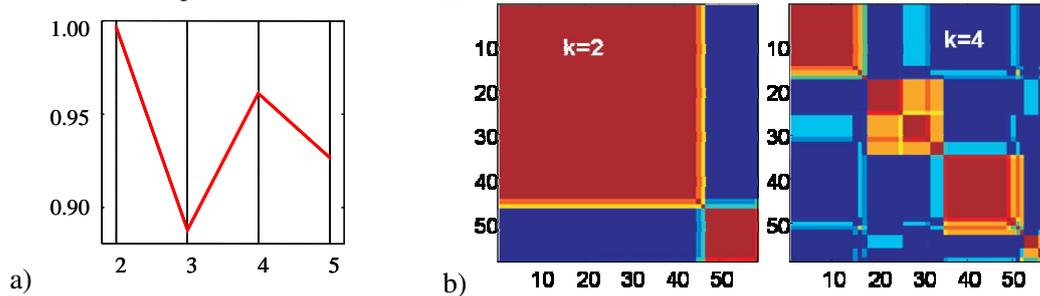


Tabelle 11.12: Varianzaufklärung in den itemspezifischen Lösungshäufigkeiten durch das 2-Klassen-Modell.

	Klasse 1 (N=46; 79 %)		Klasse 2 (N=12; 21 %)		p	η^2
	M	SD	M	SD		
Typ-I	0,25	0,174	0,27	0,219	.683	0.00
Typ-II	0,35	0,161	0,55	0,169	<.000	0.20
mit Okkl.	0,23	0,168	0,28	0,161	.382	0.01
ohne Okkl.	0,31	0,158	0,44	0,187	.019	0.09

Zur Unterscheidung zwischen holistischen und analytischen Strategien im MRT ist eine gute Differenzierung der Daten anhand der gelösten Itemtypen notwendig. In der Gegenüberstellung der Varianzaufklärung bezüglich dieser Kriterien schnitt die 4-Klassenlösung besser ab als die 2-Klassenlösung. Tabellen 11.12 und 11.13 stellen die für die Itemtypen akkumulierten Lösungshäufigkeiten

bezüglich der jeweiligen Klassen dar. Besonders wichtig war eine Differenzierung der Klassen anhand der Typ-I-Items, da diese zur Diskrimination zwischen holistischen und analytische Strategien genutzt werden.

Die Wahl der 4-Klassenlösung wird mit der besonders hohen Varianzaufklärung von 45 % bezüglich der zur Differenzierung der räumlichen Präferenzen relevanten Typ-I-Items begründet. In diesem Modell gibt es offensichtlich eine Subgruppe (Klasse 2), die sich gegenüber allen anderen durch eine hohe Anzahl (41 %) gelöster Spiegelitems auszeichnet. Die Häufigkeiten für Typ-II-Items und Items ohne Okklusion sind für Versuchspersonen dieser Klasse ähnlich hoch. Bezüglich okkludierter Items nehmen die Häufigkeiten wie in allen anderen Klassen ab, liegen jedoch immer noch signifikant darüber. Auf Grund dieses Verhaltensmusters kann vermutet werden, dass Personen der Klasse 2 eine holistische Lösungsstrategie einsetzten.

Tabelle 11.13: Varianzaufklärung in den itemspezifischen Lösungshäufigkeiten durch das 4-Klassen-Modell.

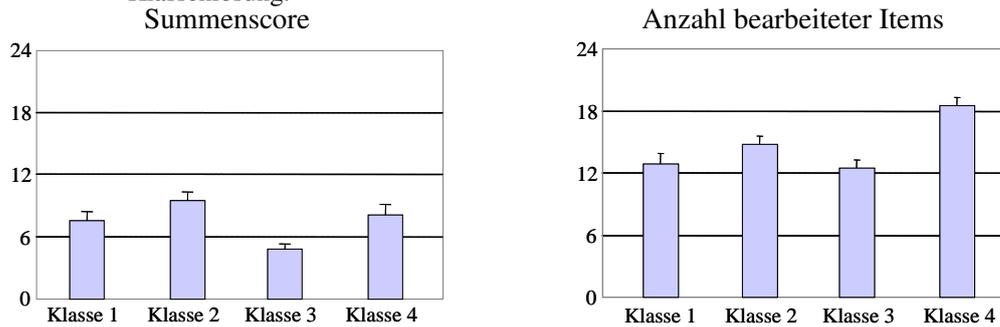
	Klasse 1 (N=17; 29%)		Klasse 2 (N=17; 29%)		Klasse 3 (N=18; 31%)		Klasse 4 (N=21; 36%)		p	η^2
	M	SD	M	SD	M	SD	M	SD		
Typ-I	0,28	0,156	0,41	0,161	0,10	0,088	0,18	0,144	.000	0.45
Typ-II	0,37	0,198	0,42	0,187	0,33	0,133	0,53	0,176	.089	0.11
mit Okkl.	0,30	0,171	0,30	0,171	0,14	0,120	0,23	0,123	.007	0.20
ohne Okkl.	0,32	0,176	0,44	0,171	0,24	0,110	0,39	0,128	.002	0.24

Ein ähnliches Profil, wenn auch in geringerer Fähigkeitsausprägung, zeigten Personen der Klasse 1. Sie konnten Typ-II-Items und Items mit Okklusion ähnlich gut lösen wie Personen der Klasse 2. Für Typ-I-Items lagen sie deutlich darunter, aber noch über den Häufigkeiten der Klassen 3 und 4. Items, die analytische Strategien erfordern, konnten von diesen Personen im Mittel in 28 % der Fälle gelöst werden.

Personen der Klasse 4 zeichneten sich durch eine besonders hohe Anzahl gelöster Typ-II-Items (53 %) und eine sehr geringe Anzahl von Typ-I-Items (18 %) aus. Diese Tatsache legte den Schluss nahe, dass Personen der Klasse 4 die MRT-Items durch Detailvergleiche bearbeiten, also eher analytisch vorgehen. Auffällig sind auch die gegenüber den Klassen 1 und 2 vorliegenden geringeren Häufigkeiten für okkludierte Items. Diese Profile zeigen auch Personen der Klasse 3, allerdings mit den von allen Personen im Mittel geringsten Häufigkeiten für alle Itemtypen. Versuchspersonen, die dieser Klasse zugeordnet wurden, hatten offensichtlich mit dem Test generell große Schwierigkeiten. Sie neigten aber eher noch dazu, Items des Typs II, also analytisch lösbare, richtig zu beantworten.

Die Auswirkung der durch die Klassifizierung identifizierten Strategien auf die Gesamtestleistung ist in Abbildung 11.11 dargestellt. Als Maß wurde der Summenscore und die Anzahl der bearbeiteten Items in den vier Klassen untersucht. Levene-Tests zeigten die Varianzhomogenität innerhalb der Subgruppen bezüglich der verwendeten Maße an (Summenscore: $F=1,02$; $p=.393$; Anzahl

Abbildung 11.11: Gesamtleistung im MRT (Summenscore und Anzahl bearbeiteter Items) bezüglich der 4-Klassenlösung.



bearbeiteter Items: $F=0,90$; $p=.448$). Mittelwertsunterschiede zwischen den Strategieklassen waren sowohl für den Summenscore als auch für die Anzahl der bearbeiteten Items signifikant (Summenscore: $F(3,54)=6,35$; $p=.001$; Anzahl bearb. Items: $F(3,54)=5,06$; $p=.004$). Die Varianzaufklärung durch die Klasseneinteilung betrug mehr als 20 % ($\eta^2=0.26$ bzw. $\eta^2=0.22$).

Post-hoc-Paarvergleiche ergaben einen signifikanten Mittelwertsunterschied von 4,6 Punkten zwischen den Klassen 2 und 3 bezüglich des Summenscores ($p=.001$). In der Anzahl der bearbeiteten Items unterschieden sich die Klassen 1 und 4 und die Klassen 3 und 4 um jeweils 5,6 bzw. 6,1 Items signifikant ($p=.019$ bzw. $p=.009$). In der Klasse 4 wurde bezüglich dieses Maßes mit $M=18,5$ ($SD=2,07$) im Mittel der höchste Wert, aber nicht der höchste Summenscore erreicht. Den erreichten Personen der Klasse 2 ($M=9,5$; $SD=3,68$) mit im Mittel weniger bearbeiteten Items ($M=14,8$; $SD=3,27$). Die geringste Gesamtleistung erbrachten Personen der Klasse 3 mit 4,8 Punkten ($SD=2,18$ Punkte) im Summenscore und 12,4 ($SD=3,65$) beantworteten Items.

Tabelle 11.14: Zusammenhänge Klasseneinteilung MRT und Geschlecht.

	Klasse 1	Klasse 2	Klasse 3	Klasse 4
männlich	10	10	7	2
weiblich	7	7	11	4
korrigiert-standardisierte Residuen	0,87	0,87	-1,14	-0,86

Zur Untersuchung der Alters- und Geschlechtsdifferenzen bezüglich der Klasseneinteilungen aus der MRT-Itemanalyse stellen die Tabellen 11.14 und 11.15 die entsprechenden kategorialen Zusammenhänge dar. Zwischen weiblichen und männlichen Versuchspersonen ergaben sich keine statistisch signifikanten Ungleichverteilungen auf die 4 MRT-Klassen ($\chi^2=2,61$; $df=3$; $p=.455$).

In den Altersgruppen ergaben sich zumindest tendenziell signifikante Zusammenhänge zu den MRT-Strategien ($\chi^2=7,10$; $df=3$; $p=.069$). Die im MRT erfolgreichste Klasse der eher holistisch arbeitenden Versuchspersonen war überproportional mit jüngeren Erwachsenen besetzt und die leistungsschwächste Klasse mit älteren. Die korrigiert-standardisierten Residuen wiesen die entsprechenden Häufungen auf diesen Faktorstufen als signifikant aus. Als generelles Maß für den Zusammenhang zwischen Alter und MRT-Lösungsstrategie in dieser Stichprobe wurde der korrigierte

Tabelle 11.15: Zusammenhänge Klasseneinteilung MRT und Altersgruppe.

	Klasse 1	Klasse 2	Klasse 3	Klasse 4
jüngere	9	12	5	2
ältere	8	5	13	4
korrigiert-standardisierte Residuen	0,46	2,19	-2,10	-0,77

Kontingenzkoeffizient $C_{\text{kor}}=0.429$ ermittelt.

11.4.3 Suchstrategien der Water-Maze-Aufgabe

Suchstrategien in den Spurdaten der virtuellen Water-Maze-Aufgabe wurden für alle 58 Versuchspersonen getrennt nach Explorationsdurchgang, Lerndurchgängen, Testdurchgang 1 und 2 mit der SNMF-Methode und dem Consensus-Clustering analysiert. Dazu wurden die hochauflösenden Trajektorien aus den Durchgängen in eine binäre 31×31 -Matrix \mathcal{B} kodiert. Ein Matrixelement b_{ij} war eins, wenn die Versuchsperson sich an der mit den Indizes i und j korrespondierenden Position im Maze aufgehalten hatte. Dieser Vorverarbeitungsschritt stellte eine starke Dimensionsreduktion der Eingangsdaten dar, war aber unbedingt notwendig, um eine Erkennung gemeinsamer Muster in einer kleineren Stichprobe zu ermöglichen.

Nach Linearisierung von \mathcal{B} entstand ein Datenvektor v mit $M=961$ dichotomen Variablen pro Person. Eine $M \times N$ -Matrix V aus $N=58$ solcher Vektoren bildete die Eingabe für den Klassifizierungsalgorithmus. Die Trajektorien aus den vier Lerndurchgängen wurden zusammengefasst verarbeitet. Dazu wurden die binären Matrizen der Spurdaten auf eine Dimension von 15×15 reduziert, linearisiert und zu einem Datenvektor der Länge $M=900$ konkateniert.

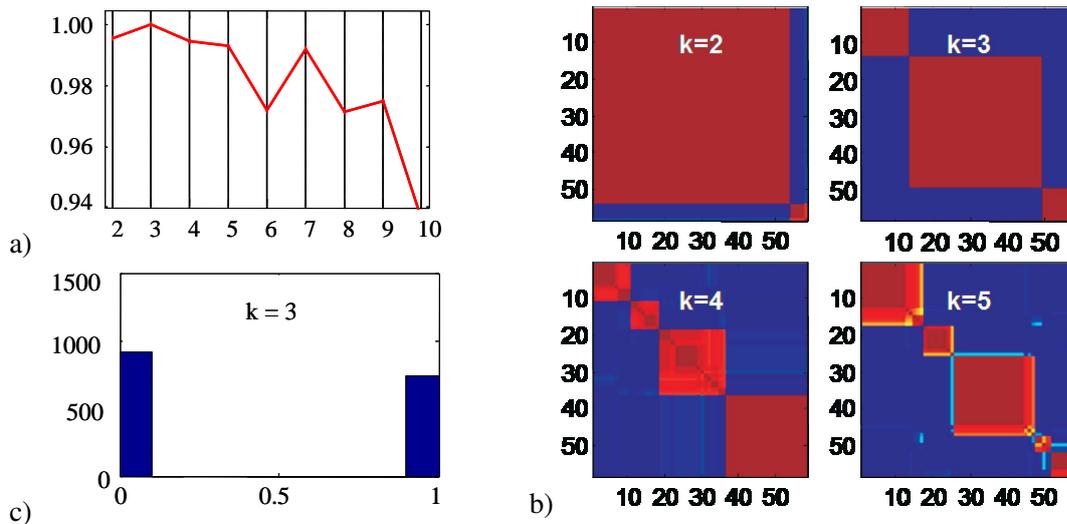
Parameter für den SNMF-Algorithmus waren $\lambda = 0.1$ und $N=800$. Im Consensus-Verfahren wurden Klassenlösungen $2 \leq k \leq 10$ mit Mittelungen über jeweils 20 SNMF-Läufe erprobt. Modellgütekriterien waren der kophenetische Korrelationskoeffizient φ_k und die Verteilung der akkumulierten Häufigkeiten in der Consensus-Matrix.

Explorationsdurchgang

Aus dem Explorationsdurchgang konnten anhand des kophenetischen Korrelationskoeffizienten Klassenlösungen mit zwei bis vier bzw. sieben Clustern als günstige angenommen werden (s. Abb. 11.12 a). Die Abbildung 11.12 b der Consensus-Matrizen unterstützt eine Wahl des Modellparameters mit $k=3$, da in diesem Fall die Klasseneinteilungen nach jedem SNMF-Lauf identisch waren, d. h. es lag eine eindeutige Klassifizierung vor. Das Histogramm der akkumulierten Häufigkeiten verdeutlicht dies, indem es für die Consensus-Matrix $k=3$ nur Einträge mit null oder eins aufweist.

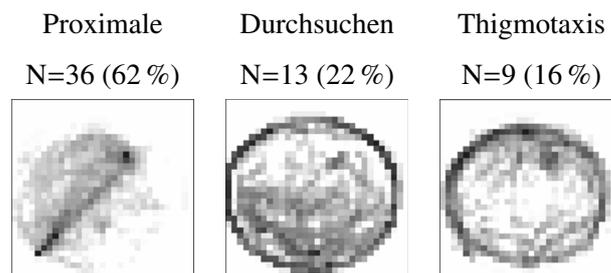
Zur Interpretation der 3-Klassenlösung werden die Basisvektoren w_i aus einem NMF-Lauf mit $k=3$ in eine 31×31 -Matrix rücktransformiert und als Graustufen-Plot visualisiert (s. Abb. 11.13). Die

Abbildung 11.12: Modellselektion für die Musteranalyse des Explorationsdurchgangs. (a) φ -k-Diagramm für $k=2$ bis $k=10$. (b) Consensus-Plots für $k=2$ bis $k=5$. (c) Akkumulierte Häufigkeiten für $k=3$.



Grauwerte der Abbildungen können als Aufenthaltswahrscheinlichkeit an den mit den Bildpunkten korrespondierenden Positionen im Maze interpretiert werden. So verdeutlichen die Muster drei typische Verhaltensweisen während der Exploration. Versuchspersonen mit der Strategie „Proximale“ zeigen direkte Trajektorien vom Start zum versteckten Zielpunkt. Diese Versuchspersonen folgten offensichtlich den proximalen Hinweisreizen, die den direkten Weg zum Zielpunkt vorgaben. Diese Strategie wendeten insgesamt 62 % der Versuchspersonen an. Die zweite Strategie wurde auf Grund der Abbildung als ein „Durchsuchen“ der gesamten Umgebung interpretiert. Wenige Versuchspersonen, insgesamt 16 %, versuchten durch systematisches Umrunden der Insel die Ziellokation zu entdecken. Dieses Verhalten, das sich durch eine reizbezogene Orientierung am Rand der Insel auszeichnet, ist typisch für „Thigmotaxis“.

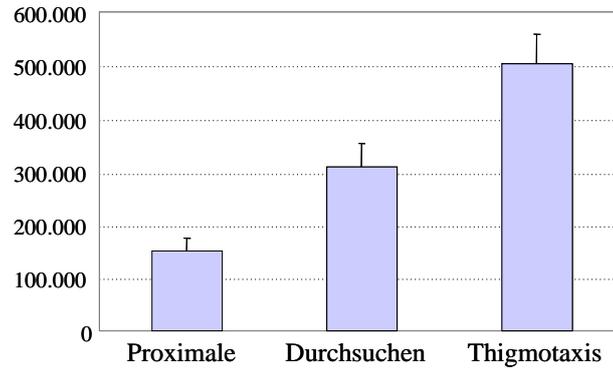
Abbildung 11.13: Identifizierte Bewegungsmuster im Explorationsdurchgang.



Die Anwendung einer dieser Explorationsstrategien führt zu deutlichen Unterschieden in den zurückgelegten Pfadlängen. Abbildung 11.14 zeigt die Mittelwertsunterschiede zwischen den identifizierten Subgruppen. Da Varianzhomogenität zwischen den drei Gruppen vorlag, konnte eine univariate ANOVA zur Bestimmung der Größe des Effekts durchgeführt werden. Insgesamt 42 % der Varianz in den Explorationsdaten konnte durch den Faktor „Suchstrategie“ aufgeklärt werden

($F=19,75; p<.001$).

Abbildung 11.14: Mittelwertsunterschiede und Standardfehler der Pfadlängen bezüglich der Explorationsklassen.



Zusammenhänge zwischen dem Faktor „Explorationsstrategie“ und der Altersgruppen- bzw. Geschlechtszugehörigkeit stellen die Tabellen 11.16 und 11.17 dar. Der korrigierte Kontingenzkoeffizient $C_{\text{kor}}=0.405$ deutet auf einen mittleren Zusammenhang zwischen Geschlecht und Explorationsstrategie. Der Zusammenhang zwischen Strategie und Geschlecht war allerdings nur tendenziell ($\chi^2=5,21; df=2; p=.074$). Insbesondere waren aber die 13 der Klasse „Durchsuchen“ zugeordneten Versuchspersonen männlich. Die korrigiert-standardisierten Residuen weisen diesen Unterschied als signifikant aus. In den Klassen „Proximale“ und „Thigmotaxis“ befinden sich zwar nicht signifikant, aber nominal mehr weibliche Versuchspersonen.

Tabelle 11.16: Zusammenhang zwischen Geschlecht und Explorationsstrategie.

	Proximale	Durchsuchen	Thigmotaxis
weiblich	20	3	6
männlich	16	10	3
korrigiert-standardisierte Residuen	1,1	-2,2	1,1

Altersunterschiede in der Explorationsstrategie konnten in dieser Stichprobe ebenfalls nur nominal beobachtet werden ($\chi^2=4,57; df=2; p=.102$). Ältere Probanden waren hier häufiger in der Klasse „Proximale“ und im Verhältnis zur Größe der Subgruppen seltener in der Klasse „Thigmotaxis“.

Tabelle 11.17: Zusammenhang zwischen Altersgruppe und Explorationsstrategie.

	Proximale	Durchsuchen	Thigmotaxis
jüngere	14	7	7
ältere	22	6	2
korrigiert-standardisierte Residuen	-1,8	0,5	1,9

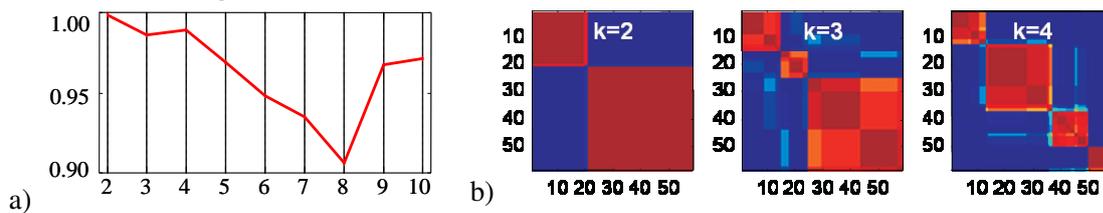
Zu bemerken ist noch, dass aufgrund der generellen Überbesetzung der Klasse „Proximal“ nur wenige Versuchspersonen für die übrigen Kategorien verblieben. Mehr als 20 % der Zellen wiesen

Werte unterhalb der kleinsten erwarteten Häufigkeit auf. Die Zuverlässigkeit des Signifikanztest kann deshalb nicht gewährleistet werden.

Lerndurchgänge

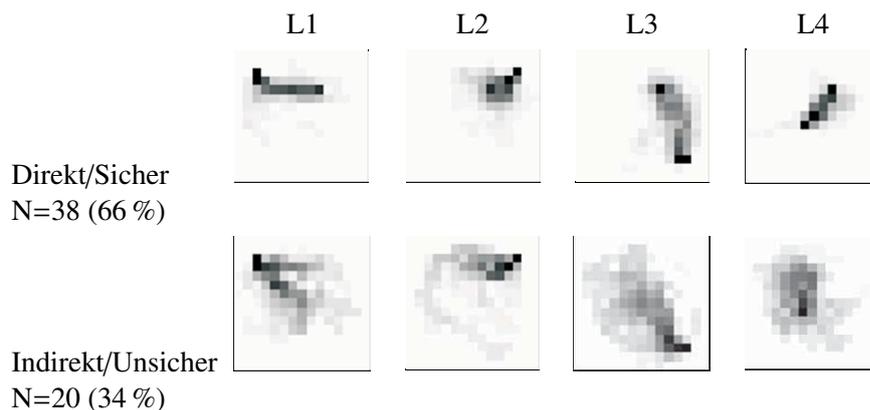
Aus der Analyse der Bewegungsmuster der Lernphase wurde eine Lösung mit zwei Klassen ausgewählt. Mit $\varphi(2) \approx 1$ und den Häufigkeitsverteilung in der Consensus-Matrix ist dieses Modell als optimalste Klassifizierung identifiziert worden (s. Abb. 11.15).

Abbildung 11.15: Modellselektion für die Musteranalyse der Spurdaten der Lerndurchgänge. (a) φ -k-Diagramm für $k=2$ bis $k=10$ und (b) Consensus-Plots für $k=2$ bis $k=4$.



Zur Interpretation der Klassen als Suchstrategien dient Abbildung 11.16. Es ist zu erkennen, dass der größte Teil der Versuchspersonen (66 %) in allen vier Lerndurchgängen Bewegungsmuster zeigte, die in direkter Weise auf die versteckte Ziellokation hinführten. Offensichtlich konnten Personen dieser Klasse unabhängig von der Startposition sicher das Ziel in jedem Durchgang auffinden. Die verbleibenden 34 % der Versuchspersonen wurden einer Klasse zugeordnet, die sich in allen Lerndurchgängen durch Bewegungsmuster hervorhoben, die eher unsicheres Suchverhalten ausdrücken. Sie bewegten sich indirekt oder häufig gar nicht auf das Ziel zu. Dies wird in der Abbildung auch durch eine größere Varianz der Bewegungsmuster innerhalb dieser Gruppe deutlich.

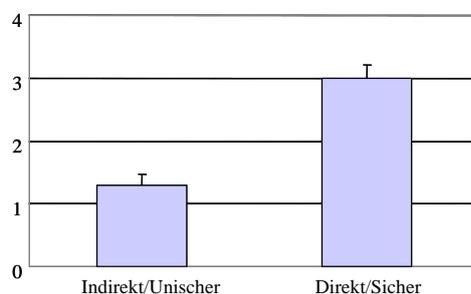
Abbildung 11.16: 2-Klassenlösung der Suchstrategien für die Lerndurchgänge.



Das direkte und sichere Auffinden des Ziels in jedem Lerndurchgang wirkte sich unmittelbar auf Lernleistung im Water-Maze aus. Abbildung 11.17 zeigt dazu, die in den Klassen gemittelte Anzahl erfolgreicher Lerndurchgänge. Während Personen mit der direkten Suchstrategie im Mittel 3,0 ($SD=1,29$) der vier Lerndurchgänge erfolgreich abschlossen, waren es in der Klasse mit indirekten

und unsicheren Strategien im Mittel nur 1,3 Durchgänge (SD=0,80). Varianz- als auch Mittelwertsunterschiede waren signifikant (Levene-Test: $F=7,63$; $p=.008$; $t(54,37)=6,16$; $p=.000$).

Abbildung 11.17: Gruppenmittelwerte und Standardfehler in der Anzahl erfolgreicher Lerndurchgänge bezüglich der Suchstrategien in den Lerndurchgängen.



Geschlechts- und Altersgruppenspezifisch bezüglich der Suchstrategien in der Lernphase fassen die Tabellen 11.18 und 11.19 zusammen. Die Signifikanztests ergaben für beide Faktoren keine signifikanten Zusammenhänge. Für die Kontingenzunterschiede zwischen weiblichen und männlichen Versuchspersonen betrug die Prüfgröße $\chi^2=1,22$ ($df=1$; $p=.269$). Die beobachteten Häufigkeiten und Residuen in Tabelle 11.18 deuten höchstens auf eine nominal stärkere Anwendung direkter Suchstrategien bei männlichen Versuchspersonen.

Tabelle 11.18: Zusammenhang zwischen Geschlecht und Strategie in den Lerndurchgängen.

	Indirekt / Unsicher	Direkt / Sicher
männlich	8	21
weiblich	12	17
korrigiert-standardisierte Residuen	-1,11	1,11

Zusammenhänge zwischen den Altersgruppen und der Ausprägung indirekten bzw. direkten Suchverhaltens in der Lernphase waren nominal noch etwas größer, aber ebenfalls nicht signifikant ($\chi^2=2,15$; $df=1$; $p=.142$). Der korrigierte Kontingenzkoeffizient zeigt mit $C_{\text{korrt}}=0.267$ insgesamt für die Faktoren „Altersgruppe“ und „Suchstrategie“ einen geringen Zusammenhang an.

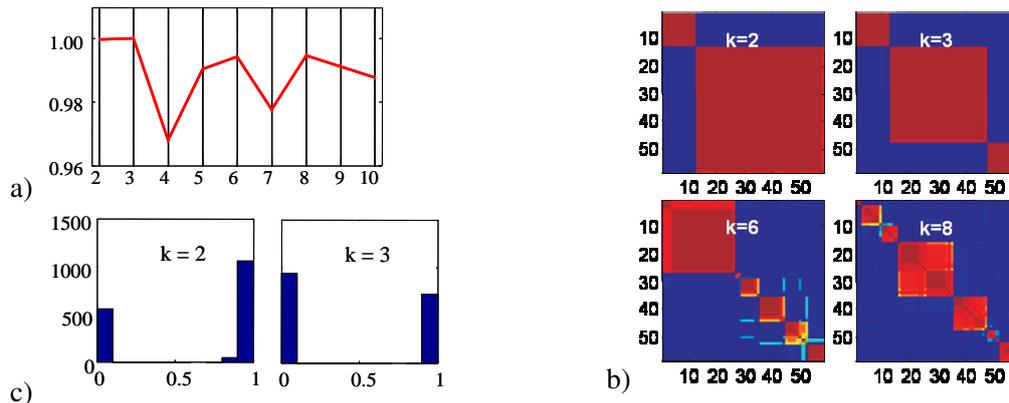
Tabelle 11.19: Zusammenhang zwischen Altersgruppe und Strategie in den Lerndurchgängen.

	Indirekt / Unsicher	Direkt / Sicher
jung	7	21
alt	13	17
korrigiert-standardisierte Residuen	-1,47	1,47

Testdurchgang 1

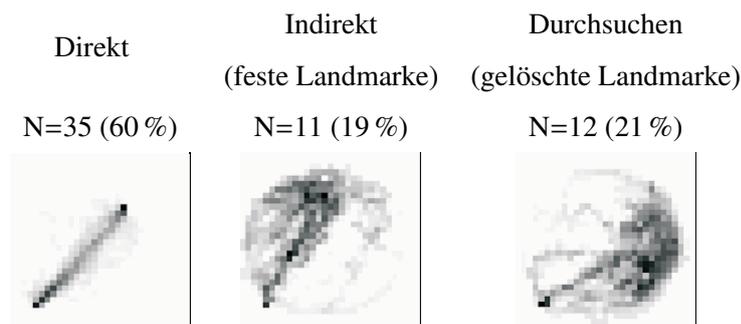
Für die Testbedingung 1 (Löschen der beweglichen Landmarken Segelboot und Ballon) ergab das automatische Klassifizierungsverfahren eine optimale Lösung mit $\varphi(k) = 1$ für $k=3$. Der Parameter $k=2$ lieferte im φ - k -Diagramm ebenfalls eine sehr gute Modellanpassung. Anhand der Consensus-Plots (s. Abb. 11.18) und der Tatsache, dass durch ein größeres k Strategien besser differenziert werden können, erschien das 3-Klassen-Modell hier plausibler.

Abbildung 11.18: Modellselektion für die Musteranalyse des Testdurchgang 1. (a) φ - k -Diagramm für $k=2$ bis $k=10$. (b) Consensus-Plots für $k=2, k=3, k=6$ und $k=8$. (c) Akkumulierte Häufigkeiten $k=2$ und $k=3$.



In Abbildung 11.19 sind die mit der Klasseneinteilung verbundenen Suchstrategien dargestellt. Es wurde in dieser Testbedingung eine sehr große Gruppe von insgesamt 60 % aller Versuchspersonen identifiziert, die im wesentlichen ein Bewegungsmuster zeigte, das vom Start direkt zum Ziel führte. Personen dieser Gruppe wurden deshalb direkten Suchstrategien zugeordnet.

Abbildung 11.19: 3-Klassenlösung der Suchstrategien im Testdurchgang 1.



Die übrigen 40 % der Stichprobe ließen sich anhand ihrer Bewegungsdaten im Testdurchgang 1 bei- nahe gleichgroßen Untergruppen zuordnen. Personen, die der Klasse „Indirekt“ zugeordnet wurden, gingen offensichtlich auf die im Norden der Insel verbliebene Landmarke zu und durchsuchten von dort beginnend systematisch die nähere Umgebung. Personen der Gruppe „Durchsuchen“ durchforsteten ebenfalls systematisch einen bestimmten Teil der Insel nach der Ziellokation, überraschen- derweise häuften sich deren Bewegungsmuster in der Nähe der gelöschten Landmarke.

Tabelle 11.20: Zusammenhang Suchstrategie und Erfolg im Testdurchgang 1.

	Direkt	Indirekt	Durchsuchen
erfolgreich	35	4	6
nicht erfolgreich	0	7	6
korrigiert-standardisierte Residuen	5,05	-3,64	-2,57

Als Leistungsmaß aus dem Testdurchgang 1 wurde die kategoriale Variable „Testerfolg“ mit der Klasseneinteilung der Suchstrategien korreliert. Tabelle 11.20 fasst diese Ergebnisse zusammen. Eindeutig ist festzustellen, dass alle Personen mit direkten Suchstrategien erfolgreich den Testdurchgang absolvierten. Andere Strategien führten jeweils nur zur Hälfte zum Erfolg im Test. Dieser Zusammenhang ist statistisch signifikant ($\chi^2=26,11$; $df=2$; $p<.001$) und drückt durch den Kontingenzkoeffizienten $C_{\text{kor}}=0.788$ die hohe Bedeutsamkeit der Suchstrategien für den Testerfolg aus.

Tabelle 11.21: Zusammenhang zwischen Geschlecht und Suchstrategie im Testdurchgang 1.

	Direkt	Indirekt	Durchsuchen
männlich	17	5	7
weiblich	18	6	5
korrigiert-standardisierte Residuen	-0,27	-0,33	0,65

Geschlechtsdifferenzen in der Anwendung von Suchstrategie im Testdurchgang 1 sind in Tabelle 11.21 dargestellt. Es bestand eine nahezu identische Verteilung der Geschlechter innerhalb der Strategieklassen ($\chi^2=0,45$; $df=2$; $p=.797$). Männer und Frauen zeigten also keine spezifischen Suchmuster in diesem Test.

Tabelle 11.22: Zusammenhang zwischen Altersgruppe und Suchstrategie im Testdurchgang 1.

	Direkt	Indirekt	Durchsuchen
jüngere	24	2	2
ältere	11	9	10
korrigiert-standardisierte Residuen	3,82	-2,22	-2,46

Dagegen stellten sich Altersdifferenzen in der Strategieanwendung nominal als sehr deutlich und statistisch signifikant heraus ($\chi^2=14,57$; $df=2$; $p=.001$). Die in Tabelle 11.22 zusammengefassten beobachteten Häufigkeiten und Residuen weisen die Unterschiede auf allen Faktorstufen der Klasseneinteilung als statistisch bedeutsam aus. Besonders auffällig ist die überproportionale Häufig jüngerer Versuchspersonen in der Klasse der direkten Suchstrategien. Der allgemeine Zusammenhang zwischen Alter und Suchstrategie im Test mit partiell gelöschten Landmarken betrug $C_{\text{kor}}=0.633$ und ist als starke Korrelation einzustufen.

Testdurchgang 2

Zur Auswahl des Modellparameters zur Klassifizierung der Bewegungsmuster aus dem Testdurchgang 2 (Löschen aller Landmarken) wurden durch den Consensus-Clusteralgorithmus optimale Lösungen für $k=2$ bzw. $k=6$ bestimmt (s. Abb. 11.20). Das Modell mit nur zwei Klassen würde anhand des Koeffizienten $\varphi(k)$ als passendste Lösung gelten. Jedoch lagen bei der Faktorisierung der Datenmatrix in nur zwei Basisvektoren die relativen Residuen über 50 % (s. Abb. 11.20 c). In der Abbildung sind die Euklidischen Normen der Residuen relativiert an der Gesamtdatenmatrix in Abhängigkeit vom Modellparameter k dargestellt. Erst für den ebenfalls optimalen Parameter $k=6$ fielen diese unter 40 %, d. h. das Modell mit wesentlich mehr Basisvektoren passte besser auf die vorliegenden Bewegungsdaten aus der Testbedingung 2.

Abbildung 11.20: Modellselektion für Testdurchgang 2. (a) φ - k -Diagramm für $k=2-10$ (b) Consensus-Plots $k=2$ und $k=6$ (c) Diagramm der relativen mittleren Residuen der Matrix-Faktorisierungen (d) Häufigkeitsverteilungen in den Consensus-Matrizen für $k=2$ und $k=6$.

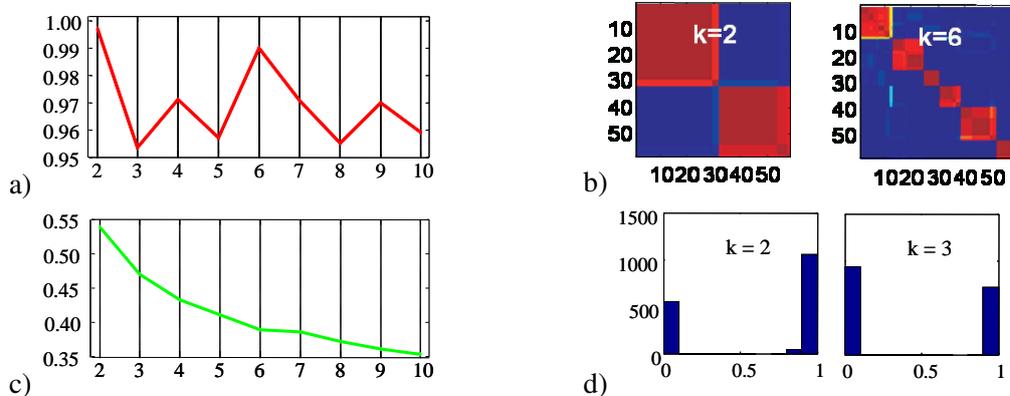
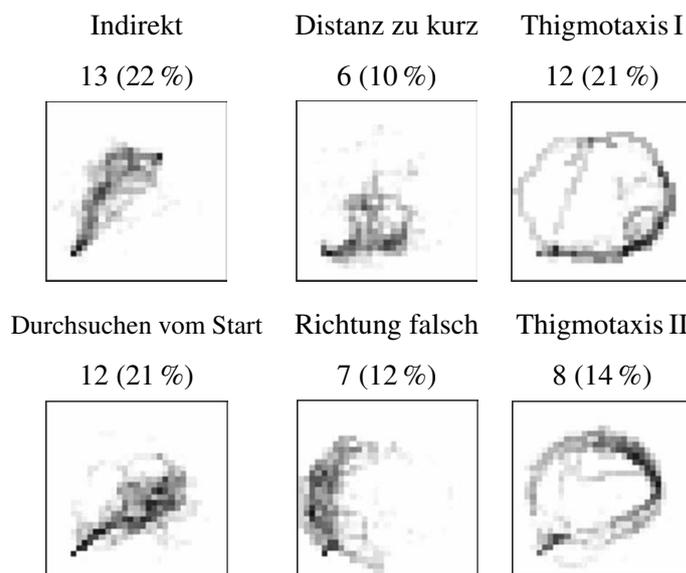


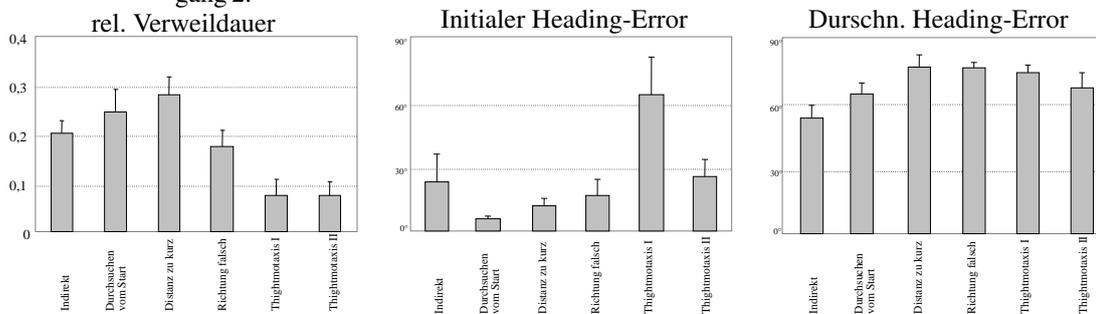
Abbildung 11.21: 6-Klassenlösung der Suchstrategien während der Testbedingung 2.



Zur Interpretation des Modells wurden in Abbildung 11.21 die ermittelten Basisvektoren als Be-

wegungsmuster geplottet. Die Versuchspersonen verteilen sich mit zehn bis 22 % von der Gesamtstichprobe relativ gleichmäßig auf die sechs identifizierten Klassen. Zur Beschreibung der Suchstrategien wurden neben den Bewegungsmustern hier zusätzlich die in Abbildung 11.22 dargestellten quantitativen Maße herangezogen.

Abbildung 11.22: Gruppenmittelwerte und Standardfehler der relativen Verweildauern im Zielquadranten und initialen und durchschnittlichen Heading-Error bezüglich der sechs Klassen im Testdurchgang 2.



Im Vergleich der Klassenmittelwerte der relativen Verweildauer und der Heading-Error zusammen mit den Bewegungsmustern ergaben sich die in Abbildung 11.21 angegebenen Bezeichnungen der Suchstrategien. In der Klasse „Indirekt“ sind Versuchspersonen mit moderater relativer Verweildauer im Zielquadranten (20 % der Navigationszeit) und moderatem initialen Heading-Error (23°). Die durchschnittlichen Heading-Error, als Maß für Kursabweichungen, sind in dieser Gruppe am geringsten (53°). Zusammen mit der Abbildung des Bewegungsmusters wird deutlich, dass es sich um Personen handelt, die Richtung und Distanz zum Ziel gut wiedererinnern und sich dann auf einer über die Personen relativ stabilen Trajektorie indirekt dem Ziel nähern. Personen mit dieser Suchstrategie waren in diesem Durchgang auch am erfolgreichsten. Acht von 13 Versuchspersonen fanden so das Ziel in dieser schwierigen Testbedingung (s. Tab. 11.23).

Tabelle 11.23: Suchstrategie und Erfolg im Testdurchgang 2.

	Indirekt	Durchsuchen vom Start	PI mit Fehlern	Thigmotaxis
erfolgreich	8	6	2	6
nicht erfolgreich	5	6	11	14
korrigiert-standardisierte Residuen	1,99	0,97	-1,90	-0,90

Die zweite Suchstrategie wurde als „Durchsuchen vom Start“ beschrieben. Personen mit dieser Strategie erinnerten sich sehr genau an die Richtung zum Ziel, was im sehr geringen mittleren initialen Heading-Error dieser Gruppe zum Ausdruck kam (6°; SD=4,2°). Diese Personen begannen aber nach dem relativ genauen Start, ab der Hälfte des Weges in Richtung des Ziels systematisch zu suchen. Dadurch stieg der durchschnittliche Heading-Error und die Verweildauer im Zielgebiet. Offensichtlich schien die Entfernung zum Ziel nicht genau bekannt zu sein. Immerhin die Hälfte, sechs von zwölf mit dieser Strategie identifizierten Versuchspersonen, erreichten so das Ziel (s. Tab. 11.23).

Eine ähnliche Suchstrategie versuchten Personen der Klasse „Distanz zu kurz“. Sie begannen ebenfalls, allerdings mit einem etwas größeren initialen Heading-Error von 12° (SD=12,8°), in Richtung des Ziels zu laufen. Mussten aber ab ungefähr der Hälfte des Weges systematisch das Zielgebiet durchsuchen, möglicherweise weil die wahre Distanz unterschätzt wurde. Dadurch hatten sie im Gruppenmittel eine hohe Verweildauer im Zielquadranten und mit die größten durchschnittlichen Heading-Error. Aufgrund des größeren initialen Richtungsfehlers fand keine der sechs Personen mit dieser Strategie das Ziel.

Eine vierte Gruppe von Versuchspersonen zeigte in ihren Verhaltensdaten nicht unbedingt eine Fehleinschätzung der Distanz zum Ziel sondern der Richtung. Sie begannen zwar mit einem moderaten initialen Heading-Error von 17° (SD=17,2°), wichen dann aber in ihrem Bewegungsmuster in einen zielfernen Bereich der Insel aus (Verweildauer im Zielquadrant: 18 % der Navigationszeit). Zwei von fünf Personen erreichten so mehr zufällig das Ziel. Die letzten beiden Gruppen wurden in eine Kategorie zusammengefasst, in der offensichtlich Personen sind, die Pfadintegration zur Zielsuche im Testdurchgang 2 einsetzten, allerdings mit Fehlern in Richtungs- bzw. Distanzeinschätzungen.

Die letzten beiden Gruppen zeigten typisches Verhalten der Thigmotaxis und unterscheiden sich lediglich in ihren Gruppenmittelwerten der initialen Heading-Error von 65° bzw. 25°. Wegen der ähnlichen Bewegungsmuster wurden sie ebenfalls in eine Kategorie zusammengefasst. Die Verweildauer im Zielquadrat war auf Grund der systematischen Suche am Rand der Insel in dieser Gruppe mit sieben Prozent am geringsten. Aber immerhin sechs von 14 Personen erreichten mit der Thigmotaxis-Strategie das Ziel im Testdurchgang 2.

Tabelle 11.23 fasst Zusammenhänge zwischen Strategie und Testerfolg zusammen. Diese waren mit $C_{\text{kor}}=0.470$ in einem mittleren Bereich und nur tendenziell signifikant ($\chi^2=7,161$; $df=3$; $p=.067$). Die indirekte Suchstrategie führte signifikant häufiger zum Testerfolg als andere Strategien. Zusammenhänge der Suchstrategien zu den Faktoren „Geschlecht“ und Altersgruppe sind in den Tabellen 11.24 und 11.25 dargestellt.

Tabelle 11.24: Zusammenhang Geschlecht und Strategie im Testdurchgang 2.

	Indirekt	Durchsuchen vom Start	PI mit Fehlern	Thigmotaxis
männlich	7	6	8	8
weiblich	6	6	5	12
korrigiert-standardisierte Residuen	0,31	0,00	0,94	-1,11

Bezüglich der Geschlechtszugehörigkeit konnten keine signifikanten Häufungen in bestimmten Strategieklassen gefunden werden ($\chi^2=1,569$; $df=3$; $p=.666$). Nominal waren mehr weibliche Versuchspersonen in der Klasse mit Thigmotaxis-Strategien.

In der Verteilung jüngerer und älterer Erwachsener auf die einzelnen Strategien wurden sehr deutliche Unterschiede entdeckt ($\chi^2=12,15$; $df=3$; $p=.007$). Insgesamt konnte ein Zusammenhang zwischen Altersgruppen und Suchstrategie im Test 2 von $C_{\text{kor}}=0.588$ festgestellt werden. Hauptsäch-

Tabelle 11.25: Zusammenhang Altersgruppe und Strategie im Test 2.

	Indirekt	Durchsuchen vom Start	PI mit Fehlern	Thigmotaxis
jüngere	6	9	9	4
ältere	7	3	4	16
korrigiert-standardisierte Residuen	-0,17	2,08	1,72	-3,13

lich verantwortlich für diesen Effekt ist die Verteilung in der Klasse der Thigmotaxis-Strategien. Mehr als 53% der älteren Erwachsenen versuchten sich mit diesem Vorgehen im Testdurchgang 2. Jüngerer Erwachsenen wurden dagegen signifikant häufiger in der Klasse „Durchsuchen vom Start“ und tendenziell häufiger in der Klasse „Pfadintegration mit Fehlern“ aufgefunden.

11.4.4 Zusammenhänge zwischen den Klasseneinteilungen

Zur Beantwortung der Frage, ob die voran in verschiedenen räumlichen Tests identifizierten Strategien situationsinvariante Verhaltensmuster darstellen oder spezifisch auf eine Situation angepasstes Verhalten widerspiegeln, wurden korrelativ Zusammenhänge überprüft. Für einen ersten Überblick wurden die korrigierten Kontingenzkoeffizienten, die sich aus der paarweisen Kreuztabellierung aller sechs durchgeführten Klasseneinteilungen ergaben, in einer Korrelationsmatrix zusammengefasst (s. Tab. 11.26). Für signifikante Korrelationen wurden die Kontingenzen genauer untersucht. Auf Grund geringer Teststärken der kategorialen Zusammenhangsanalysen wurden auch tendenziell signifikante Korrelationen mit in die Betrachtung einbezogen.

Tabelle 11.26: Korrelationen zwischen den identifizierten Suchstrategien im VWM, den Bearbeitungsstrategien im MRT und im VPT. Darstellung aller kategorialer Zusammenhänge (korrigierte Kontingenzkoeffizienten).

	1.	2.	3.	4.	5.	6.
1. Exploration	-	0.158	0.236	0.496 ⁺	0.318	0.492 ⁺
2. Lernen		-	0.558 ^{**}	.122	.064	0.337
3. Test 1			-	0.366	0.510 [*]	0.383
4. Test 2				-	0.220	0.634 ^{**}
5. VPT					-	0.380
6. MRT						-

⁺<0.10, ^{*}<0.05, ^{**}<0.01

Durch diese Kriterien wurden innerhalb der Water-Maze-Aufgabe Zusammenhänge zwischen den Suchstrategien im Explorations- und dem Testdurchgang 2 und Zusammenhänge zwischen dem

in den Lerndurchgängen beobachtetem Verhalten und der Strategie im Testdurchgang 1 identifiziert. Zur Klassifizierung der Bearbeitungsstrategien aus dem MRT bestanden Zusammenhänge zum Explorationsverhalten und den Suchstrategien im Testdurchgang 2 der Water-Maze-Aufgabe. Die Klasseneinteilung aus der Zeigeaufgabe korrelierte mit der Suchstrategie im Testdurchgang 1.

Tabelle 11.27: Zusammenhang Explorationsstrategien und Suchstrategie im Testdurchgang 2.

	Indirekt	Durchsuchen vom Start	PI mit Fehlern	Thigmotaxis
Proximale	9	7	4	16
Durchsuchen+Thigmotaxis	4	5	9	4
korrigiert-standardisierte Residuen	0,60	-0,30	-2,64	2,04

In der folgenden genaueren Betrachtung der Häufigkeitsverteilungen wurden zur Erhöhung der Teststärke und Durchführbarkeit gültiger Signifikanztests mit dem χ^2 -Verfahren einzelne Faktorstufen sinnvoll zusammengefasst. Tabelle 11.27 zeigt die Häufigkeiten der Explorationsstrategien verteilt auf die Kategorien der Suchstrategien im Testdurchgang 2. Die Explorationsstrategien „Thigmotaxis“ und „Durchsuchen“ wurden hier zu einer Kategorie zusammengefasst, da sie inhaltlich ähnlich sind. Der Signifikanztest für diese Zusammenfassung ergab mit $\chi^2=8,50$ ($df=3$; $p=.037$) und $C_{\text{kor}}=0.505$ einen mittleren Zusammenhang zwischen diesen Suchstrategien. Die Residuen in der Tabelle 11.27 verdeutlichen einen positiven Zusammenhang zwischen dem Verhalten im Explorationsdurchgang den proximalen Hinweisen zu folgen und der Erkundung der Insel durch Thigmotaxis im Testdurchgang 2. Statistisch bedeutsam ist ebenfalls eine Überbesetzung der Klasse der Personen, die sich in der Exploration nicht an den proximalen Hinweisen orientieren und sich dann im Testdurchgang 2 mit fehlerhafter Pfadintegration bewegten.

Tabelle 11.28: Zusammenhänge Suchstrategie in den Lerndurchgängen und der Suchstrategie im Testdurchgang 1.

	Direkt	Indirekt	Durchsuchen
Direkt/Sicher	28	3	7
Indirekt/Unsicher	7	8	5
korrigiert-standardisierte Residuen	2,86	-2,96	-0,59

Die Verhaltensmuster in den Lerndurchgängen im VWM standen in direktem Zusammenhang mit den im Testdurchgang 1 beobachteten Suchstrategien (s. Tab. 11.28). Die Häufigkeiten verdeutlichen, dass Personen, die bereits während der Lernphase sicher und direkt zum Ziel fanden, auch mit den Veränderungen im Testdurchgang 1 ähnliche Bewegungsmuster zeigten. Sieben Personen, die kein sicheres Lernverhalten zeigten, konnten im Test dennoch direkt das Ziel ansteuern. Personen mit unsicherem Lernverhalten wendeten im Testdurchgang signifikant häufiger indirekte Suchstrategien als Personen mit sicherem Lernverhalten. In der Kategorie der Teststrategie „Durchsuchen“ befanden sich statistisch gleich viele Personen mit sicherem wie mit unsicherem Lernverhalten.

Tabelle 11.29: Kontingenztafel der Suchstrategien im Testdurchgang 1 und der Klasseneinteilung aus der Kompass-Zeigeaufgabe.

	Orientiert	Desorientiert	Raten
Direkt	15	6	14
Indirekt + Durchsuchen	5	13	5
korrigiert-standardisierte Residuen	1,66	-3,13	1,45

Signifikante Zusammenhänge zwischen den Strategien im VWM und der Klasseneinteilung, die bezüglich des Desorientierungseffekts im VPT getroffen wurden, bestanden nur zu den Suchmustern in der Testbedingung 1. Zur Durchführung des χ^2 -Tests wurden die Kategorien „Indirekt“ und „Durchsuchen“ in der Kontingenztafel zusammengefasst (s. Tab. 11.29). Die Korrelation bezüglich dieser Kategorien betrug $C_{\text{kor}}=0.537$ und ist mit $\chi^2=9,78$ ($df=2$; $p=.008$) signifikant.

Die 4-Klasseneinteilung der MRT-Bearbeitungsstrategien wurde zur Weiterverarbeitung für die Zusammenhangsanalysen in eine dichotome Variable reduziert. Diese bildet die vier Kategorien direkt auf das Konstrukt einer räumlichen Präferenz für eher holistische bzw. eher analytische Bearbeitungsstrategien ab. Dazu wurden die Klassen 1 und 2 zu einer Kategorie holistischer Strategien und die Klassen 3 und 4 zu einer Kategorie eher analytischer Strategien zusammengefasst.

Tabelle 11.30: Kontingenztafel Präferenz im MRT und Suchstrategie in der Exploration im VWM.

	Proximale	Durchsuchen+Thigmotaxis
analytisch	20	4
holistisch	16	18
korrigiert-standardisierte Residuen	2,80	-2,80

In Tabelle 11.30 ist die Verteilung analytisch bzw. holistisch arbeitender Versuchspersonen auf die Suchstrategien im Explorationsdurchgang dargestellt. Ein Zusammenhang für analytische Bearbeitungsstrategien im MRT und der Bevorzugung proximaler Hinweise während der Exploration des VWM und *vice versa* ist statistisch signifikant ($\chi^2=7,86$; $df=1$; $p=.005$). Der korrigiert Korrelationskoeffizient beträgt $C_{\text{kor}}=0.490$.

Tabelle 11.31: Kontingenztafel Präferenz im MRT und Suchstrategie im Testdurchgang 2.

	Direkt	Durchsuchen vom Start	PI mit Fehlern	Thigmotaxis
analytisch	5	1	2	16
holistisch	8	11	11	4
korrigiert-standardisierte Residuen	-0,24	-2,61	-2,16	4,33

Der stärkste Zusammenhang dieser korrelativen Analysen fand sich zwischen der Präferenz für ana-

lytische bzw. holistische Strategien im MRT und der Wahl der Suchstrategie in der Testbedingung 2 des VWM ($\chi^2=21,38$; $df=3$; $p<.001$). Die Korrelation betrug hier nach Korrektur $C_{\text{kor}}=0.734$. Tabelle 11.31 veranschaulicht die dazugehörigen Häufigkeiten in den einzelnen Faktorstufen. Insbesondere sind analytisch arbeitende Versuchspersonen im VWM häufiger mit Thigmotaxis-Strategien identifiziert wurden als holistisch arbeitende. Letztere versuchten signifikant häufiger die Testbedingung 2 durch die Suchstrategien „Durchsuchen vom Start“ und „PI mit Fehlern“ zu lösen.

Abschließend wurde zur Bearbeitung der aus der Fragestellung aufgeworfenen Hypothese korrelativ der Zusammenhang zwischen den räumlichen Präferenzen analytisch vs. holistisch und egozentrisch vs. allozentrisch „überprüft“. Zur Entwicklung einer ähnlichen dichotomen Variable zur Abbildung der ego- vs. allozentrischen Präferenz müssen mehrerer Kriterien berücksichtigt werden. Aus dem VWM wurden Personen aus der Testbedingung 1, die kein direktes Suchverhalten zeigten, für die Kategorie der egozentrischen ausgewählt. Aus den direkten Bewegungsmustern konnte allerdings nicht sicher geschlossen werden, dass diese Personen ihre Strategien aus allozentrischem Überblickswissen hergeleitet hatten. Deshalb wurden Personen, die im VPT nicht der Gruppe „Orientiert“ angehörten, ebenfalls der Kategorie einer eher egozentrischen Präferenz zugewiesen. Alle übrigen galten dann als Personen mit einer sicheren Präferenz für allozentrische Strategien.

Tabelle 11.32: Zusammenhang zwischen den identifizierten räumlichen Präferenzen.

	egozentrisch	allozentrisch
analytisch	19	5
holistisch	24	10
korrigiert-standardisierte Residuen	-0,73	0,73

Nach diesen Kriterien wurden 28 % aller Versuchspersonen einer allozentrischen Präferenz zugeordnet. Diese verteilen sich gleichmäßig auf die Kategorien der Präferenz für analytische bzw. holistische Bearbeitungsstrategien ($\chi^2=0,54$; $df=1$; $p=.462$). Ein Zusammenhang zwischen diesen allgemeineren räumlichen Präferenzen konnte also nicht nachgewiesen werden.

11.4.5 Zusammenhänge zu Strategiefragebögen

Die Auswertung der eingesetzten Fragebögen zur Strategieanwendung im VPT und im VWM wurde nur deskriptiv durchgeführt, da die Items nicht zu einem Summenwert zusammengefasst und inferenzstatistisch untersucht werden konnten. Sie dienten so lediglich der Gegenüberstellung von Antworthäufigkeiten in den einzelnen Items und den Strategieklassen der automatischen Analyse-methode.

Tabelle 11.33 zeigt die Auswertung des Strategiefragebogens zur Kompass-Zeigeaufgabe. Es sind die Antworthäufigkeiten unter der Gruppeneinteilung dargestellt, die aus der Analyse zum Desorientierungseffekt gewonnen wurde. Das erste Item war als offene Frage formuliert und wurde meist unbeantwortet gelassen. Die Tabelle beginnt deshalb mit dem 2. Item. 33 % der Versuchspersonen

Tabelle 11.33: Deskriptive Auswertung des Strategiefragebogens und Häufigkeiten von Ja/Nein-Antworten in den Gruppen zur Kompass-Zeigeaufgabe.

		Gesamt	Orientiert	Desorientiert	Geraten
		58 (100 %)	20 (100 %)	19 (100 %)	19 (100 %)
2.	Ja	19 (33 %)	7 (35 %)	4 (21 %)	8 (42 %)
	Nein	39 (67 %)	13 (65 %)	15 (79 %)	11 (58 %)
3.	Ja	17 (29 %)	5 (25 %)	7 (37 %)	5 (26 %)
	Nein	41 (71 %)	15 (75 %)	12 (63 %)	14 (74 %)
4.	Ja	51 (88 %)	20 (100 %)	15 (79 %)	16 (84 %)
	Nein	7 (12 %)		4 (21 %)	3 (16 %)
5.	Ja	48 (83 %)	18 (90 %)	14 (74 %)	16 (84 %)
	Nein	10 (17 %)	2 (10 %)	5 (26 %)	3 (16 %)
6.	Ja	38 (66 %)	14 (70 %)	13 (68 %)	11 (58 %)
	Nein	20 (34 %)	6 (30 %)	6 (32 %)	8 (42 %)
7.	Ja	21 (36 %)	8 (40 %)	3 (16 %)	10 (53 %)
	Nein	37 (64 %)	12 (60 %)	16 (84 %)	9 (47 %)

gaben im Fragebogen an, dass sie sich die Versuchsanordnung aus der Vogelperspektive vorgestellt haben (2. Item). Die wenigsten Personen davon befanden sich in der Klasse der Personen mit „Desorientierung“ im VPT (fett gedruckt). Nur 29 % gaben an, den Weg in Gedanken rückwärts gelaufen zu sein (3. Item). 88 % hatten sich bei jeder Drehung überlegt, wie sich die Richtung zum Startpunkt verändert (4. Item). Alle Personen, die keine Desorientierung zeigten, beantworteten diese Frage mit „Ja“. 83 % der Versuchspersonen haben die aktuelle Position mit der Ausgangsposition verglichen (5. Item). Auch diese Strategie geben am häufigsten Personen an, die keine Desorientierung zeigten. Insgesamt 66 % hatten sich, laut Befragung, auf ihr Gefühl während des VPT verlassen (6. Item). Personen, die eher einer Rate-Strategie zugeordnet wurden, antworteten hier allerdings nominal häufiger mit „Nein“. Insgesamt 64 % der Versuchspersonen fielen die schnellen Durchgänge im VPT schwerer (7. Item). Insbesondere Personen der Gruppe „Desorientiert“ gaben dies im Fragebogen an.

Zu den Suchstrategien im Explorationsdurchgang des Water-Maze waren im Fragebogen drei Items zu beantworten (s. Tab. 11.34). Das erste Item war wieder als offene Frage formuliert und wurde ebenfalls meist unbeantwortet gelassen. Aus den Antworten zum zweiten Item ging hervor, dass 40 % der Versuchspersonen die proximalen Hinweise zur Orientierung nutzten. Die meisten dieser Personen wurden durch die automatische Musteranalyse auch einer entsprechenden Strategiegruppe zugewiesen (fett gedruckt). 72 % fanden die Zeit zum Umschauen hilfreich, nachdem sie das Ziel erreicht hatten (3. Item). Es beantworteten aber nominal weniger Versuchspersonen dieser Frage mit „Ja“, die eine Strategie einsetzten, mit der sie die gesamte Insel bereits durchsucht hatten.

Tabelle 11.34: Deskriptive Auswertung des Fragebogens zur Water-Maze-Aufgabe und den Strategien im Explorationsdurchgang.

		Gesamt	Durchsuchen	Proximal	Tightmotaxis
		58 (100 %)	13 (100 %)	36 (100 %)	9 (100 %)
2.	Ja	23 (40 %)	3 (23 %)	17 (47 %)	3 (33 %)
	Nein	35 (60 %)	10 (77 %)	19 (53 %)	6 (67 %)
3.	Ja	42 (72 %)	8 (62 %)	27 (75 %)	7 (78 %)
	Nein	16 (28 %)	5 (38 %)	9 (25 %)	2 (22 %)

Tabelle 11.35: Deskriptive Auswertung Strategiefragebögen und Suchstrategie in den Lerndurchgängen.

		Gesamt	Indirekt/Unsicher	Direkt/Sicher
		58 (100 %)	20 (100 %)	38 (100 %)
4.	Ja	54 (93 %)	17 (85 %)	37 (97 %)
	Nein	4 (6 %)	3 (15 %)	1 (3 %)
6.	Ja	58 (100 %)	20 (100 %)	38 (100 %)
	Nein			
7.	Leuchtturm und Segelboot	26 (45 %)	3 (15 %)	23 (61 %)
	Leuchtturm plus andere Landmarken	17 (29 %)	7 (35 %)	10 (26 %)
	nur Leuchtturm	12 (21 %)	7 (35 %)	5 (13 %)
	Leuchtturm nicht verwendet	3 (5 %)	3 (15 %)	
8.	Ja	8 (14 %)	3 (15 %)	5 (13 %)
	Nein	50 (86 %)	17 (85 %)	33 (87 %)

In den Lerndurchgängen gaben nur insgesamt vier (6 %) der Personen an, keine Strategie eingesetzt zu haben (4. Item in Tabelle 11.35). Alle Personen orientierten sich laut Befragung an den Landmarken im VWM (6. Item). In der Frage nach den verwendeten Landmarken konnten die insgesamt 15 Antwortmöglichkeiten zu vier Kategorien zusammengefasst werden, die alle in Bezug auf die Verwendung des Leuchturms zur Navigation standen (7. Item). 45 % der Personen verwendeten den Leuchtturm und das Segelboot. Diese Landmarken waren auch jene, die den Zielquadranten und seine Ausdehnung direkt markierten. Insbesondere Personen mit direkten/sicheren Lernmustern erkannten offensichtlich diesen Zusammenhang und gaben dies im Fragebogen an. Personen, die weitere bzw. andere Landmarken oder nur den Leuchtturm verwendeten, wurden in den Lerndurchgängen häufiger mit indirekten und unsicheren Strategien identifiziert. Die 21 Personen, die nur den Leuchtturm zur Orientierung verwendeten, hatten explizit festgestellt, dass es sich um eine

fixe Landmarke handelt (8. Item).

Tabelle 11.36: Deskriptive Auswertung Strategienfragebogen und Suchstrategien im Testdurchgang 1.

	Gesamt	Direkt	Indirekt	Durchsuchen
	58 (100 %)	35 (100 %)	11 (100 %)	12 (100 %)
nur Leuchtturm	41 (71 %)	27 (77 %)	8 (73 %)	6 (50 %)
10. Leuchtturm plus Windrad	13 (22 %)	8 (23 %)	2 (18 %)	3 (25 %)
Windrad oder Raten	4 (7 %)		1 (9 %)	3 (25 %)
11. Ja	30 (52 %)	14 (40 %)	8 (73 %)	8 (67 %)
Nein	28 (48 %)	21 (60 %)	3 (27 %)	4 (33 %)

Die Items, die sich auf die Suchstrategien in den Testdurchgängen beziehen, sind in den Tabellen 11.36 und 11.37 dargestellt. Die Antworten des als offene Frage formulierten 10. Items konnten zu drei Kategorien zusammengefasst werden. 71 % der Versuchspersonen gaben an, sich im Testdurchgang 1 nur am Leuchtturm orientiert zu haben. Hier antworteten nur 50 % der Personen, die die Insel in der Nähe der gelöschten Landmarke durchsuchten mit „Ja“. 52 % fanden den Testdurchgang 1 schwer in der Bearbeitung. Prozentual die wenigsten Ja-Antworten bei dieser Frage gaben Personen mit direkten Suchstrategien (11. Item).

Tabelle 11.37: Deskriptive Auswertung Strategienfragebogen und Suchstrategien im Testdurchgang 2.

	Gesamt	Indirekt	Durchsuchen vom Start	PI mit Fehlern	Thigmotaxis
	58 (100 %)	13 (100 %)	12 (100 %)	13 (100 %)	20 (100 %)
PI oder ähnliche	16 (28 %)	7 (54 %)	3 (25 %)	4 (31 %)	2 (10 %)
12. Kreisen	11 (19 %)	1 (8 %)	1 (8 %)	5 (38 %)	4 (20 %)
unbekannte Strategie	31 (53 %)	5 (38 %)	8 (67 %)	4 (31 %)	14 (70 %)
13. Ja	44 (76 %)	8 (62 %)	8 (67 %)	12 (92 %)	16 (75 %)
Nein	14 (23 %)	5 (38 %)	4 (33 %)	1 (8 %)	4 (25 %)

Die zur Testbedingung 2 explizit angegebenen Strategien konnten zu drei größeren Kategorien zusammengefasst werden (12. Item). 28 % gaben tatsächlich an, Pfadintegration oder Berechnungsmethoden eingesetzt zu haben. 19 % berichteten in wiederholten kleinen Kreisen (nicht Thigmotaxis) über die Insel gelaufen zu sein. Der größte Teil der Versuchspersonen (53 %) konnte aber explizit keine Angaben zur verwendeten Strategie machen. In der automatischen Analyse der Bewegungsmuster wurden letztere aber häufiger mit Thigmotaxis-Strategien identifiziert. Ebenso waren Personen, die Pfadintegration als Strategie explizit angaben, häufiger in der Klasse mit indirekten Suchstrategien. Die meisten Personen fanden die Testbedingung 2 schwer. Insbesondere Personen

die mit Suchstrategien identifiziert wurden, die auf fehlerhafter Pfadintegration beruhten, betraf diese.

11.5 Zusammenfassung

Durch das Screening allgemeiner kognitiver Fähigkeiten innerhalb der Stichproben konnte eine Vergleichbarkeit der Alters- und Geschlechtsgruppen in verbaler Intelligenz, aktueller Stimmung und visuo-konstruktiven Fähigkeiten festgestellt werden. Altersunterschiede in den Tests zur visuo-motorischen Leistungsfähigkeit und der kognitiven Umstellfähigkeit waren signifikant. Im Summenscore zur Vorerfahrung mit Raumvorstellungstests und der Umgang mit Computern unterschieden sich jüngere männliche Versuchspersonen deutlich von allen anderen Gruppen.

In der Raumvorstellung differenzierten die Ergebnisse in alterssensitive bzw. geschlechtssensitive Tests. In den Tests zur räumlichen Wahrnehmung (Feldabhängigkeit) und zum visuell-räumlichen Gedächtnis gingen auf die Altersgruppenzugehörigkeit 32 % bzw. 24 % der Varianz zurück. Für die Facette mentale Rotation bestanden geringere Altersdifferenzen in den Tests *Spiegelbilder* und *Figurenauswahl*. Tendenzielle Geschlechtsdifferenzen für diese Raumvorstellungsfacette ergaben sich nur im MRT.

Gruppenunterschiede in der Pfadintegrationsleistung, gemessen über die Schätzfehler der Kompass-Zeigeaufgabe, konnten für die Altersgruppen nachgewiesen werden. Der Effekt war mit nur neun Prozent Varianzaufklärung gering. Es zeigte sich in der Stichprobe ebenfalls ein Alterseffekt für die Reaktionszeiten (zwölf Prozent der Varianz).

Ergebnisse aus der virtuellen Water-Maze-Aufgabe dokumentieren große Alters- und moderate Geschlechtsdifferenzen bezüglich der Aufgabe. Dies zeigte sich bereits im Training zur virtuellen Navigation. Ältere Versuchspersonen hielten sich viel länger in der Trainingsumgebung auf als jüngere. Frauen nutzten ebenfalls mehr Trainingszeit als Männer. Junge Männer navigierten mit Abstand am schnellsten in der virtuellen Umgebung. Aber in allen Gruppen wurde während des Trainings ein mittlerer Geschwindigkeitszuwachs beobachtet.

Im Explorationsdurchgang blieb der Geschwindigkeitsvorteil junger männlicher Versuchspersonen bestehen. Frauen der älteren Altersgruppe legten die kürzesten Wege (gemessen über Pfadlängen) während der Exploration zurück. Die Anzahl der Umkehrpunkte war bei jüngeren Männer am geringsten. Insgesamt differenzierten diese Leistungsmaße aber nicht sehr gut zwischen den Alters- und Geschlechtsgruppen.

In der Lernphase der Aufgabe konnten nur jüngere Männer alle Durchgänge zu fast 100 % abschließen. Ältere Frauen verbesserten sich zwar, auf die Gruppe bezogen, kontinuierlich. Ihre Fehlerraten lagen in allen Lerndurchgängen aber über 50 %. Die Fehlerraten ältere Männer waren etwas geringer, schwankten jedoch ebenfalls um die 50 % Marke. In der Zusammenfassung als Summenscore der erfolgreichen Lerndurchgänge zeigten sich die starken Altersunterschiede in dieser Phase des Experiments am deutlichsten. 37% der Varianz in diesem Leistungsmaß gingen auf die Altersgrup-

penzugehörigkeit zurück. Auch in den Leistungsmaßen initialer Heading-Error und relativer Umweg bestanden diese Altersdifferenzen. Jeweils 20 % Varianz konnte durch den Faktor „Altersgruppe“ aufgeklärt werden. Geschlechtsdifferenzen in der Lernphase der Water-Maze-Aufgabe konnten für die Anzahl der erfolgreichen Lerndurchgänge gezeigt werden. Der Effekt betrug $d=0.62$.

Den Testdurchgang 1 absolvierten die meisten Versuchspersonen erfolgreich (Fehlerrate 22 %). Ältere Personen waren allerdings signifikant weniger erfolgreich als jüngere. Die initialen Heading-Error jüngerer Frauen waren die niedrigsten in diesem Testdurchgang im Vergleich zu allen anderen Gruppen. Eine Beurteilung getrennt nach erfolgreichen und nicht erfolgreichen Personen ergab keine signifikanten Unterschiede für die initialen Heading-Error. Die relative Verweildauer im Zielquadranten war für erfolgreiche deutlich größer als für nicht erfolgreiche Personen, d. h. Personen, die das Ziel nicht erreichten, suchten im Mittel auch deutlich seltener in der Nähe des Ziels.

Im Testdurchgang 2 lag die Gesamtfehlerrate bei 62 %. Es wurden im Gegensatz zu den anderen Bedingungen keine Alters- oder Geschlechtseffekte beobachtet. Für die Initialen Heading-Error wurden keine Differenzen zwischen erfolgreichen und nicht erfolgreichen Personen ermittelt. Relative Verweildauern differenzierten dagegen gut zwischen erfolgreichen und nicht erfolgreichen Personen. Bei nicht erfolgreichen wurde, wie in anderen Phasen der Water-Maze-Aufgabe, eine Staffelung der Gruppenmittelwerte für Alter und Geschlechts deutlich (Männer vor Frauen; jüngere vor älteren).

Zur weiteren Untersuchung der Zuverlässigkeit und Validität der entwickelten Water-Maze-Aufgabe wurden korrelative Analysen durchgeführt. Es ergaben sich mittlere positive Korrelationen zwischen der Anzahl der erfolgreichen Lerndurchgänge im VWM und den Summenscores aller Raumvorstellungstest. Der höchste Zusammenhang bestand dabei zum Konstrukt der Feldabhängigkeit. Relative Verweildauern aus dem Testdurchgang 1 korrelierten mit keinem der Raumvorstellungstests signifikant. Relative Verweildauern des Testdurchgangs 2 korrelierten mit dem Test zum visuell-räumlichen Gedächtnis. Dies galt insbesondere für die nach Alterseffekten bereinigte Partialkorrelation.

Die Pfadintegrationsleistung (gemessen über die Zeigeaufgabe) korrelierte nur mit der Anzahl der erfolgreichen Lerndurchgänge, aber nicht mit den Leistungsmaßen aus den Testdurchgängen. Innerhalb der Water-Maze-Aufgabe bestanden nur schwache Korrelationen zwischen der Anzahl der erfolgreichen Lerndurchgänge und den Testbedingungen. Die Testdurchgänge 1 und 2 waren nahezu unkorreliert.

In den qualitativen Analysen wurden Lösungsstrategien im MRT, der Desorientierungseffekt in der Kompass-Zeigeaufgabe und die Suchstrategien in der Water-Maze-Aufgabe ausgewertet. Für die Bearbeitungsstrategien im MRT konnten auch in der relativ kleinen Stichprobe vier Klassen mit unterschiedlichem Lösungsverhalten identifiziert werden. Diese unterschieden sich in den zur Lösung der Aufgaben eingesetzten Strategien und der Effizienz in der Gesamtleistung. Die meisten Personen gehörten einer Klasse mit vorzugsweise analytischen Strategien an. Eine weitere Klasse von Personen hatte offensichtlich besondere Schwierigkeiten mit dem MRT-Material. In dieser waren signifikant mehr ältere Personen. Jüngere waren dagegen häufiger in einer Klasse mit

vermutlich holistischen Strategien. Eine vierte Klasse unterschied sich von dieser dadurch, dass Personen möglicherweise auch holistische Strategien anwendeten, jedoch wesentlich langsamer. Geschlechtsdifferenzen konnten allerdings in dieser Stichprobe nicht auf die Anwendung unterschiedlich effizienter Bearbeitungsstrategien zurückgeführt werden.

Aus den Schätzfehlern der Pfad 3-Bedingung in der Kompass-Zeigeaufgabe wurden wiederum drei Gruppen für die Schätzleistung gebildet (vgl. Experiment 2, Abschn. 10.2.3, S. 112). Statistisch signifikant war die Überbesetzung der Gruppe „Desorientiert“ durch ältere bzw. weibliche Versuchspersonen. Jüngere bzw. männliche Versuchspersonen waren dagegen häufiger in der Gruppe „Orientiert“.

Durch Musteranalysen der Spurdaten aus dem Water-Maze konnten in allen Phasen der Aufgabe jeweils zwei bis vier qualitativ unterschiedliche Suchstrategien identifiziert werden. In der Exploration folgten die meisten Versuchspersonen den proximalen Hinweisen. Einige Personen zeigten Thigmotaxis, wenige durchsuchten das gesamte Inselterritorium. In den Lerndurchgängen konnten zwei Gruppen mit entweder direkten oder indirekten Suchstrategien unterschieden werden. Direkte Strategien standen auch mit sicherem, indirekte Strategien mit unsicherem Lernverhalten in Zusammenhang. Subgruppen mit direkten und indirekten Suchstrategien waren auch im Testdurchgang 1 dominant. Eine dritte Strategie, die sich als Durchsuchen an der Position einer gelöschten Landmarke beschreiben lässt, wurden hier ebenfalls identifiziert. Indirekte und direkte Suchstrategien wurden auch im Testdurchgang 2 angewandt. Letztere wurden hier aber mit Durchsuchsstrategien kombiniert. Zwei weitere Strategien beruhten auf Thigmotaxis und vermutlich fehlerhafter Pfadintegration.

Bei der korrelativen Analyse zur Anwendung der Suchstrategien in verschiedenen Testsituationen ergab sich folgendes Bild: Das Verhalten in den Lerndurchgängen und das Verhalten im Testdurchgang 1 war durch die konsistente Anwendung von direkten bzw. indirekten Suchstrategien gekennzeichnet. Personen, die in der Exploration Durchsuchen und Thigmotaxis anwendeten, wechselten im Testdurchgang 2 am häufigsten zur Strategie fehlerhafter Pfadintegration. Personen, die in der Exploration den proximalen Hinweisen folgten, wendeten im Testdurchgang 2 häufiger Thigmotaxis an.

Vorwiegend ältere Versuchspersonen folgten in der Exploration den proximalen Hinweisen, wendeten in der Lernphase und im Testdurchgang 1 indirekte Suchstrategien oder Durchsuchen und im Testdurchgang 2 Thigmotaxis an. Jüngere Versuchspersonen durchsuchten in der Exploration häufiger mit Thigmotaxis und im Testdurchgang 2 mit direkten Strategien. In der Exploration wurden vorwiegend männliche Personen mit der Strategie Durchsuchen identifiziert.

Räumliche Präferenzen wurden durch die Aggregation von Strategien aus dem MRT, dem Water-Maze und der Zeigeaufgabe abgebildet. Für analytisch arbeitende Versuchspersonen wurden so Zusammenhänge zum Explorationsverhalten ermittelt. Diese Personen wandten in der Exploration und im Testdurchgang 2 eher die proximale Strategie bzw. Thigmotaxis an. Zusammenhänge zwischen den Präferenzen analytisch vs. holistisch und egozentrisch vs. allozentrisch bestanden nicht.

Diskussionsteil

12 Einzeldiskussionen

12.1 Experiment 1 - Mentale Rotation

Ziel des Experiments war die Untersuchung von Lösungs- und Bearbeitungsstrategien und deren Bedeutung für die Geschlechtseffekte in verschiedenen Tests, die die Facette mentale Rotation operationalisieren. Zur Identifikation der Strategien wurde ein neuartiges, strukturentdeckendes Klassifizierungsverfahren eingesetzt (Gao & Church, 2005). Als Messinstrumente kamen der *Mental Rotations Test* (MRT; Peters et al., 1995) und die *Würfelaufgaben* aus dem Intelligenzstrukturtest (I-S-T 2000; Amthauer et al., 1999) zum Einsatz. Für beide Tests wurden bereits mit Modellen der Item-Response-Theorie latente Subgruppen unter den Versuchspersonen entdeckt, die sich in der Anwendung analytischer bzw. holistischer Lösungsstrategien unterschieden (Geiser et al., 2006; Hosenfeld et al., 1997).

Während der MRT jedoch einer der zuverlässigsten Tests zur Demonstration von Geschlechtsdifferenzen ist (Kimura, 1999; Halpern, 2000; Masters & Sanders, 1993), sind in den *Würfelaufgaben* Geschlechtseffekte in vergleichbarer Höhe nicht vorhanden (Hosenfeld et al., 1997). Auch in der vorliegenden Untersuchung bestätigte sich der quantitative Vorteil männlicher Versuchspersonen im MRT. Die für diesen Test typischen großen Geschlechtsdifferenzen konnten in allen hier untersuchten Altersgruppen nachgewiesen werden und lagen über der für Erwachsene oft berichteten Effektstärke von $d=0.7$ (Linn & Petersen, 1985). Ein enorm großer Geschlechtseffekt bestand in der Stichprobe der Studenten mit $d=2.0$. Dieser geht vermutlich auf einen Bias durch das Studienfach zurück. Wie Peters et al. (1995) und Lehmann (2000) zeigen konnten, sind Studenten naturwissenschaftlich-technischer Fachrichtungen Studenten sozial- und geisteswissenschaftlicher Studienfächer im MRT überlegen. Die ungünstige Verteilung der Studienrichtung auf die Studentinnen und Studenten muss hier zur Verstärkung des Geschlechtseffekts geführt haben.

Darüber hinaus wurde mit $d=1.5$ eine sehr große Geschlechterdifferenz im MRT für Schüler im achten Schuljahr (Alter: 14 – 15 Jahre) ermittelt. Ab diesem Lebensalter war auch ein sprunghafter Anstieg der Geschlechterdifferenzen zu verzeichnen. Dies geht möglicherweise auf eine beginnende stärkere hormonelle Differenzierung bei Jungen und Mädchen in diesem Alter zurück. Höhere Östrogenkonzentrationen, wie sie zu diesem Zeitpunkt bei den Mädchen möglich gewesen sein könnten, führen zu einer Verringerung dieser Raumvorstellungsfähigkeit (Hausmann et al., 2000). Erhöhte Androgenkonzentrationen (vorrangig Testosteron), kurz vor Beginn der Pubertät liegen diese bei Jungen in zwei- bis dreifach erhöhter Konzentration vor (Keidel, 1970), beeinflussen dagegen die MRT-Leistung positiv (Bell & Saucier, 2004).

Andere Faktoren, wie Schulleistung, Schulabschluss, die „induktive Denkfähigkeit“ und räumlich-perzeptive Fähigkeiten, wurden kontrolliert und ergaben eine weitgehende Vergleichbarkeit männlicher und weiblicher Versuchspersonen. Bezüglich des Umgangs mit Computerspielen waren Schüler und Studenten gegenüber den Schülerinnen und Studentinnen allerdings deutlich überlegen. Für eine häufigere Anwendung von Computerspielen seitens männlicher Versuchspersonen wird eine positive Auswirkung auf deren MRT-Leistung vermutet, allerdings nur für Spiele, die die entsprechenden Fähigkeiten erfordern (Quaiser-Pohl, Geiser & Lehmann, 2005). Die Art der verwendeten Spiele wurde in der durchgeführten Untersuchung nicht erfasst. Wie dem auch sei, da der Unterschied in der Häufigkeit von Computerspielen zwischen Jungen und Mädchen derart groß war und Mädchen meistens überhaupt nicht am Computer spielten, ist für die Aufklärung der Geschlechtsunterschiede im MRT die Art der Spiele hier nicht weiter relevant gewesen. Lediglich für die Leistungsunterschiede innerhalb der Gruppe der Jungen könnten diese eine Rolle gespielt haben. Somit nimmt der Faktor „Computererfahrung“ eher eine vermittelnde Rolle zwischen Geschlecht und MRT-Leistung ein.

Allgemein würde eine dauerhafte Manifestierung von Geschlechtsdifferenzen auf einem solchen intermediärendem Faktor auch eher für die Persistenz von Geschlechtsdifferenzen in der Raumvorstellungsfacetten mentale Rotation als für ein allmähliches Verschwinden sprechen (Masters & Sanders, 1993). Allerdings gilt dies nicht uneingeschränkt für alle Facetten, ja nicht einmal für Tests, die ähnliche Fähigkeiten operationalisieren. Im Test *Würfelaufgaben* wurden auch in der vorliegenden Studie keine signifikanten Geschlechtsdifferenzen nachgewiesen. Dies bestätigte die von Hosenfeld et al. (1997) gemachten Beobachtungen. Innerhalb des I-S-T werden zwar über die *Würfelaufgaben* räumliche Fähigkeiten operationalisiert, jedoch zählen dazu neben der Bewegungsvorstellung auch räumlich-topographische und nicht-räumliche Fähigkeiten, wie Symbolverständnis und induktives Schließen, die im I-S-T zu räumlich-figuralen Fähigkeiten zusammengefasst und allgemein auf dem Faktor „figurale Intelligenz“ abgebildet werden (Amthauer et al., 1999). Trotzdem wurden im durchgeführten Experiment quantitative Zusammenhänge zwischen beiden Tests festgestellt. Das Interessante daran: In einem eher exploratorischen Vorgehen wurden stärkere Zusammenhänge zwischen den Leistungsmaßen bei einer geschlechtsspezifischen Betrachtung entdeckt. Es scheint, dass männliche Versuchspersonen in beiden Tests ähnlichere Leistung erbringen als weibliche Versuchspersonen.

Die wichtigste Hypothese des Experiments ging davon aus, dass im MRT und in den *Würfelaufgaben* von den Probanden unterschiedliche Lösungs- und Bearbeitungsstrategien eingesetzt werden. Diese Strategien sollten sich vor allem in ihrer Effizienz unterscheiden und somit quantitative Unterschiede im Test erklären. In diesen qualitativen Analysen wurden Methoden der automatischen Mustererkennung und Klassifizierung (SNMF und Consensus-Clustering) zur Entdeckung der im MRT-A und den *Würfelaufgaben* eingesetzten Lösungsstrategien erprobt. Es konnten auf den Datensätzen der ca. 200 Schüler jeweils für den MRT und die *Würfelaufgaben* Modelle mit drei Klasse identifiziert werden.

Zur Modellselektion wurden die in Monti et al. (2003) vorgeschlagenen Kriterien φ -k-Diagramme und Consensus-Plots verwendet. Da bisher unbekannt war, wie die Heuristik auf dichotomen Daten-

sätzen mit wenigen Variablen arbeitet, wurden zusätzliche Informationen über das Ziel der Klassifizierung in den Auswahlprozess des Modellparameters aufgenommen. Dies erfolgte anhand quantitativer Maße zur Einschätzung der Differenzierungsfähigkeit der Modelle hinsichtlich spezieller Itemtypen.

Die im MRT identifizierten Klassen wurden in ihrer Bedeutung für die Art der verwendeten Strategie anhand von Ratio-Scores und Lösungshäufigkeiten der Itemtypen, der Itemprofile, der Gesamtleistung im Test und der Korrelationen mit Außenkriterien interpretiert.

Aus diesen Betrachtungen wurde eine Klasse (Klasse 3) deutlich, die weder Schwierigkeiten mit den Items des Tests noch mit der Zeitbegrenzung hatte. Die Strategien dieser Klasse wurden unter dem Begriff „Schnelles Rotieren“ zusammengefasst. Personen dieser Klasse bearbeiteten und lösten im Vergleich zu den anderen Klassen die meisten Items im Test. Personen einer weiteren zeigten ähnliches Verhalten, konnten jedoch nicht alle Items im Zeitlimit lösen (Klasse 1). Es wurde vermutet, dass Personen dieser Klasse mentale Rotation ausführen, aber wesentlich langsamer. Personen dieser Klasse bearbeiteten insgesamt auch die wenigsten Items im Test. Zahlenmäßig war diese Klasse mit 43 % der Probanden die größte der identifizierten Klassen.

Personen der Klasse 2 wiesen, ähnlich wie Personen der Klasse 1, Schwierigkeiten mit dem Zeitlimit auf. Aber diese Klasse unterschied sich von den anderen beiden Klassen in den nach Itemtypen gemittelten Ratio-Scores und Lösungshäufigkeiten. Typ-I-Items und Items mit Okklusion lösten Personen dieser Klasse seltener, aber sie erzielten ähnliche Resultate für Typ-II-Items. Deshalb wurde geschlussfolgert, dass Personen dieser Klasse vorzugsweise Items lösten, die sich mit Detailvergleichen bearbeiten lassen. Insgesamt bearbeitete diese Klasse im Mittel mehr Items als die Klasse 1, erzielte aber genau die gleiche Punktzahl.

Für die Validität dieser Klasseneinteilung sprechen die klassenspezifischen Korrelationen mit den Außenkriterien. So korrelierte die Leistung aus dem Test zur „induktiven Denkfähigkeit“ mit dem MRT-Score nur in der Klasse der analytischen Strategien (Klasse 2) signifikant und in Höhe eines mittleren Zusammenhangs. Geringere, aber signifikante Korrelationen mit räumlich-perzeptiven Fähigkeiten, erfasst über den Untertest *Flächenzählen*, bestanden insbesondere in der Klasse „Langsames Rotieren“. Dies könnte bedeuten, dass auch in dieser Klasse die Analyse räumlicher Details eine größere Rolle spielte als für Personen der Klasse „Schnelles Rotieren“. Der Unterschied zur Klasse der analytischen Strategien bestand möglicherweise darin, dass Figuren in ihren Teilflächen rotiert wurden.

Die in den *Würfelaufgaben* identifizierten Klassen wurden durch Analyse der itemspezifischen Lösungshäufigkeiten und der Summenscores interpretiert. Ratio-Scores und Itemprofile lieferten hier keine zusätzlichen Informationen. Spezifische Antwortmuster wurden vor allem in den Lösungshäufigkeiten für „Raumwürfel“ und „Flächenwürfel-Plus“ ausgemacht. Es ergab sich eine in der Logik etwas andere Art der Zusammensetzung der Klassen als im MRT. Probanden schienen entweder „Raumwürfel“ oder „Flächenwürfel“ und insbesondere „Flächenwürfel-Plus“ gut bearbeiten zu können. Besonders deutlich unterschieden sich hier die Klassen 1 und 3. Die Klasse 1 war die einzige Klasse, deren Personen im Mittel mehr als 50 % der Aufgaben zu „Raumwürfeln“ lösten.

Alle anderen Klassen lagen unter 50 %. Die Klasse 3 löste dagegen als einzige über 50 % der Aufgaben zu „Flächenwürfel-Plus“. Diese Klasse war auch mit fast 60 % der Versuchspersonen die größte identifizierte Klasse. Zwischen beiden Klassen, aber in den itemspezifischen Lösungshäufigkeiten eher wie Klasse 1, lag noch eine zahlenmäßig sehr kleine Klasse. In der Interpretation von Hosenfeld et al. (1997) würde das Vorgehen der Personen in Klasse 1 und eventuell auch in Klasse 2 der Anwendung von holistischen Strategien entsprechen. Die große Klasse 3 würde der Gruppe von Personen entsprechen, die mit analytischen Strategien Raumvorstellungstests bearbeiten.

Durch die Klasseneinteilung konnte aber keine Differenzierung der Summenscores erfolgen. Offensichtlich kompensiert der Summenscore in den Würfelaufgaben unterschiedliche Bearbeitungsstrategien noch stärker als dies im MRT der Fall ist. Lediglich in der Anzahl der bearbeiteten Items lag die Klasse der analytischen Strategien etwas unter der der anderen Klassen und wies auf Grund der deutlich größeren Anzahl von Personen auch eine größere Streuung dieses Merkmals auf. Insgesamt wurden aber in allen Klassen fast alle Items bearbeitet, weshalb Ratio-Scores hier auch keine zusätzliche Information brachten.

Die Validität der automatischen Klassifizierungsverfahren SNMF und Consensus-Clustering konnte durch plausible und mit bekannten Befunden übereinstimmende Resultate demonstriert werden. Sowohl für den MRT als auch für die *Würfelaufgaben* wurden Subgruppen von Personen identifiziert, die sich durch die Anwendung holistischer bzw. analytischer Strategien auszeichnen. Dazu gehören auch die bestätigten Zusammenhänge zwischen Geschlechtszugehörigkeit und der mit dieser Methode ermittelten Klassifizierungen.

Zusammenhänge zu den Klasseneinteilungen und der Geschlechtszugehörigkeit zeigten sich sehr deutlich im MRT. Signifikant mehr männliche Versuchspersonen wurden der Klasse „Schnelles Rotieren“ zugeordnet. Mädchen waren dagegen häufiger in der Klasse „Langsames Rotieren“. In der hier als Gruppe mit analytischen Strategien identifizierten Personen waren die Geschlechter gleichverteilt. In diesen Punkten stimmen die vorliegenden Ergebnisse mit den in Geiser (2004) dargestellten Zusammenhängen zwischen Geschlechtszugehörigkeit und der Strategiewahl im MRT erstaunlich gut überein. In den *Würfelaufgaben* wurde kein signifikanter Zusammenhang zwischen Strategiewahl und Geschlecht deutlich. Ein Befund der auch in der von Hosenfeld et al. (1997) durchgeführten Analyse zu Bearbeitungsstrategien in den *Würfelaufgaben* vorgelegt wurde.

Deutliche Zusammenhänge bestanden allerdings zwischen dem Schuljahrgang und der MRT-Strategie. Insgesamt wurde eine Zunahme der effizientesten Strategien ab dem achten Schuljahr beobachtet. Über 60 % der Jungen und etwas mehr als 30 % der Mädchen wurden mit der Strategie „Schnelles Rotieren“ klassifiziert. In den Schulstufen fünf bis sieben waren weniger als 25 % einer Jahrgangsstufe in dieser Strategiekategorie. Wobei auch in diesen Jahrgangsstufen Jungen diese Strategiekategorie dominierten. In den Jahrgängen fünf und sechs wurde kein einziges Mädchen mit dieser Strategie identifiziert. Durch diese Verteilung der Strategien nahm unter dem Einfluss der MRT-Strategie auf den Summenscore auch der Geschlechtseffekt im achten Schuljahr zu. Die beobachtete Diskrepanz zwischen den Geschlechtern in der Verteilung auf die effizientesten Strategie war am stärksten in diesem Alter, so dass zunehmende Leistungsdifferenzen eindeutig auf die veränderte Verteilung der Strategien zurückzuführen waren.

Es konnte also gezeigt werden, dass in diesem Alter eine Differenzierung der Leistungsfähigkeit der mentalen Rotation von Jungen und Mädchen stärker auf eine qualitative Veränderung zurückzuführen ist. Es ist nicht nur die Ausprägung der eindimensionalen Fähigkeit, die sich verändert, sondern der Prozess der Informationsverarbeitung wird gänzlich anders organisiert. Unter Umständen wirkt insbesondere Testosteron bei der Veränderung und Ausbildung entsprechender neuronaler Strukturen. Dass es sich lediglich um die Auswirkung der fortgeschritteneren Schulbildung (Kohorteneffekt) handelt, kann ausgeschlossen werden, da der Effekt für Jungen und Mädchen unterschiedlich stark wirkt.

Nicht bestätigt haben sich die erwarteten positiven Zusammenhänge zwischen der räumlichen Präferenz und der Strategiewahl in verschiedenen Tests zur mentalen Rotation. Die Zusammenhänge, korreliert über das hypothetische Konstrukt einer räumlichen Präferenz analytisch vs. holistisch, waren nicht signifikant und in der Tendenz sogar konträr.

Für die Beurteilung wurden die Strategieklassen „Schnelles Rotieren“ und „Langsames Rotieren“ aus dem MRT auf dem Pol einer eher holistischen Präferenz zusammengefasst. Die Klasse der analytischen Bearbeitungsstrategien verblieb komplett auf dem Pol der analytischen Präferenz. Zusammenhänge zwischen diesen Kategorien und den Lösungsstrategien in den *Würfelauflagen* waren nur tendenziell und im Vorzeichen negativ. Es ergab sich eine nominale Dominanz von Personen der Strategiegruppe „Raumwürfel“ mit einer eher analytischen Präferenz im MRT. Das könnte auf der einen Seite bedeuten, dass die Art des Testmaterials maßgeblich die Wahl der Strategie beeinflusst, so dass die entsprechenden Verteilungen zustande kamen. Auf der anderen Seite könnte aber auch die Art der Aggregation der Präferenzklassen nicht valide sein. Das wiederum ließ sich mit dieser Art der Analyse der Antwortmuster nicht genauer entscheiden.

Aus der Auswertungen der Strategiefragebögen konnten dazu ebenfalls keine zuverlässigen Erklärungen abgeleitet werden. Nach Zusammenfassung der Fragebogenitems zu inhaltlichen Faktoren bestanden auf dieser Ebenen keine Zusammenhänge zu den implizit detektierten Strategien und den explizit angegebenen. Insgesamt muss sogar festgestellt werden, dass die in der Kategorie holistisch vs. analytisch aus den Fragebögen ermittelten Strategien mit keiner der unabhängigen Variablen („Strategie“ und „Geschlecht“) in Zusammenhang standen.

Sekundäres Ziel des Experiments war die Untersuchung quantitativer Leistungsdifferenzen im Computertests und dem Papier-Bleistift-Test des MRT. Es konnten erste Hinweise erbracht werden, dass die verwendete eigene Adaption des Computertests zur Papier-Bleistift-Version abweichende Ergebnisse liefert. Im Computertest wurden tendenziell niedrigere Summenscores für weibliche Versuchspersonen gemessen als in der ursprünglichen Version. Als Fazit sollte deshalb die Interaktion mit dem Geschlecht beim Einsatz des entsprechenden Test berücksichtigt werden.

Diese Ergebnisse begründen sich auf einem Prä-Post-Test, in dem eine Parallelisierung der Versuchspersonen nach MRT-Vortestleistung, der „induktiven Denkfähigkeit“, der Computererfahrung, dem Alter und dem Geschlecht erfolgte. Somit war eine Vergleichbarkeit der Post-Test-Gruppen Papier-Bleistift und Computertest gewährleistet. Der Post-Test erfolgte nach genau vier Wochen.

Für diesen Post-Test wurden in Kovarianzanalysen differentielle Effekte der Testversion innerhalb der Subgruppe der Schülerinnen nachgewiesen. Dieser Unterschied manifestierte sich in einem signifikanten Unterschied der Regressionsgeraden zwischen der Papier-Bleistift-Gruppe und der Computergruppe. Mädchen, die den Computertest absolvierten, profitierten von allen Gruppen am wenigsten von der Testwiederholung. Der Geschlechtseffekt verblieb, auch nach der zeitnahen Testwiederholung in der Computertestgruppe bei fast $d=0.7$. Möglicherweise würde bei einer Durchführung ohne Prä-Test der Geschlechtseffekt in der Computerversion generell höher ausfallen als im ursprünglichen Papier-Bleistift-Test. In der Kovarianzanalyse wurden durch diesen differentiellen Effekt der Testversion 14 % der Varianz in der Stichprobe der Schülerinnen aufgeklärt.

Zur Analyse der Unterschiede im Computertest könnte man vermuten, dass auch hier durch den Test unterschiedliche Strategien induziert wurden. Ein Post-Hoc-Test zur Verteilung der Strategien aus dem Prä-Test erbrachte zunächst keine Unterschiede zwischen beiden Gruppen. Möglicherweise verwendeten Mädchen dann aber im Computertest doch andere Strategien, ausgelöst durch die spezielle Testsituation. Zur Aufklärung dieses Sachverhalts können auch die impliziten Analysemethoden herangezogen werden. Entsprechende Auswertungen wurden in die vorliegende Arbeit aus Platzgründen zunächst nicht mit aufgenommen, sind aber bereits durchgeführt worden. Die Ergebnisse dieser Analysen zeigten, dass nur im Computertest Mädchen signifikant häufiger analytische als holistische und Jungen häufiger holistische als analytische Strategien einsetzten (Schönfeld, Lehmann, Ungewiß & Lepow, 2007). Dieser Unterschied bestand im Papier-Bleistift-Test so nicht. Dies unterstreicht die Interaktion von Testsituation, Strategiewahl und dem Faktor „Geschlecht“ bei der Bearbeitung des MRT.

Ein weiterer erster Ansatz zur Erklärung der Testunterschiede konnte aus den Fragebögen abgeleitet werden. Hier gaben Mädchen der Computertest-Gruppe häufiger an, bei der Lösung der Items geraten zu haben als Mädchen der anderen Gruppe. Diesen Hinweisen könnte man ebenfalls mit impliziten Itemanalysen nachgehen. Ein solches Rateverhalten würde sich in bestimmten Mustern innerhalb der Items äußern. Eine genauere Analyse auf dieser Ebene wurde beispielsweise bei Voyer und Hou (2006) durchgeführt und ergab, dass im MRT üblicherweise Jungen stärker zum Raten neigen als Mädchen. Möglicherweise wird diese Verhaltensweise ebenfalls durch die Art der Testdarbietung, d. h. Computer statt Papier-Bleistift, beeinflusst. Zusätzlich könnten für diese Fragestellung aus dem Computertest Reaktionszeiten auf Itemebene ausgewertet werden. In der verwendeten Version des Computertests war diese Möglichkeit allerdings noch nicht vorgesehen.

Darüber hinaus stellte der Post-Test am Computer eine neue Situation dar, auf die Mädchen möglicherweise anders reagierten als Jungen. Somit bestand eine Interaktion des Post-Tests mit dem Faktor „Neuheit“, so dass der Vergleich Computertest vs. Papier-Bleistift-Test auch eher ein Vergleich Neuheit vs. Wiederholung darstellt.

Für generelle Schlussfolgerungen hinsichtlich der Testversion sind jedoch auch die Einschränkungen des Designs zu berücksichtigen. So bestand in der Stichprobe ein deutlicher Deckeneffekt durch die zeitnahe Messwiederholung. Außerdem handelte es sich bei Schülern eines mathematisch-naturwissenschaftlichen Spezialgymnasiums um eine Stichprobe einer sehr eingeschränkten Popu-

lation. Die Reliabilität der Resultate ist also noch mit weiteren Stichproben zu prüfen. Ein Hinweis lieferte in der eigenen Untersuchung die Stichprobe der Studenten, in der der Trend genau umgekehrt war. Allerdings war die Stichprobe vor allem im Computertest zu klein, um daraus gesicherte Erkenntnisse zu gewinnen.

Am Schluss der Diskussion sollen aber auch inhaltliche und methodische Kritikpunkte stehen. Auf Grund der nachgewiesenen Interaktion der Faktoren „Geschlecht“ und „Strategie“ in Raumvorstellungstests ist eine getrennte Interpretation der Effekte schwer möglich. Deshalb kann auch mit dieser Untersuchung nicht geklärt werden, warum Mädchen in diesen Tests weniger effiziente Strategien nutzen. Der Erkenntnisgewinn bleibt kategorial. Dennoch können die nicht gefunden Zusammenhänge zwischen den Lösungsstrategien im MRT-A und den *Würfelaufgaben* und die stärkere Geschlechtsspezifität im Computertest als Hinweise darauf interpretiert werden, dass Strategien generell eher situationsspezifisch eingesetzt werden, aber möglicherweise Mädchen sensibler auf Testmaterial und -situation reagieren als Jungen. Woraus folgt, dass die Stabilität der Entscheidung für eine Strategie ein Personenmerkmal darstellt, das in seiner Ausprägung aber vom Geschlecht der Person abhängt.

Für den MRT und die *Würfelaufgaben* sind hier möglicherweise auch die Entscheidungsprozesse innerhalb der Items für unterschiedliche Strategien mitverantwortlich. So mussten sich Probanden im MRT immer für zwei richtige Figuren entscheiden, während in den *Würfelaufgaben* jeweils nur eine richtige Antwort gegeben werden musste. Bei Entscheidungsprozessen wird generell nicht von Geschlechtsdifferenzen ausgegangen. So konnten u. a. Kerkman, Wise und Harwood (2000) bei Paarvergleichen zur mentalen Rotation von gänzlich anderem räumlichen Testmaterial keinen Geschlechtseffekt nachweisen. Das könnte bedeuten, dass zwar in Raumvorstellungstests unterschiedliche Lösungsstrategien zum Einsatz kommen, sie aber nur in Interaktion mit dem Faktor „Geschlecht“ treten, wenn die Entscheidungsprozesse zur Auswahl der Antwortalternativen eine bestimmte Komplexität überschreiten.

Neben inhaltlichen Einschränkungen der Untersuchungen sind auch methodische zu nennen. An der Klassifizierungsmethode ist zunächst kritisch zu bemerken, dass sie trotz der eindrucksvollen Konvergenz mit bereits vorliegenden Ergebnissen in ihren Modellauswahlkriterien wesentlich schwächer ist als die in Geiser et al. (2006) und Hosenfeld et al. (1997) verwendeten Modelle der IRT. So gibt es kein Maß für die Wahrscheinlichkeit der Existenz einer Klasse und der Zuverlässigkeit der einzelnen Klassenzuweisungen. Außerdem wurden für die Interpretation der Klassen keine absoluten Kriterien angewandt, sondern die inhaltliche Bestimmung der Strategien erfolgte komparativ nur innerhalb der jeweiligen Stichprobe. Damit weist die Methode insgesamt eine schwache Auswertungsobjektivität auf. Daraus ergibt sich auch eine fehlende Möglichkeit zur quantitativen Überprüfung der Reliabilität der Klassifizierung.

Für die Methode spricht jedoch die Identifikation von drei Klassen mit augenscheinlich sehr ähnlichen Strategieprofilen zur Untersuchung von Geiser et al. (2006). Weshalb hier keine fünf Klassen identifiziert wurden, könnte am geringeren Stichprobenumfang liegen. Außerdem zeichneten sich die fehlenden zwei Klassen durch besonders schlechte Performanz aus. Innerhalb der hier unter-

suchten Stichprobe gab es statistisch nicht genügend Personen, die wirklich schlechte Resultate im MRT erzielten.

Perspektivisch könnte jedoch eine „Verbesserung“ des MRT in Richtung eines eindimensionalen Fähigkeitstests durchgeführt werden, der generell nur aus Typ-I-Items besteht. Die Konsequenz dieser Modifikation wäre, dass Personen die keine mentale Rotation ausführen, praktisch null Punkte erzielen. Allerdings könnten die nicht vorhandenen Typ-II-Items eine andere Art der Strategiewahl bewirken, indem analytische Strategien von den Versuchspersonen eventuell gar nicht erst in Erwägung gezogen werden. Die Überprüfung dieser Hypothese wäre Aufgabe einer weiteren Untersuchung.

12.2 Experiment 2 - Pfadintegration

Im Experiment wurde die Frage untersucht, ob mit Methoden virtueller Umgebungen, die Bewegungsinformationen ausschließlich über visuelle Quellen (optischer Fluss) vermitteln, Orientierungsleistungen operationalisiert werden können, die auf Pfadintegrationsprozessen beruhen. Das eingesetzte Messinstrument war eine selbst entwickelte Kompass-Zeigeaufgabe, über die die Leistungsfähigkeit einer Versuchsperson über die absolute Differenz zwischen der geschätzten und der tatsächlichen Heimkehrrichtung ermittelt wurde. Der Transport durch die virtuelle Umwelt erfolgte passiv. Zur Beantwortung der Fragestellung wurden Hypothesen zu allgemeinen und differentiellen Effekten der Zeigeaufgabe überprüft.

Bestätigt wurde die Hypothese, dass der absolute Schätzfehler in der Kompass-Zeigeaufgabe mit der Komplexität der Pfade zunimmt. Bestimmt man die Komplexität eines Pfades über die Anzahl der Abzweigungen und deren Winkel, so traf dies uneingeschränkt zu. Der Einfluss der Darbietungsgeschwindigkeit auf die abhängigen Variablen konnte auf Grund methodischer Probleme im vorliegenden Experiment nicht geklärt werden.

Mittlere Schätzfehler für Bedingungen, in denen nur 90°-Turns absolviert wurden, lagen innerhalb eines akzeptablen Bereichs, wie er auch bei Lawton und Morrin (1999) für virtuelle Zeigeaufgaben berichtet wurde. Die Streuungen um den Mittelwert waren aber relativ groß, was auch auf die Möglichkeit größerer Ausreißerwerte innerhalb der Schätzungen zurückzuführen ist. In den Häufigkeiten lagen die „guten“ Schätzungen, zu denen alle bis ca. 36° Schätzfehler gezählt wurden, allerdings sehr hoch.

Bedingungen mit 135°-Turns führten zu den mit Abstand größten mittleren Schätzfehlern und noch größeren Streuungen. Der mittlere Schätzfehler lag beinahe als gleichverteiltes Merkmal vor. Im Vergleich zu Pfaden mit 90°-Turns stellte die Durchführung von 135°-Turns offensichtlich eine massive Manipulation der Versuchsbedingung dar. Schätzfehler stiegen bei der Verwendung von 135°-Turns statt der voran genannten 90°-Turns um 50°. Im Vergleich dazu stieg der Schätzfehler um weniger als 10° bei Verwendung von drei statt zwei 90°-Turns.

Im Ergebnis lässt sich feststellen, dass die Komplexität der Pfade den Schätzfehler in der erwarteten Richtung beeinflusst. Unter der Komplexität kann hier allerdings nur ein ordinalskaliertes Merkmal

verstanden werden, dass mit Anzahl und Größe des Winkel der Abzweigungen (Turns) ansteigt. Das gilt zumindest für die hier verwendeten Faktorstufen zwei bzw. drei Abzweigungen und 90° bzw. 135°. Einschränkend gilt allerdings, dass bei den im Versuch verwendeten Pfaden mit der Anzahl der Abzweigungen auch die Weglänge variierte.

Variierende Darbietungsgeschwindigkeiten äußerten sich dagegen nicht in den Schätzfehlern. Es zeigten sich hohe Korrelationen zwischen Bedingungen mit gleichen Pfaden und unterschiedlichen Darbietungsgeschwindigkeiten. Dies lag vermutlich im Versuchsaufbau begründet. Für die Variation der Versuchsbedingung „Geschwindigkeit“ wurden die gleichen Pfade in der gleichen Reihenfolge verwendet, so dass von Übungs- bzw. Erinnerungseffekten ausgegangen werden muss, die einen möglichen Einfluss der Geschwindigkeit kompensierten. Dass Versuchspersonen die Pfade in der Wiederholung wiedererkannten, zeigte sich auch im Fehlen eines Trainingseffekts, wie er von Lawton und Morrin (1999) in der gleichen Aggregation von Versuchsbedingungen beobachtet wurde. Die Schlussfolgerung, dass die Darbietungsgeschwindigkeit generell keinen Einfluss auf die Richtungsschätzungen hat, konnte deshalb nicht gezogen werden.

Die zweite Hypothese des Experiments, in der angenommen wurde, dass die Schätzfehler generell mit dem Alter der Versuchspersonen ansteigen, konnte bestätigt werden. Das traf sowohl für den mittleren Gesamtfehler als auch für die mittleren Schätzfehler in den einzelnen Bedingungen zu. Eine in die gleichen Richtung formulierte Hypothese bezüglich der Reaktionszeiten konnte dagegen nicht bestätigt werden.

Die Differenz der Gruppenmittelwerte des Gesamtfehlers betrug zwischen jüngeren und älteren (mittlere und ältere Gruppe) je 20°. Differenzen in dieser Größenordnung wurden u. a. auch in Foreman, Stanton-Fraser, Wilson, Duffy und Parnell (2005) berichtet. Dort lagen die Unterschiede zwischen jüngeren Erwachsenen und älteren bei zirka 30°. Allerdings waren ältere Versuchspersonen bereits in einem Alter von 62 bis 82 Jahren. In den hier vorliegenden Ergebnissen zur Schätzgenauigkeit unterschieden sich bereits jüngere Erwachsene (20-29 Jahre) von Erwachsenen im Alter zwischen 38 bis 43 Jahren. Möglicherweise nimmt die Fähigkeit zur PI in virtuellen Umgebungen schneller ab als erwartet. Obwohl der Alterseffekt in der virtuellen Kompass-Zeigeaufgabe mit dem im Test zur räumlichen Wahrnehmung (Feldabhängigkeit) und zum visuell-räumlichen Gedächtnis vergleichbar war, wurde in den klassischen Raumvorstellungstests eine derartig große Leistungsdifferenz zwischen jüngerer und mittlerer Altersgruppe nicht beobachtet. In der Zeigeaufgabe könnte somit der Faktor Computererfahrung, der in diesem Experiment nicht erfasst wurde, eine Rolle gespielt haben.

Ähnliche Überlegungen ergeben sich aus der Auswertung der Reaktionszeiten. Der von Lemay et al. (2004) berichtete Alterseffekt wurde nicht bestätigt. Aus den Beobachtungen ging hervor, dass insbesondere ältere Erwachsene im Umgang mit Computern nicht so versierte Versuchspersonen sind und mit der implementierten Drag-&-Drop-Technik zur Bedienung der Kompassnadel Schwierigkeiten hatten. Die erfassten Reaktionszeiten können nicht als zuverlässiges Maß angesehen werden, da sie offensichtlich mit motorischen Fertigkeiten konfundiert sind. Bei der Entwicklung der Zeigeaufgabe wurde aber davon ausgegangen, dass Drag-&-Drop eine einfache intuitive

Technik zur Bedienung von VR-Objekten darstellt (Bowman et al., 2001). Hier zeigt sich einmal mehr die Diskrepanz zwischen allgemeinen Hinweisen zur Entwicklung von VRs und den differentiellen Effekten seitens der Nutzer.

Nur teilweise wurde die Hypothese bestätigt, dass männliche Versuchspersonen in ihren Richtungsschätzungen weiblichen überlegen sind. Ein entsprechender Geschlechtseffekt, der auch noch nominal geringer war als vermutet, wurde nur für Pfade mit zwei rechtwinkligen Abzweigungen nachgewiesen. Männliche Versuchspersonen gaben in diesen Bedingungen um ca. 14° bessere Schätzungen ab als weibliche Versuchspersonen. Die Mittelwertsdifferenz betrug allerdings weniger als 0.4 Standardabweichungen. Der Effekt kehrte sich bei Abzweigungen mit 135° -Turns tendenziell sogar um. Diese Ergebnisse weichen deutlich von den in Lawton und Morrin (1999) berichteten großen Geschlechtsdifferenzen für Zeigeaufgaben in virtuellen Umgebungen ab. Die Schätzfehler weiblicher Versuchspersonen lagen bei Lawton und Morrin (1999) mit einer Differenz von 20° über denen der männlichen in allen von den Autoren untersuchten Bedingungen. Dieser Unterschied zu den Ergebnissen der eigenen Studie könnte in den unterschiedlichen Implementierungen der Zeigeaufgabe begründet sein.

Der deutlichste Unterschied zwischen der Zeigeaufgaben von Lawton und Morrin (1999) und der eigenen bestand in der Art der Abfrage. So mussten die Probanden im Experiment von Lawton und Morrin (1999) die Richtung auf einer Pappscheibe (außerhalb der virtuellen Umgebung) einstellen. Diese Art der Abfrage erfordert zusätzliche Umrechnungen der mentalen Repräsentation der Heimkehrrichtung in das außerhalb der virtuellen Umgebung liegende Referenzsystem des Zeigeinstruments. Eventuell liegen in dieser Fähigkeit stärkere geschlechtsspezifische Differenzen vor. Dass die Art der Abfrage für die Genauigkeit relevant ist, konnten Avraamides, Klatzky, Loomis und Golledge (2004) in Zusammenhang mit der Untersuchung von verschiedenen Response- und Transportmethoden zeigen. Einen Hinweis darauf, dass Responsemethoden in Zeigeaufgaben eventuell auch mit dem Faktor „Geschlecht“ interagieren, liefern die Reaktionszeiten aus dem eigenen Experiment. In der Auswertung dieser ergaben sich für weibliche Versuchspersonen im Mittel zehn Sekunden kürzere Antwortzeiten als für männliche Versuchspersonen. Das könnte bedeuten, dass mit Hilfe der hier gewählten Respons-Methode Frauen ihre Schätzungen schnell und intuitiv abgaben. Fehleranfällige mentale Transformationen wurden nicht durchgeführt und waren für eine korrekte Schätzung nicht notwendig. Zu berücksichtigen sind allerdings die bereits diskutierten Probleme in der Erfassung der Reaktionszeiten.

Korrelative Untersuchungen zur diskriminanten und konvergenten Validität der Zeigeaufgabe bestätigten die Dissoziation des Konstrukts Pfadintegration von anderen räumlichen Fähigkeiten. So korrelierte der Schätzfehler aus einzelnen Pfaden mit Scores der Raumvorstellungstests nur schwach (diskriminante Validität). Korrelationen des Gesamtfehlers mit dem LPS-Untertest 10 (räumliche Wahrnehmung und Erkennen von Mustern bzw. Feldabhängigkeit) und dem RCF-delay (visuell-räumliches Gedächtnis) lagen im mittleren Bereich (konvergente Validität). Diese waren vergleichbar zu den Korrelationen der Scores der beiden Raumvorstellungstests untereinander. Nach Partialisierung der Alterseffekte korrelierte nur noch der Gesamtfehler mit dem Leistungsmaß aus dem Test zur Feldabhängigkeit/räumlichen Wahrnehmung. Die mittleren Fehler aus Pfad 1 und 2 korrelierten

ebenfalls in der Größenordnung eines mittleren Zusammenhangs, sowie hoch mit dem Gesamtfehler (interne Konsistenz). Bemerkenswert muss aber, dass der mittlere Gesamtfehler durch den hohen Anteil des absoluten Fehlers aus Pfad 3 bestimmt wird.

Exploratorisch wurde eine Klassifizierung durch Ausnutzung des Desorientierungseffekts durchgeführt. Hinweise auf mögliche Strategien, die sich auf Überblickswissen berufen, konnten erbracht werden. Dazu zählten die Art der Verteilung der Fehler und systematische Verschiebungen in den Verteilungen der Alters- und Geschlechtsgruppen. Es wurde der von Wang und Spelke (2002) und Waller und Hodgson (2006) beschriebene allgemeine Desorientierungseffekt in Bedingungen mit 135°-Turns zur Differenzierung der Strategien in der Zeigeaufgabe eingesetzt. Die „Identifikation“ erfolgte durch eine künstliche Kategorisierung anhand der Schätzfehler in drei gleich große Gruppen.

Für Versuchspersonen der ersten Gruppe, die mit ihren absoluten Schätzfehlern im ersten Drittel der Stichprobe lagen, wurde vermutet, dass sie neben Pfadintegrationsprozessen auch Überblickswissen zur Bestimmung der Heimkehrrichtung nutzten. Diese Personen konnten dadurch offensichtlich den Desorientierungseffekt, der bei Drehungen von 135° auftreten sollte, kompensieren und erzielten geringere Schätzfehler in diesen Versuchsbedingungen als alle anderen Versuchspersonen

Für Personen, die mit ihren Schätzfehlern im zweiten und dritten Drittel der Stichprobe lagen, wurde angenommen, dass sie durch die 135°-Turns die Orientierung in der virtuellen Umgebung verloren. Die mittlere Gruppe versuchte vermutlich durch Pfadintegration die Richtung zu bestimmen. Deren Schätzfehler verteilten sich gleichmäßig um einen Mittelwert von 95° und wiesen somit den typischen Desorientierungseffekt auf. Für Personen der dritten Gruppe wurde wiederum eine Systematik der Schätzfehler mit einer Häufung bei einem bestimmten Wert festgestellt. Daraus wurde abgeleitet, dass diese Personen die Richtung eher geraten hatten.

In Bedingungen mit ausschließlich 90°-Turns unterschieden sich die Gruppen mit dem orientierten und dem Rateverhalten nicht. Personen, die desorientiert waren, wiesen allerdings schon in diesen Bedingungen im Mittel größere Schätzfehler auf. Gruppenspezifische Korrelationen verschiedener Testbedingungen waren gering, was die von Waller und Hodgson (2006) geäußerte Vermutung auf spontane Wechsel zwischen egozentrischen und allozentrischen Repräsentationen stützt. Eine Spezifizierung von Lösungsstrategien in einer Zeigeaufgabe ist also prinzipiell durch Ausnutzung des Desorientierungseffekt möglich.

Auch die ungleichmäßige Verteilung von Alters- und Geschlechtsgruppen auf die gewonnene Klassifizierung lässt einen systematischen Effekt vermuten. So wurden weibliche und ältere Versuchspersonen öfter der Gruppe zugewiesen, die den Desorientierungseffekt zeigte. Männliche Versuchspersonen waren öfter in der Gruppe mit Rateverhalten. Inhaltlich bedeutet dies, dass auch aus der ausschließlich visuellen Bewegungswahrnehmung zwei mentale Repräsentationen des Raumes aufgebaut werden. Die eine ist eine transiente, relativ genaue und implizite Repräsentation, die online aufgebaut wird; die andere eine dauerhafte, weniger genaue und offline konstruierte (Waller & Hodgson, 2006). Unklar bleibt, ob diese Unterschiede im Aufbau oder dem Abruf der Repräsentationen bestehen. Weiterhin kann keine zuverlässige Aussage darüber getroffen werden, welche

Referenzsysteme (egozentrische oder allozentrische) tatsächlich benutzt wurden. Beobachtbar bleiben nur die unterschiedlichen Verhaltensantworten, die hier als die Anwendung unterschiedlicher Strategien interpretiert wurden.

Als Fazit zur Methode der Kompass-Zeigeaufgabe kann festgestellt werden: Es handelt sich bei der Zeigeaufgabe um einen schnell durchführbaren Test mit effizienten Einsatzmöglichkeiten, der PI-Leistung in virtuellen Umgebungen erfasst. Dieser stellt Anforderungen, die auch im Alltag erforderlich sind.

Zur Beantwortung der Eingangs gestellten Frage, ob Orientierungsleistungen, die auf visueller PI beruhen, durch Zeigeaufgaben in virtuellen Umgebungen operationalisiert werden können, ist zu sagen, dass dies nicht uneingeschränkt möglich war. Als Ursache sind hier hauptsächlich methodische Schwächen der Implementierung der Kompass-Zeigeaufgabe zu nennen. In der vorliegenden Implementierung streut das hauptsächlich gemessene Leistungsmerkmal beträchtlich. Eine mögliche Verbesserung der Versuchsanordnung wäre die Einbeziehung mehrerer Bedingungen mit Pfad 1-ähnlicher Komplexität. Diese Pfade stellen offensichtlich Items mittlerer Schwierigkeit dar, deren stärkere Anwendung die Messgenauigkeit der Zeigeaufgabe auf das Konstrukt Pfadintegration deutlich erhöhen würde.

Weiterhin ist zur Kontrolle der Übungs- und Erinnerungseffekte eine Randomisierung der Reihenfolge und die Vermeidung gleicher Pfade notwendig. In Anbetracht von Altersdifferenzen und der invaliden Messung der Reaktionszeit sollte auch die Bedienung des Instruments angepasst werden. Die Durchführung von Probedurchgängen sollte unter diesem Gesichtspunkt, entgegen der Empfehlung von Bowman et al. (2001), in psychometrischen Tests in Erwägung gezogen werden.

Zur Verwendung des Desorientierungseffekts als Klassifikationskriterium stellt die hier verwendete Methode sicher nur einen ersten Versuch dar. Einerseits geht durch die Methode der künstlichen Kategorisierung Information verloren, andererseits konnten so *post hoc* Informationen aus dem ansonsten gleichverteilten Merkmal der Schätzfehler für Pfade mit 135°-Turns gewonnen werden. Die Ergebnisse dieser Strategieanalyse sollten aber in einer weiteren Untersuchung wenigstens mittels Fragebogen validiert werden.

12.3 Experiment 3 - Place-Learning

Im Experiment 3 wurde der Frage nachgegangen, ob Zusammenhänge zwischen den räumlichen Fähigkeiten Place-Learning, Pfadintegration und mentale Rotation bestehen, und sich diese auch in qualitativen Zusammenhängen für die räumlichen Präferenzen nachweisen lassen. Dazu wurden die räumlichen Fähigkeiten mentale Rotation durch den *Mental Rotations Test* (MRT, Peters et al., 1995), Pfadintegration durch eine Kompass-Zeigeaufgabe und Place-Learning durch eine Morris-Water-Maze-ähnliche Navigationsaufgabe operationalisiert. Räumliche Präferenzen wurden aus den Klassifikationen von Lösungs- bzw. Suchstrategien abgeleitet. Die Untersuchung von Alters- und Geschlechtsunterschieden diente auch in diesem Experiment der Überprüfung der Re-

liabilität und Validität der selbstentwickelten Water-Maze-Aufgabe und der Strategieanalyseverfahren.

Zunächst konnten die erwarteten quantitativen Altersdifferenzen für die Lernphase und den Testdurchgang 1 der Water-Maze-Aufgabe bestätigt werden. Die Ergebnisse belegen den Vorteil jüngerer Erwachsener (Alter 20 bis 29 Jahre) gegenüber älteren Erwachsenen (Alter 50 bis 59 Jahre) in dieser raumkognitiven Aufgabe und sind konsistent mit anderen Arbeiten in diesem Bereich (Driscoll et al., 2005; Moffat & Resnick, 2002; Newman & Kasznaik, 2000). Für die Anzahl erfolgreicher Lerndurchgänge war die Varianzaufklärung durch die Altersgruppenzugehörigkeit vergleichbar mit dem Alterseffekt im Test zur räumlichen Wahrnehmung/Feldabhängigkeit. Alterseffekte für die initialen Richtungsschätzfehler (Heading-Error) und die relativen Umwege, Maße welche die Genauigkeit der mentalen Repräsentation für die Ziellokation bzw. die Effizienz der Navigation erfassen, lagen in der gleichen Größenordnung vor wie für den Test zum visuell-räumlichen Gedächtnis. Alterseffekte in den relativen Verweildauern aus dem Testdurchgang 2 bestanden ebenfalls. Dieses Maß erfasst auch die Genauigkeit der mentalen Repräsentation des Ziels, kann also als direkte Abbildung der Place-Learning-Leistung interpretiert werden. Im Testdurchgang 1, in dem eine Orientierung anhand der Landmarken, auf eine eher nicht-räumliche Art erfolgen konnte, bestanden Altersunterschiede in diesem Leistungsmaß nicht. Das passt zu den in Laurence et al. (2002) publizierten Place-Learning-Defiziten in virtuellen Umgebungen für ältere Versuchspersonen. Offensichtlich können bei dieser Untersuchungsmethode ältere Versuchspersonen Relationen innerhalb der virtuellen Umgebung in Bezug auf Landmarken sehr gut repräsentieren und zur erfolgreichen Orientierung verwenden. Sie sind aber nicht in der Lage, eine räumliche Vorstellung abzurufen, welche die genaue Position der versteckten Ziellokation repräsentiert.

Geschlechtseffekte in der Water-Maze-Aufgabe bestanden, fielen allerdings niedriger aus als erwartet. Eine Effektstärke von $d=0.62$ für die Anzahl der erfolgreichen Lerndurchgänge lag unter den Effektstärken für die in Astur et al. (2004) verwendeten Leistungsmaße der Water-Maze-Aufgabe. Das könnte zum einen an der anderen Operationalisierung des Place-Learnings gelegen haben und zum anderen an der Zusammensetzung der Stichprobe, denn auch im MRT fiel der Geschlechtseffekt mit $d=0.48$ relativ gering aus.

Für die Beurteilung qualitativer Unterschiede in der Bearbeitung der Water-Maze-Aufgabe lieferten Musteranalysen Basisfunktionen, die quasi als Prototypen des Suchverhaltens der Personen innerhalb einer Klasse (Subgruppe) interpretiert wurden. Aus diesen Analysen wurde deutlich, dass sich Versuchspersonen in jeder Phase der Aufgabe in der Anwendung ihrer Lösungsstrategien unterschieden. Die Strategien selbst unterschieden sich in ihrer Effizienz und waren somit für quantitative Leistungsunterschiede mit verantwortlich.

Bereits für den Explorationsdurchgang konnte auf ein differierendes Verhalten der Versuchspersonen auf Grund ihrer Bewegungsmuster geschlussfolgert werden. Die Suchstrategien dieser Phase unterschieden sich prinzipiell darin, ob Personen die proximalen Hinweisreize zur Erkundung der Insel verwendeten oder sich von ihnen lösten. Insbesondere ältere Versuchspersonen wurden mit diesen Strategien identifiziert. Jüngere Versuchspersonen konnten sich dagegen öfter von den proxi-

malen Hinweisen lösen und die Insel durch eine an Thigmotaxis angelehnte Suchstrategie erkunden. Interessant dabei ist, dass auch der Inselrand einen proximalen Cue darstellt. Nur wenige Personen konnten sich, durch eine dritte Art Strategie, auch von diesem lokalen Hinweis lösen. Die vorwiegend männlichen Versuchspersonen dieser Gruppe zeigten in den Spurdaten Bewegungsmuster, die aus dem Durchsuchen des gesamten Inselterritoriums resultierten.

Konsistent zu den Musteranalysen des Explorationsverhaltens gaben Personen, die proximale Hinweisreize benutzten, in Fragebögen auch häufiger an, diese bewusst genutzt zu haben. Personen, die sich mit Durchsuchung in der Exploration fortbewegten, berichteten in den Fragebögen zu den Suchstrategien, dass für sie die zusätzliche Zeit zum Umschauen nicht mehr erforderlich war. Diese Unterschiede im ersten Durchgang der Water-Maze-Aufgabe wurden auch in Moffat, Kennedy, Rodrigue und Raz (2007) allerdings für die Pfadlängen berichtet. Bemerkenswert dabei ist, dass Altersdifferenzen zu diesem Zeitpunkt der Aufgabe nicht auf Unterschiede im räumlichen Gedächtnis zurückzuführen sind, sondern durch das initiale Verhalten der einzelnen Versuchsperson bestimmt sind. Moffat et al. (2007) argumentierten, auf der Basis von Tiermodellen nach Barnes (1979), dass ältere Versuchspersonen zu Beginn eher zu egozentrischen Response-Strategien und jüngere eher zu allozentrischen Ortsstrategien neigen.

Hier konnte nun gezeigt werden, dass jüngere und ältere Erwachsene sich zumindest tendenziell, bei den zur Exploration bevorzugten Hinweisreizen, und somit in ihren initialen Suchstrategien unterscheiden. Das von Personen beider Altersgruppen gezeigte Orientierungsverhalten beruhte jedoch auf egozentrischen Strategien. Nur zwischen den Geschlechtergruppen bestanden Unterschiede in den Häufigkeiten für Response- und Ortsstrategien. Die Interaktion beider Faktoren ist mit den hier durchgeführten Analysemethoden nicht überprüft worden, so dass die Faktoren nicht getrennt interpretiert werden können. Außerdem wussten die Versuchspersonen durch die Testinstruktion, dass die proximalen Hinweise (blau-weiße Bälle) zur Ziellokation führen. Die Entscheidung der Versuchspersonen für eine bestimmte Strategie war also in dieser Richtung beeinflusst.

Aus den Bewegungsmustern der Lerndurchgänge wurden insgesamt zwei Gruppe von Personen identifiziert, die sich in ihren Bewegungsmustern komplett unterschieden. Die eine Gruppe bewegte sich in jeder der Bedingungen direkt auf das Ziel zu. Diese Versuchspersonen konnten offensichtlich, auch mit wechselnden Startbedingungen, sicher die Ziellokation finden. Sie zeigten also auch bei veränderten Bedingungen ein gleichbleibendes Niveau. Diese Stabilität im Verhalten über die Lerndurchgänge hinweg wurde als das unter dem Begriff unimodale räumliche Präferenz eingeführte Konstrukt interpretiert. Personen der zweiten Gruppe zeigten in ihren Bewegungsmustern das bimodale Verhalten. Sie steuerten in jedem der Durchgänge nur indirekt die Ziellokation an. Personen dieser Gruppe fielen auch durch eine geringe Anzahl erfolgreicher Lerndurchgänge auf, so dass auf ein eher unsicheres wechselndes Suchverhalten geschlussfolgert wurde. In der Verteilung jüngerer und älterer Versuchspersonen auf diese beiden Strategiegruppen bestand auch der erwartete, wenn auch hier nur nominale, Zusammenhang zwischen Suchverhalten in der Lernphase und dem Faktor „Alter“. Entsprechend der Hypothese von Burke und Barnes (2006) über uni- und bimodales Suchverhalten wurden ältere Probanden öfter der Kategorie indirekter, unsicherer Strategien zugeordnet.

Laut eigener Angaben verwendeten alle Personen irgendeine Art Strategie während der Lernphase. Personen, die mit direkten sicheren Suchstrategien identifiziert wurden, verwendeten am häufigsten die Landmarkenkonstellation Segelboot-Leuchtturm. Um das Ziel genau ansteuern zu können, führten diese Personen möglicherweise eine Art Triangulation zur Kursbestimmung und -korrektur durch. In deren Berechnungen müssten die räumlichen Relationen der zwei Landmarken zueinander und zur eigenen Position einbezogen worden sein.

Suchmuster aus dem Testdurchgang 1 demonstrierten drei verschiedene Arten von Suchstrategien. Es gab eine sehr große Gruppe, die trotz der fehlenden bedeutsamen Landmarke „Segelboot“ eine direkte Suchstrategie zum Ziel einsetzen konnte, d. h. Richtung und Position waren richtig encodiert und konnten auch ohne zweite Landmarke abgerufen werden. Auch Positionskorrekturen während der Navigation konnten relativ fehlerfrei vorgenommen werden. Dadurch sind aber in diesem Test die direkten Suchstrategien sowohl als ortsbezogene Place-Learning-Strategie als auch als reizbezogene Response-Strategie zu interpretieren, d. h. egozentrische und allozentrische Strategien können von Personen dieser Strategiekategorie gleichermaßen angewandt worden sein. Indirekte Suchstrategien in Testbedingung 1 lassen wiederum auf eine fehlende Encodierung der räumlichen Position des Ziels schließen. Personen versuchten hier zunächst den Leuchtturm, als mit dem Ziel assoziierte Landmarke, anzusteuern und anschließend mit beliebigen Suchbewegungen in der Nähe des Ziels diese zu erreichen. Place-Learning wurde von diesen Versuchspersonen im eigentlichen Sinne nicht durchgeführt. Vorwiegend für ältere Versuchspersonen wurden diese Suchstrategien nachgewiesen.

Am verblüffendsten im Testdurchgang 1 war eine dritte Strategie, die sich als Durchsuchen an der Position der gelöschten Landmarke Segelboot am besten beschreiben lässt. Versuchspersonen hatten offensichtlich ihr räumliches Bezugssystem in der virtuellen Insel auf diese Landmarke aufgebaut. Nach Löschung wurde nur noch die ungefähre Position dieser wiedererinnert. Die Position des Ziels war nicht abgespeichert worden. Dies musste durch systematisches Suchen kompensiert werden. So waren nur 50 % der Personen, die mit dieser ineffizienten Strategie identifiziert wurden, im Test erfolgreich. Der Leuchtturm spielte in dieser Suchstrategie als Hinweisreiz keine oder nur eine untergeordnete Rolle. Dies deckte sich mit den Angaben, die die Versuchspersonen in den Fragebögen zu den in diesem Testdurchgang verwendeten Landmarken gaben. Die Personen, die diese Suchstrategie verwendeten, konnten sich am seltensten daran erinnern, den Leuchtturm bewusst zur Orientierung genutzt zu haben. In dieser Strategieggruppe überwog wiederum die Anzahl ältere Versuchspersonen die der jüngeren.

Die vielfältigen Möglichkeiten, die Testbedingung 2 zur Verhaltensgenerierung bot, spiegelte sich auch in einer besonders hohen Zahl identifizierter Strategien wieder. Es wurden vier sich unterscheidende Suchstrategien ermittelt. Zwei davon waren die eher erfolgreichen. Die erste davon, die indirekte Suchstrategie, wurde teilweise als Abruf einer erlernten Bewegungssequenz und teilweise als Abruf einer visuellen Vorstellung der räumlichen Lage der Ziellokation in Bezug auf die Startposition interpretiert. Der Navigationsvorgang bei dieser Art Strategie gliederte sich sehr wahrscheinlich in eine Sequenz von zwei Handlungsplänen: (1) Einschlagen einer indirekten Richtung zum Ziel und (2) die sukzessive Korrektur während der Bewegung. Statistisch war diese die erfolgreichste Strategie. Die zweiterfolgreichste Strategie unterschied sich doch recht deutlich von

diesem Bewegungsmuster. Die Versuchspersonen schlugen hier eine direkte Richtung zum Ziel ein und suchten, wie in einem schmalen Korridor, links und rechts von diesem einmal eingeschlagenen Weg. Die Erfolgchancen dieser Strategie betragen 50:50. Personen dieser Strategiekategorie hatten offensichtlich die Richtung zum Ziel richtig encodiert und entsprechend der Testinstruktion gehandelt. Diese Art der Strategie, d. h. das Einschlagen einer direkten Zielrichtung mit anschließendem systematischem Durchsuchen in einem kleineren Zielbereich wurde auch im VR-Water-Maze von Kallaia et al. (2005) als Suchstrategie „Flächendeckendes Durchsuchen“ (*Enfilading*) identifiziert. Am häufigsten wurde diese Suchstrategie von jüngeren Personen eingesetzt.

Die nicht erfolgreichen Strategien im Testdurchgang 2 beruhten vermutlich auf einer fehlerhaften Pfadintegrationsstrategie oder auf Thigmotaxis. Letztere wurde am häufigsten von älteren Versuchspersonen eingesetzt. Diese Strategien stellen in dieser sehr schwierigen Testbedingung die wenigen Möglichkeiten zur Verhaltensgenerierung durch visuelles Suchen bzw. die Rettung auf die verbliebenen lokalen Hinweisreize (Inselrand) dar.

In einem weiteren Schritt der Strategieanalysen wurden auch in diesem Experiment die Lösungs- und Bearbeitungsstrategien im MRT sowie in der Kompass-Zeigeaufgabe bestimmt. Gemeinsam mit der Klassifizierung der Suchstrategien aus der Water-Maze-Aufgabe wurden diese Ergebnisse genutzt, um der Frage nach einer aufgabenübergreifenden Strategieanwendung in der Raumkognition nachzugehen.

Aus der Kompass-Zeigeaufgabe wurden, wie bereits in Experiment 2 beschrieben, drei Subgruppen von Personen abgeleitet (s. Abschn. 10.2.3). Die erste identifizierte Personengruppe zeichnete sich dadurch aus, dass sie neben PI-Prozessen auch Überblickswissen in dieser Aufgabe nutzte (Gruppe „Orientiert“). Eine zweite Gruppe zeigte in dieser Aufgabe den typischen Desorientierungseffekt (Gruppe „Desorientiert“). Auf Grund ihres Antwortmusters wurde bei Personen der dritten Gruppe davon ausgegangen, dass sie die Richtung geraten hatten (Gruppe „Geraten“). Im Experiment 3 wurde nun auch durch Fragebögen die Validität der so ermittelten Klassifizierung überprüft. Es ergaben sich konsistente und inkonsistente Ergebnisse zwischen Selbst-Report und der durchgeführten Klassifizierung. Konsistent waren folgende Angaben: Personen, die der Gruppe mit Desorientierung zugewiesen wurden, gaben seltener an, sich den Park aus der Vogelperspektive vorgestellt zu haben und berichteten häufiger, dass ihnen die schnelleren Durchgänge schwerer fielen als die langsameren. Das lieferte ein Indiz dafür, dass Personen mit Schätzfehlern um den Ratewert von 90° keine zusätzliche Repräsentation der Umgebung verwendeten, sondern über eine berechnungsintensive Methode die Richtung zum Ausgangspunkt korrigierten. Personen der Klasse „Orientiert“ überlegten öfter, wie sich die Richtung zum Startpunkt änderte. Laut eigener Angaben führen diese Personen Pfadintegrationsprozesse aus. Dazu bevorzugten sie die Verrechnung von Richtungen. Ein Widerspruch bestand in den Angaben der Personen aus der Klasse „Geraten“, die sich laut eigener Angaben nicht auf ihr Gefühl verließen. Offensichtlich waren auch die sehr großen Schätzfehler eher durch überlegte Handlungen zustande gekommen. So lässt sich nur vermuten, dass auch diese Personen die typischen Auswirkungen von Desorientierung zeigten. Es ist also festzuhalten, dass die Interpretation der dritten Gruppe vermutlich nicht zutreffend ist. Für die weiteren Analysen spielte allerdings nur die Zuweisung zur Gruppe „Orientiert“ eine Rolle.

Aus dem MRT wurden Strategieklassen ermittelt, die später zur Ableitung der räumlichen Präferenz holistisch vs. analytisch verwendet wurden. Es konnten auch in der relativ kleinen Stichprobe vier Klasse mit unterschiedlichem Lösungsverhalten identifiziert werden. Die meisten Personen gehörten einer Klasse mit vorzugsweise analytischen Strategien an, da sie durch das gezielte Lösen von Typ-II-Items auffielen. Den höchsten Summenscore erzielte eine Gruppe von Personen, die offensichtlich holistische Strategien einsetzten, da sie sowohl für Typ-I-Items als auch für Items mit Okklusion im Mittel die höchsten Lösungshäufigkeiten aufwiesen. Ein ähnliches Lösungsprofil, jedoch in schwächerer Ausprägung, wiesen Personen einer weiteren Klasse auf, die somit ebenfalls zu den eher holistisch aber langsamer arbeitenden Personen gezählt wurden. In dieser Stichprobe wurde auch, wie in Geiser (2004) beschrieben, eine Gruppe von Personen identifiziert, die offensichtlich generelle Schwierigkeiten mit dem MRT hatte (Klasse 3). Geschlechtsdifferenzen konnten allerdings nicht auf die Anwendung unterschiedlich effizienter Bearbeitungsstrategien zurückgeführt werden. Bezüglich der Altersgruppen verteilten sich jedoch ältere Personen statistisch signifikant häufiger auf die Klasse 3. Jüngere waren dagegen häufiger in der Klasse mit holistischen Strategien vertreten. Die Zusammensetzung der Stichprobe hinsichtlich der Strategien könnte auch, neben der geringen Größe und der daraus resultierenden geringen Teststärke, eine mögliche Erklärung für die gering ausgeprägten Geschlechtsunterschiede im MRT sein.

Zur Bearbeitung der Fragestellung der aufgabenübergreifenden Strategieanwendung wurden die bisher ermittelten Klassifizierungen in das aus der eigenen theoretischen Sicht entwickelte Konzept der räumlichen Präferenzen transformiert. Aus den MRT-Strategien wurden dazu die Kategorien analytisch und holistisch als die beiden Pole des bipolaren Präferenzkonstrukts gebildet. Die gleiche Aggregation wurde für Strategien aus dem Water-Maze und der Zeigeaufgabe durchgeführt. Es wurde je eine Kategorie für eine eher egozentrische vs. allozentrische Präferenz gebildet. Die Zusammenhänge auf diesem kategorialen Niveau waren insgesamt eher schwach. Lediglich für analytisch arbeitende Versuchspersonen wurden Zusammenhänge zum Explorationsverhalten in der Water-Maze-Aufgabe ermittelt. Diese Personen setzten auch im Explorationsdurchgang und im Testdurchgang 2 eher auf lokale Hinweisreize, indem sie die proximale Strategie bzw. Thigmotaxis anwendeten. Es sind aber auch die Effekte der stärkeren Häufigkeit älterer Versuchspersonen in diesen Kategorien zu berücksichtigen. Zusammenhänge zwischen den Präferenzen analytisch vs. holistisch und egozentrisch vs. allozentrisch konnten nicht bestätigt werden.

Eine mögliche Interpretation dieses Befundes wäre die tatsächliche Unabhängigkeit beider Verarbeitungsprozesse. So ist die Präferenz egozentrisch vs. allozentrisch auf das Referenzsystem zur Repräsentation räumlicher Information bezogen. Die Präferenz holistisch vs. analytisch umfasst dagegen mentale Aggregationsmethoden für räumliche Informationen. Dagegen sprechen jedoch die Korrelationen auf dem Niveau einzelner Strategien, da hier bereits entsprechende empirische Zusammenhänge zwischen Bearbeitungsstrategien bei Raumvorstellungstests und der Nutzung von Hinweisreizen aus der virtuellen Umgebung des Water-Maze nachgewiesen wurden. Nicht vorhandene Zusammenhänge sind vermutlich eher auf methodische Probleme durch die *at hoc* ermittelten Kategorien zur Operationalisierung der Präferenzkonstrukte zurückzuführen.

Die abschließenden Punkte zur Diskussion beziehen sich auf die Validität der eingesetzten Verfahren. Annahmen zur Dissoziation der Water-Maze-Aufgabe zu anderen Raumvorstellungstests konn-

ten nicht ganz bestätigt werden. Die Aufgabe stellt offensichtlich in jeder Phase unterschiedlich gewichtete mehrdimensionale Anforderungen. Dazu gehören Fähigkeiten des visuell-räumlichen Gedächtnisses, der mentalen Rotation, der feldunabhängigen räumlichen Wahrnehmung und Pfadintegration.

Die oft berichteten Zusammenhänge zwischen MRT-Summscore und der Anzahl der Lerndurchgänge bestand auch hier (Driscoll et al., 2005). Es konnte aber auch gezeigt werden, dass diese nicht nur für den MRT, sondern generell für Papier-Bleistift-Tests zur Raumvorstellung zutreffen. In den meisten anderen Untersuchungen war das Testmaterial nicht genügend auf weitere Papier-Bleistift-Tests ausgeweitet worden. Außerdem lassen sich die meisten Korrelationen zu den verwendeten Raumvorstellungstests durch Alterseffekte aufklären. Die Zeitlimits spielten hier vermutlich eine sehr große Rolle. Ausgehend von geringfügigen Differenzen in den Bearbeitungszeiten des TMT, akkumulierten sich diese zu einem generellen Leistungsdefizit älterer Versuchspersonen in länger andauernden Raumvorstellungstests. Dennoch auch zur Pfadintegrationsleistung, die ja über die Zeigeaufgabe ohne Zeitlimit gemessen wurde, bestanden Zusammenhänge zur Anzahl der Lerndurchgänge. Offensichtlich konvergieren mehrere räumliche Fähigkeiten in der Lernphase des Water-Maze. Das unterstreicht die Multi-Dimensionalität der Aufgabe.

Im Testdurchgang 1 ergaben sich divergente Resultate mit Tests zur mentalen Rotation, Feldunabhängigkeit und den Schätzfehlern der Zeigeaufgabe. Es bestanden außerdem nur geringe Altersdifferenzen in dem Maß, so dass die Validität des Maßes und der Testbedingung für die Operationalisierung von Place-Learning-Fähigkeiten als gering anzusehen ist. Eine geringe positive Korrelation mit der Anzahl der erfolgreichen Lerndurchgänge spricht ebenfalls für eine geringe interne Konsistenz der Testbedingung 1 zur gesamten Water-Maze-Aufgabe. Durch diesen Test und den verwendeten Leistungsparameter sind offensichtlich die räumlich-mnestischen Fähigkeiten des Place-Learning nicht genügend operationalisiert worden.

Unerwartet zeigten sich die Ergebnisse im Testdurchgang 2, der ursprünglich als Pfadintegrationsbedingung in der Water-Maze-Aufgabe implementiert wurde. Das Leistungsmaß aus dem Testdurchgang korreliert nur sehr schwach mit den Schätzfehlern der Zeigeaufgabe, dafür aber hoch mit dem Papier-Bleistift-Test zum visuell-räumlichen Gedächtnis. Demzufolge scheint die operationalisierte Fähigkeit im VWM-Test konvergent zu dieser Fähigkeit und divergent zum Konstrukt Pfadintegration zu sein. Als Schlussfolgerung aus diesen beiden Korrelationen muss gelten, dass die virtuelle Umgebung und der Ort des Ziels eher in seiner visuellen Gesamtheit abgespeichert werden, als durch Richtungsvektoren und Distanzen. Die Leistung im Testdurchgang 2 korreliert auch nach Bereinigung von Alterseffekten mit dem Test zum visuell-räumlichen Gedächtnis. Die Gemeinsamkeit beider Verfahren besteht im spontanen Abrufen visuell-räumlicher Informationen aus dem Langzeitgedächtnis. Dies beinhaltet die Fähigkeit zum freien Wiedererinnern einer zuvor gesehenen räumlichen Anordnung und scheint sowohl für zwei-dimensionales Material wie für virtuelle drei-dimensionale Umgebungen zu gelten.

Die Konsistenz der gesamten Water-Maze-Aufgabe über die verschiedenen Bedingungen und Leistungsmaße hinweg scheint dagegen eher gering zu sein. Es bestehen nur geringe Korrelationen zwischen der Anzahl der erfolgreichen Lerndurchgänge und den relativen Verweildauern. Beide

Testbedingungen sind sogar in ihren Anforderungen statistisch unabhängig von einander. Es bestanden keine bedeutsamen und kaum von Null verschiedene Korrelationen.

Plausible Zusammenhänge zwischen den einzelnen Phasen der Water-Maze-Aufgabe bestehen dagegen deutlicher auf der qualitativen Ebenen der Suchstrategien. Verhalten in den Lerndurchgängen und Verhalten im Testdurchgang 1 waren durch die konsistente Anwendung von direkten bzw. indirekten Suchstrategien gekennzeichnet. Personen mit indirekten Suchstrategien erwiesen sich in ihrer Performanz auch in der Gruppierung aus der Zeigeaufgabe als konsistent. Diese wurden häufiger als andere Personen der Gruppe „Desorientiert“ zugeordnet, woraus geschlussfolgert wird, dass diese Personen insbesondere beim Erwerb von allozentrischen und dauerhaften mentalen Repräsentationen der Umgebung Defizite aufwiesen. Hier schien wiederum der Altersunterschied die entscheidende Rolle zu spielen. In der Verteilung der Personen auf beide Suchstrategien zeigte sich dieser Alterseffekt. Explorationsverhalten und Suchstrategie im Testdurchgang 2 wiesen ebenfalls Parallelen auf. Die Explorationsstrategien „Durchsuchen“ und „Thigmotaxis“ wurden im Test am häufigsten zur Strategie „Pfadintegration mit Fehlern“. Personen, die in der Exploration den proximalen Cues folgten, nutzten im Test den Rand der Insel durch Thigmotaxis. Auch hier bestand ein deutlicher Alterseffekt in den Kategorien. Die Zusammenhänge zwischen den Suchstrategien beider Durchgänge lagen in der Ähnlichkeit der Situationen begründet. Ohne räumliche Hinweise, auf die die Versuchspersonen die ganze Zeit achteten, ist der Testdurchgang 2 wiederum die Exploration einer unbekanntem Umgebung.

Bei der Einbeziehung der qualitativen Zusammenhänge in die Betrachtungen zur Validität der selbst entwickelten Water-Maze-Aufgabe muss auch auf die Validität der automatischen Musteranalysen zur Identifikation der Suchstrategien eingegangen werden. So sind die hier ermittelten Klassifikationen von Suchstrategien vergleichbar zu den in Astur et al. (2004) und Kallaia et al. (2005) berichteten Ergebnissen. Aufschlussreich dabei ist, dass in jeder dieser Arbeiten unterschiedliche Setups der Water-Maze-Aufgabe und unterschiedliche Analysemethoden zur Anwendung kamen. Trotzdem konnten, auch in der vorliegenden Arbeit, offensichtlich für das Verhalten in einem virtuellen Water-Maze typische invariante Bewegungsmuster identifiziert werden. Darüber hinaus scheint es sich sogar um noch generelle Verhaltensinvariante zu handeln, denn auch in Daten von Mäusen und Ratten aus realen Water-Maze-Versuchsanordnungen sind diese Bewegungsprototypen vorhanden (u. a. Janus, 2004). Die vorliegende Arbeit reproduziert mit dem eingesetzten Musteranalyseverfahren (SNMF und Consensus-Clustering) diese in der Literatur berichteten Ergebnisse und liefert somit erste Hinweise zur Validität der Anwendung dieses Verfahrens auf Spurdaten aus dem Morris-Water-Maze.

Methodische Einschränkungen des Experiments beziehen sich vor allem auf die Kritik an einigen Leistungsmaßen der Water-Maze-Aufgabe und dem Schwierigkeitsgrad einiger Durchgänge. So ist festzuhalten, dass sich bestimmte Startpositionen während der Lerndurchgänge ungünstig auf den Schwierigkeitsgrad auswirkten. Zu beobachten war dies an den Leistungsmaßen relativer Umweg und initialer Heading-Error, die stark zwischen den Durchgängen wechselten. Besonders kritisch ist die Interaktion des Schwierigkeitsgrads einiger Durchgänge mit dem Alter und Geschlecht der Versuchsperson. Der Testdurchgang 2 muss auf Grund viel zu geringer Lösungshäufigkeiten ebenfalls

als zu schwierig eingestuft werden. In einer rein quantitativen Beurteilung der VWM-Leistung stellt dies ein Problem dar.

Testdurchgang 2, im Ansatz als Pfadintegrationsbedingung in der Water-Maze-Aufgabe implementiert, offerierte durch die komplett fehlenden Landmarken sehr viele Freiheitsgrade für die Verhaltensgenerierung. Die Erfolgsraten in dieser Bedingung lagen insgesamt unterhalb der Ratewahrscheinlichkeit, was auf die Schwierigkeit der virtuellen Navigation ohne Landmarken hinweist. Es wurden im Gegensatz zu den anderen Bedingungen keine Alters- oder Geschlechtseffekte beobachtet. Personen fanden in diesem Test offensichtlich eher zufällig das Ziel.

Bezüglich einzelner Leistungsmaße muss festgestellt werden, dass sie für bestimmte Bedingungen (in der Lernphase) nicht valide waren. Diese konnten sogar in der Aggregation nicht verwendet werden. Außerdem scheint insbesondere der initiale Heading-Error generell ein unsicheres Maß zu sein, was durch die großen Streuungen auffiel. Dies geht vermutlich hauptsächlich auf die willkürliche Festlegung des Messzeitpunkts (20 Sekunden nach Initiierung der Bewegung) zurück. Bei Berücksichtigung dieser Punkte muss die Water-Maze-Aufgabe hinsichtlich ihrer Konstruktvalidität noch weiter verbessert werden.

13 Gesamtdiskussion

13.1 Diskussion inhaltlicher Ziele

In der vorliegenden Arbeit wurde der Frage nachgegangen, in wieweit Lösungs- und Bearbeitungsstrategien Anwendung finden und in Zusammenhang stehen mit Alters- und Geschlechterdifferenzen bei Aufgaben mit raumkognitivem Kontext.

In insgesamt drei Experimenten wurden für drei Teilbereiche raumkognitiver Fähigkeiten die Anwendung von Lösungsstrategien untersucht. Dazu konnten die hier erstmals in der Verhaltensforschung eingesetzten Verfahren SNMF und Consensus-Clustering (Gao & Church, 2005; Brunet et al., 2004; Monti et al., 2003) als Methoden der automatischen Musteranalyse sowohl zur Identifikation von Lösungs- und Bearbeitungsstrategien in Raumvorstellungstests als auch bei der Identifikation von Suchstrategien aus Bewegungsdaten der Water-Maze-Aufgabe überzeugen. Die analysierten Strategien sind vergleichbar zu den in Referenzarbeiten ermittelten Verhaltensmustern (Astur et al., 2004; Kallaia et al., 2005; Geiser, 2004; Hosenfeld et al., 1997). Über die Beurteilung der Effizienz der beobachteten Lösungsstrategien konnten wesentliche Anteile der Leistungsdifferenzen in räumlichen Tests aufgeklärt werden.

In den ersten beiden Hypothesen der Arbeit wurde die Frage aufgeworfen, ob Alters- und Geschlechtseffekte in Zusammenhang mit den identifizierten Lösungs- bzw. Suchstrategien stehen, und somit die organismischen Faktoren lediglich vermittelnden Charakter für die Anwendung gruppenspezifischer Strategien haben. Aus dem Experiment 1 konnte dazu folgende Erkenntnis beigetragen werden: In der Altersgruppe der 12 bis 15-jährigen wenden Jungen in der Raumvorstellungsfacetten mentale Rotation eher holistische (ganzheitliche) Bearbeitungsstrategien an als Mädchen. Der Geschlechtseffekt ist beständig, entwickelt sich aber wahrscheinlich mit dem Zeitpunkt stärkerer hormoneller Differenzierung während der Adoleszenz. Aus dem Experiment 2 ergab sich, dass jüngere Erwachsene (Alter: 20-29 Jahre) schneller Überblickswissen in Situationen zur räumlichen Orientierung erwerben als ältere. Die Besonderheit in diesem Experiment war, dass der Aufbau der allozentrischen Repräsentation offensichtlich ausgehend von Routenwissen bzw. Wissen über Richtungen und Distanzen begann, da die virtuelle Umgebung keinerlei Landmarkeninformation vermittelte. Dieser Altersunterschied bestand schon ca. ab dem 40. Lebensjahr. Im Experiment 3 wurde gezeigt, dass ältere Versuchspersonen (Alter: 50-59 Jahre) egozentrische Suchstrategien in virtuellen Umgebungen präferieren. Diese Strategien wurden als egozentrisch interpretiert, da sie auf der Verwendung von lokalen Hinweisen und Reiz-Reaktions-Verhalten beruhen. In der Raumvorstellungsfacetten mentale Rotation basierte das Verhalten älterer Versuchspersonen öfter auf der

Anwendung analytischer als auf der Anwendung holistischer Strategien. Geschlechtsunterschiede wurden im Explorationsverhalten deutlich: Männliche Personen explorieren öfters als weibliche Personen unbekannte Umgebungen weiträumiger und orientieren sich losgelöst von lokalen Hinweisreizen.

Die Faktoren Strategie, Alter und Geschlecht interagieren sowohl in ihrer Wirkung auf die abhängigen Variablen als auch untereinander auf kategorialen Niveau. Eine Trennung der Effekte ist nicht so einfach möglich, wie zunächst vermutet. Die am Ende der Untersuchungen vertretene Auffassung beinhaltet, dass Strategien und Geschlecht in der Raumvorstellung Faktoren sind, die gegenseitig verstärkend bzw. abschwächend auf die abhängigen Variablen wirken, d. h. es handelt sich um Moderatoren der Raumkognitionsleistung. Diese Schlussfolgerung stützt sich hauptsächlich auf die Ergebnisse aus dem Experiment 1. Die wahrscheinlichste Richtung eines kausalen Zusammenwirkens beider Faktoren ergab sich aus dem Experiment 3. Der schwache Geschlechtseffekt im MRT und die Gleichverteilung der Geschlechter auf effiziente und weniger effiziente Strategien kann dahingehend interpretiert werden, dass hier Lösungsstrategien den Geschlechtseffekt moderieren.

Strategien sind aber ebenfalls der Variation durch andere Faktoren ausgesetzt. Das verdeutlichen die beobachteten Entwicklungs- bzw. Alterseffekte. Die Zunahme des Geschlechtseffekts zwischen dem 12. und 15. Lebensjahr geht einher mit einer massiven Zunahme schneller holistischer Strategien bei männlichen Versuchspersonen. Im Erwachsenenalter ab dem 50. Lebensjahr, wie hier untersucht, kommen bereits stärker analytische Strategien zum Einsatz. Dieser wirkte sich jedoch nicht direkt auf einen quantitativen Effekt aus, was darin begründet liegt, dass Altersdifferenzen im MRT generell als geringer anzusehen sind als Geschlechtsdifferenzen (Geiser et al., in Press). Weiterhin sind die, u. a. in Lehmann (2000) und Rönicke (2003), beobachteten Leistungsdifferenzen zwischen verschiedenen Studienfachrichtungen und Berufsgruppen sehr wahrscheinlich über die Anwendung unterschiedlich effizienter Lösungsstrategien beeinflusst.

Für ältere Erwachsene ergab sich eine interessante Parallele in der häufigeren Anwendung analytischer Strategien in der Raumvorstellung und die Verwendung proximaler Hinweisreize in der räumlichen Orientierung. In einer stark generalisierten Interpretation könnte das heißen: Die Fokussierung der Aufmerksamkeit erfolgt bei älteren eher auf Details einer räumlichen Szene. Das bedeutet, es liegt bereits bei Personen ab dem 50. Lebensjahr eine stärker selektierte Verarbeitung der räumlichen Information vor. Das könnte in einer Kompensation der geringeren Verarbeitungsgeschwindigkeiten und in der Einschränkung bei der Speicherung umfangreicher Details begründet sein. Die vergleichbar großen Alterseffekte in diesen Tests belegen diese Möglichkeit. Möglich ist aber auch, dass dabei Erfahrungswerte älterer Versuchspersonen eine Rolle spielen. Die Fokussierung auf bestimmte nützliche Hinweise kann sich auf der Basis vorangegangener Orientierungssituationen begründen, in denen diese Strategien in der realen Umwelt erfolgreich waren. Vermutlich stellt die Präferenz für egozentrische Orientierungshilfen bei älteren Versuchspersonen den Abruf von Algorithmen dar, die sich in mehreren Jahrzehnten als erfolgreiche Handlungsvarianten verfestigt haben.

Folglich lassen sich Leistungsunterschiede zwischen jüngeren und älteren Personen in virtuellen Navigationsaufgaben viel stärker auf Strategieunterschiede zurückführen. Diese wiederum sind

Ausdruck von Alterseffekten in exekutiven Funktionen, wie dass in Moffat et al. (2007) bereits herausgestellt wurde. Somit unterscheiden sich ältere Erwachsene bei diesen Aufgabe sehr stark in Prozessen der Zielsetzung, der Aufmerksamkeit und der Handlungskontrolle. Das ist ein klarer Beleg für die modulierende Wirkung von Strategien auf Aufmerksamkeitsprozesse, wie sie im postulierten Strategiemodell vorgeschlagen wurden.

In zwei weiteren Hypothesen der Arbeit wurden Zusammenhänge zwischen den Strategien in verschiedenen Testsituationen mit raumkognitivem Kontext sowie zwischen der Strategiewahl und der Präferenzen in der Verarbeitung räumlicher Informationen vermutet. Darüber hinaus wurde die Frage aufgeworfen, ob Strategien situationsspezifisch eingesetzt werden oder eher stabile Personeneigenschaften darstellen.

Die Ergebnisse der hierzu durchgeführten kategorialen Zusammenhangsanalysen ergaben, dass es eine analytische vs. holistische Präferenz so nicht gibt. Eine individuelle, oder wie die hier untersuchte gruppenspezifische, Strategiewahl bezieht sich eher auf einen speziellen Test. Die egozentrische vs. allozentrische Präferenz ist dagegen als die Bevorzugung bestimmter Hinweisreize aus einer räumlichen Umgebung interpretiert worden. Lokale Hinweisreize führen dabei zu egozentrischen Raumrepräsentationen. Globale Landmarken, sofern sie bei der Encodierung von Relationen verwendet werden, führen zum Aufbau allozentrischer Bezugssysteme. Zusammenhänge zwischen diesen beiden räumlichen Präferenzen bestanden ebenfalls nicht. Es handelt sich offensichtlich aus der Sicht der Informationsverarbeitung um unterschiedliche Prozesse. Der eine Prozess betrifft die Organisation von Arbeitsgedächtnisinhalten, z. B. durch Chunkbildung und Informationsverdichtung bei holistischen Strategien, der andere bezieht sich auf ein Referenzsystem, das nur mit der eigenen Perspektive in Zusammenhang steht. Oder anders ausgedrückt geht es um die Bewegungsvorstellung einzelner Objekte bzw. die Einordnung der eigenen Person im Raum.

Bei dieser Interpretation darf aber nicht die Schwäche der Operationalisierung des Präferenzkonstrukts unberücksichtigt bleiben. Die Bildung von jeweils zwei Kategorien aus einer Reihe verschiedener Strategien ist willkürlich und höchst interpretativ. Sehr wahrscheinlich sind nicht gefundene Zusammenhänge in dieser Zusammenfassung auch darauf zurückzuführen. Die Frage, ob Präferenzen über verschiedene Testsituationen stabil sind, kann deshalb nicht sicher beantwortet werden. Hinweise einer Interaktion mit dem Alter der Versuchsperson ergeben sich dennoch auf der Basis der Experimente 1 und 3. Die Raumvorstellungsstrategien 12 bis 15-jähriger stellen sich als situationsspezifisch heraus. Junge Erwachsene reagieren in der Water-Maze ebenfalls spezifischer auf verschiedene Testsituationen in der Wahl ihrer Suchstrategie. Für ältere Erwachsene wurde dagegen ein Zusammenhang zwischen der Verwendung detailbezogener Strategien in Raumvorstellungstest und Aufgaben zur räumlichen Orientierung deutlich. Wie bereits festgestellt, scheinen ältere Personen zur Lösung raumkognitiver Aufgaben feste Handlungsmuster zu verfolgen. Für die Beantwortung der aufgeworfenen Frage heißt das: Mit zunehmendem Alter wechselt die Anwendung von Strategien von situationsspezifisch zu einer Personeneigenschaft.

Insofern ordnen sich die vorliegenden Ergebnisse in die Schlussfolgerungen anderer gegenwärtiger Arbeiten ein, die eine kritische Konfundierung der Water-Maze-Aufgabe mit Aufmerksamkeits-

und Exekutivfunktionen vor allem bei der Untersuchung altersbedingter Veränderung sehen (Moffat et al., 2007; Newman & Kasznaik, 2000). Die gleich großen Alterseffekte im Rey-Osterrieth-Figur-Test und dem verwendeten Place-Learning-Test untermauern dies. Der RCF wird teilweise in der neuropsychologischen Diagnostik zu den Tests Exekutiver Funktionen gezählt (Lezak, 1995). Neurobiologische Evidenz für diese Sicht auf die Place-Learning-Aufgabe besteht darüber hinaus, da insbesondere strukturelle anatomische Veränderung im präfrontalen Kortex (PFC) mit normalem Altern in Zusammenhang gebracht werden. Die einzige im Alter stabil nachgewiesene Abnahme der neuronalen Dichte findet im PFC statt (Van Petten, 2004; Buckner, 2004). Dies bildet möglicherweise das neuronale Korrelat für die Abnahme flexibler Strategiegenerierung hin zu einer selektiven, aber auch verstärkten Anwendung von Algorithmen im Alter. Die vorliegende Arbeit liefert einen weiteren empirischen Beleg dafür.

Diese Erkenntnisse sind natürlich auch für Untersuchungen mit der allgemeine Fragestellung zur Art des räumlichen Langzeitgedächtnisses zu berücksichtigen. Jedes, der in der Theorie zur Definition der räumlichen Informationsspeicher angeführten Modelle, könnte richtig sein, da sie jeweils auf eine bestimmte Subgruppe zutreffen können, die selektiv Strategien einsetzen, mit denen Informationen gefiltert und verarbeitet werden. Eine deklarative, nicht-räumliche Speicherung (Squire, 1986) könnte mit der Abhängigkeit der Strategie von Landmarken in Verbindung stehen. Die Encodierung eines räumlichen Kontextes, wie in der Multiple-Trace-Theorie (Moscovitch et al., 2006), ist möglicherweise mit Repräsentationen von Ortsinformationen, die eher auf der Speicherung räumlicher Primitiven beruhen, in Einklang zu bringen. Die Identifikation von Strategien ist in diesem Forschungsbereich also absolut notwendig. Sie dient der Identifikation von in ihrem Verhalten homogenen Subgruppen. Bisher besteht die Möglichkeit, dass die Beurteilung der empirischen Daten zu diesen Modellen nicht auf der Beurteilung homogener Stichproben beruht. Eine homogene Beurteilung der Verhaltensmaße einer heterogenen Stichprobe (mehrere Populationen) ist nicht möglich. Die Identifikation von homogenen Subgruppen wird hier umso mehr benötigt, um kontrollierte Studien durchführen zu können, und dazu sind klare Kriterien zur Differenzierung notwendig. Das Problem dabei bleibt nach wie vor bestehen: Die Strategien in raumkognitiven Aufgaben beeinflussen nicht immer direkt die gemessenen Leistungsmaße. Die Erkennung von Strukturen, Mustern und Profilen im Verhalten ist dazu erforderlich. Dazu sind Ansätze auf allen Ebenen der Datenanalyse notwendig. Die Anwendung von Strategieanalysemethoden, wie sie hier demonstriert wurde, leistet einen Beitrag, um diesen Widerspruch erfolgreich zu lösen.

13.2 Diskussion methodischer Ziele

Im Rahmen der vorliegenden Arbeit wurde flexible Software zur Durchführung von psychometrischen Tests und Verhaltensexperimenten entwickelt. Damit existieren für weitere Untersuchungen computergestützte Verfahren, die auf verschiedene Designs und Fragestellungen leicht angepasst werden können. Die Automatisierung der Erfassung, Aggregation und Analyse reduziert zudem Fehler im Prozess der Verarbeitung der empirischen Daten. Dadurch kann der Grad der statistischen Validität entsprechender Untersuchungen erhöht werden. Während der Implementierung der

Verfahren wurden darüber hinaus Unstimmigkeiten im Testmaterial des MRT-B und der Analyse der Itemstruktur des MRT-A in vorangegangenen Arbeiten entdeckt und korrigiert.

Über die Implementierung eines alternativen Verfahrens zur Musteranalyse in Verhaltensdaten konnten sowohl Itemanalysen von Raumvorstellungstests als auch Trajektorien von Spurdaten aus Navigationsexperimenten durchgeführt werden. Das verwendete Verfahren ist in der Richtung exploratorischer Faktorenanalysen einzuordnen und basiert auf einem intuitiv verständlichem Prinzip von Analysen der Building-Blocks innerhalb eines Datensatzes und einer bekannten Clusterheuristik. Die Modellparameterwahl erfolgte dabei teilweise durch grafische Mittel. Die mit dem Verfahren durchgeführten Strategieanalysen in den Itemantworten waren ein Ansatz zur Lösung der internen Probleme mehrdimensionaler Skalen innerhalb von Raumvorstellungstests. Erstmals wurde das Verfahren zur Identifikation von Suchstrategien in Trajektoriedaten eingesetzt.

Es konnte in der Arbeit gezeigt werden, dass das Analyseverfahren (SNMF und Consensus-Clustering) prinzipiell funktioniert. Jedoch muss auch festgestellt werden, dass es noch nicht konkurrenzfähig ist zu den in diesem Forschungsbereich eingesetzten probabilistischen Modellen (z. B. Latent-Class-Analysen). Es fehlt die statistische Absicherung der Zuverlässigkeit der Klassenzuweisungen und der Zuverlässigkeit für die Existenz der ermittelten Klassen. Kophentischer Korrelationskoeffizient und Consensus-Plots können nur Ausgangspunkt für weitere Verfeinerungen der Methode sein. Etablierte Maße zur Modellselektion, z. B. Informationskriterien, müssen hier mit berücksichtigt werden. Der große Vorteil der Mustererkennung liegt aber auch in der Flexibilität der Anwendung auf nahezu beliebige Daten. Eine Begrenzung auf dichotome Itemmuster ist nicht unbedingt notwendig. Dies könnte insbesondere bei der Analyse der Antwortmuster aus dem MRT in Zukunft stärker ausgenutzt werden.

Virtuelle Umgebungen konnten auch hier ihre Nützlichkeit und Effizienz in der Operationalisierung zur Erfassung räumlicher Fähigkeiten unter Beweis stellen. Jedoch unterliegt die Entwicklung dieser eher experimentellen Versuchsanordnungen genau so den strengen Regeln der Testkonstruktion, wenn sie als psychometrische Verfahren eingesetzt werden sollen. Hier steht der Entwicklungsprozess im Konflikt zwischen diesen messtheoretischen Anforderungen und der Erzeugung einer Präsenz der Versuchsperson in der virtuellen Umgebung. Oft wird zur Erhöhung der Präsenz, der entscheidende Vorteil der Methode, die Kontrollierbarkeit der Wirkung aller experimenteller Faktoren aufgegeben. Die Ergebnisse der eigenen Arbeiten verdeutlichen in diesem Zusammenhang die Bedeutung der angebotenen virtuellen Objekte als Stimuli für die Wahlmöglichkeiten zur Verhaltenssteuerung. Das Anbieten vieler Stimuli zur realistischen Gestaltung der virtuellen Umgebung führt zur Einschränkung in der Kontrollierbarkeit der Strategieanwendung und zur Abschwächung von Reliabilität und Konstruktvalidität der Messverfahren.

Andererseits ist die Erzeugung von Präsenz zweifelsohne immens wichtig, da die Vorstellung virtueller Räume offensichtlich sehr eng an Raumvorstellungsfähigkeiten geknüpft ist. Als Fazit in dieser Hinsicht muss also für den Entwickler VR-gestützter Tests gelten: Es sind so viel wie nötig Objekte in eine virtuelle Umgebung einzubetten, um den Anspruch auf Präsenz des Nutzers zu wahren, aber auch so wenig wie möglich, um die Kontrollierbarkeit des experimentellen Designs zu erhalten.

Literaturverzeichnis

- Albuquerque, A. L. P., Melo, R. & Velho, L. (2003). Connecting the Presence's Factors for Guiding Measurements. In *Proceedings of PRESENCE 2003, 6th Annual International Workshop on Presence*. Aalborg, Denmark.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (1999). *I-S-T 2000: Intelligenz-Struktur-Test 2000*. Göttingen: Hogrefe.
- Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Croz, J. D., Greenbaum, A., Hammarling, S., McKenney, A. & Sorensen, D. (1999). *LAPACK User's Guide*. Philadelphia, USA.
- Astur, R. S., Ortiz, M. L. & Sutherland, R. (1998). A characterization of performance by men and women in a virtual Morris water task: A large and reliable sex difference. *Behavioral Brain Research*, 93, 185–190.
- Astur, R. S., Taylor, L. B., Mamelak, A. N., Philpott, L. & Sutherland, R. (2002). Humans with hippocampus damage display severe spatial memory impairments in a virtual Morris water task. *Behavioral Brain Research*, 132(1), 77–84.
- Astur, R. S., Tropp, J., Sava, S., Constable, R. T. & Markus, E. (2004). Sex differences and correlations in a virtual Morris water task, virtual radial arm maze, and mental rotation. *Behavioral Brain Research*, 151(1-2), 103–115.
- Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (eds), *The psychology of learning and motivation: Advances in research and theory* (vol. 2). New York: Academic Press.
- Avraamides, M. N., Klatzky, R. L., Loomis, J. M. & Golledge, R. (2004). Use of cognitive vs. perceptual heading during imagined locomotion depends on the response mode. *Psychological Science*, 15(6), 403–408.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, 4(11), 417–423.
- Barnes, C. A. (1979). Memory deficits associated with senescence: a neurophysiological and behavioral study in the rat. *Journal of Comparative and Physiological Psychology*, 93, 74–104.

- Barnes, C. A., Suster, M. S., Shen, J. & McNaughton, B. L. (1997). Multistability of cognitive maps in the hippocampus of old rats. *Nature*, 388, 272–275.
- Barratt, E. S. (1953). An Analysis of Verbal Reports of Solving Spatial Problems as an Aid in Defining Spatial Factors. *The Journal of Psychology*, 36, 17–25.
- Bell, S. & Saucier, D. (2004). Relationship among environmental pointing accuracy, mental rotation, sex, and hormones. *Environment and Behavior*, 36(2), 251–265.
- Bühner, M., Ziegler, M., Krumm, S. & Schmidt-Atzert, L. (2006). Ist der I-S-T 2000 R Raschskalierbar? *Diagnostica*, 52(3), 119–130.
- Bird, C. M. & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience*, 9, 182–194.
- Bülthoff, H. H. & van Veen, H. A. (1999). *Vision and Action in Virtual Environments: Modern Psychophysics in Spatial Cognition Research* (Technical Report Nr. 77). Max-Planck-Institut für biologische Kybernetik.
- Boll, T. J. & Reitan, R. M. (1973). Effect of age on the Trail Making Test. *Perceptual and Motor Skills*, 36, 691–694.
- Bortz, J. & Lienert, G. A. (2003). *Kurzgefaßte Statistik für die Klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben* (2. Aufl.). Berlin: Springer Verlag.
- Bowman, D. A., Johnson, D. & Hodges, L. F. (2001). Testbed Evaluation of Virtual Environment Interaction Techniques. *Presence: Teleoperators and Virtual Environments*, 10(1), 75–95.
- Bowman, D. A., Koller, D. & Hodges, L. F. (1998). A Methodology for the Evaluation of Travel Techniques Immersive Virtual Environments. *Virtual Reality*, 3, 120–131.
- Brandt, T., Schautzer, F., Hamilton, D. A., Bruning, R., Markowitsch, H. J., Kalla, R., Darlington, C., Smith, P. & Strupp, M. (2005). Vestibular loss causes hippocampal atrophy and impaired spatial memory in humans. *Brain*, 128, 2732–2741.
- Brody, D. L. & Holtzman, D. M. (2006). Morris water maze search strategy analysis in PDAPP mice before and after experimental traumatic brain injury. *Experimental Neurology*, 197(2), 330–340.
- Brunet, J.-P., Tamayo, P., Golub, T. & Mesirov, J. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 4164–4169.
- Buckner, R. L. (2004). Memory and Executive Function in Aging and AD: Multiple Factors that Cause Decline and Reserve Factors that Compensate. *Neuron*, 44, 195–208.
- Burke, S. N. & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*, 7, 30–40.

- Burkitt, J., Widman, D. & Saucier, D. M. (2007). Evidence for the influence of testosterone in the performance of spatial navigation in a virtual water maze in women but not in men. *Hormones and Behavior*, *51*, 649–654.
- Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience*, *Advance online publication*, 1–8.
- Caviness, V. S., Kennedy, D. N., Richelme, C., Rademacher, J. & Filipek, P. A. (1996). The human brain age 7-11 years: A volumetric analysis based on magnetic resonance images. *Cerebral Cortex*, *6*, 726–736.
- Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale: Lawrence Erlbaum Associates.
- Collins, D. W. & Kimura, D. (1997). A large sex difference on a two-dimensional mental rotation task. *Behavioral Neuroscience*, *111*, 845–849.
- Cooper, L. A. (1976). Individual differences in visual comparison processes. *Perception & Psychophysics*, *19*(5), 433–444.
- D'Hooge, R. & De Deyn, P. P. (2001). Applications of the Morris water maze in the study of learning and memory. *Brain Research Reviews*, *36*, 60–90.
- Donoho, D. & Stodden, V. (2003). When Does Non-Negative Matrix Factorization Give a Correct Decomposition into parts? In *Proceedings of the Seventeenth Annual Conference on Neural Information Processing Systems*, (pp. 8–13). Vancouver and Whistler, British Columbia, Canada.
- Driscoll, I., Hamilton, D. A., Petropoulos, H., Yeo, R. A., Brooks, W. M., Baumgartner, R. N. & Sutherland, R. J. (2003). The Aging Hippocampus: Cognitive, Biochemical and Structural Findings. *Cerebral Cortex*, *13*, 1344–1351.
- Driscoll, I., Hamilton, D. A., Yeo, R. A., Brooks, W. M. & Sutherland, R. J. (2005). Virtual navigation in humans: the impact of age, sex, and hormones on place learning. *Hormones and Behavior*, *47*, 326–335.
- Duda, R. O., Hart, P. & Stork, D. (2000). *Pattern Classification and Scene Analysis*. (second ed. Aufl.). New York: John Wiley & Sons.
- Eichenbaum, H. (2000). Hippocampus: Mapping or memory? *Current Biology*, *10*, 785–787.
- Foreman, N., Stanton-Fraser, D., Wilson, P. N., Duffy, H. & Parnell, R. (2005). Transfer of spatial knowledge to a two-level shopping mall in older people, following virtual exploration. *Environment and Behavior*, *37*(2), 275–292.
- Freksa, C., Bauer, W., Habel, C. & Wender, K. F. (eds). (2000). *Spatial cognition II - Integrating*

- Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications.* Berlin, New York, Heidelberg: Springer-Verlag.
- Fyh, M., Molden, S., Witter, M. P., Moser, E. I. & Moser, M. B. (2004). Spatial representation in the entorhinal cortex. *Science*, *305*, 1258–1264.
- Gao, Y. & Church, G. (2005). Improving molecular cancer class discovery through sparse non-negative matrix factorization. *Bioinformatics*, *21*(21), 3970–3975.
- Geiser, C. (2004). *Itemanalyse des MRT-A mit Latent-Class-Model - Fragestellung nach Strategien und Geschlechtsdifferenzen*. Unveröffentlichte Diplomarbeit, Institut für Psychologie I, Otto-von-Guericke-Universität Magdeburg.
- Geiser, C., Lehmann, W. & Eid, M. (2006). Separating 'Rotators' from 'Non-Rotators' in the Mental Rotations Test: A multigroup latent class analysis. *Multivariate Behavioral Research*, *41*, 261–293.
- Geiser, C., Lehmann, W. & Eid, M. (in Press). A note on sex differences in mental rotation in different age groups. *Intelligence*.
- Glück, J., Machat, R., Jirasko, M. & Rollett, B. (2001). Training-related changes in solution strategy in a spatial test: An application of item response models. *Learning and Individual Differences*, *13*, 1–22.
- Golledge, R. G., Ruggles, A. J., Pellegrino, J. W. & Gale, N. D. (1993). Integrating route knowledge in an unfamiliar neighborhood: Along and across route experiments. *Journal of Environmental Psychology*, *13*, 293–307.
- Grön, G., Wunderlich, A. P., Spitze, M., Tomczak, R. & Riepe, M. (2000). Brain Activation during human navigation: gender-different neural networks as substrate of performance. *Natural Neuroscience*, *3*(4), 404–408.
- Halpern, D. (2000). *Sex differences in cognitive abilities*. (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Hamilton, D. A., Driscoll, I. & Sutherland, R. J. (2002). Human place learning in a virtual Morris water task: some important constraints on the flexibility of place navigation. *Behavioral Brain Research*, *129*, 159–170.
- Hausmann, M., Slabbekoorn, D., Van Goozen, S. H. M., Cohen-Kettenis, P. T. & Güntürkün, O. (2000). Sex Hormones Affect Spatial Abilities During the Menstrual Cycle. *Behavioral Neuroscience*, *114*(6), 1245–1250.
- Hegarty, M. & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*, 175–191.
- Held, R. & Durlach, N. (1992). Telepresence. *Presence*, *7*, 129–143.

- Herbelin, B., Vexo, F. & Thalmann, D. (2002). Sense of Presence in Virtual Reality Exposures Therapy. In *1st. International Workshop on Virtual Reality Rehabilitation (VRMHR)*. Lausanne, Switzerland.
- Hergovich, A. (1999). Vorstellung und Validierung des Gestaltwahrnehmungstests zur Messung der Felabhängigkeit. *Diagnostica*, 45(6), 20–34.
- Höll, D., Leplow, B., Schönfeld, R. & Mehdorn, M. (2002). Learning of a Real Environment and Learning of a Virtual Space: Is it possible to learn and transfer spatial information? In C. Freska, C. Habel & K. Wender (eds), *Spatial Cognition III - An interdisciplinary approach to representation and processing of spatial knowledge. Lecture Notes of Artificial Intelligence/Computer Sciences*. Berlin: Springer Verlag.
- Horn, W. (1983). *LPS: Leistungsprüfsystem*. Göttingen: Hogrefe.
- Horn, W., Lukesch, H., Mayrhofer, S. & Kormann, A. (2003). *PSB-R 6-13: Prüfungssystem für Schul- und Bildungsberatung für 6. bis 13. Klassen - revidierte Fassung*. Göttingen: Hogrefe.
- Hosenfeld, I., Strauss, B. & Köller, O. (1997). Geschlechtsdifferenzen bei Raumvorstellungsaufgaben - eine Frage der Strategie? *Zeitschrift für Pädagogische Psychologie*, 11(2), 85–94.
- Hoyer, P. (2004). Non-negative matrix factorization with sparseness constraints. *Journal of Machine Learning Research*, 5, 1457–1469.
- Iaria, G., Petrides, M., Dagher, A., Pike, B. & Bohbot, V. (2003). Cognitive Strategies Dependent On The Hippocampus And Caudate Nucleus In Human Navigation: Variability And Change With Practice. *The Journal of Neuroscience*, 23(13), 5945–5952.
- Jansen-Osmann, P. & Fuchs, P. (2006). Wayfinding Behavior and Spatial Knowledge of Adults and Children in a Virtual Environment - The Role of Landmarks. *Experimental Psychology*, 53(3), 171–181.
- Janus, C. (2004). Search strategies used by APP transgenic mice during navigation in the Morris water maze. *Learning & Memory*, 11(3), 337–346.
- Jäger, A. & Althoff, K. (1994). *Der WILDE-Intelligenz-Test*. Göttingen: Hogrefe.
- Jordan, K., Heinze, H.-J., Lutz, K., Kanowski, M. & Jäncke, L. (2001). Cortical Activations during the Mental Rotation of Different Visual Objects. *NeuroImage*, 13, 143–152.
- Jordan, K., Schadow, J., Wüstenberg, T., Heinze, H.-J. & Jäncke, L. (2003). Different cortical activations for subjects using allocentric or egocentric strategies in virtual navigation task. *Neuroreport*, 14(15), 135–140.
- Jordan, K., Wüstenberg, T., Heinze, H.-J., Peters, M. & Jäncke, L. (2002). Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia*, 40, 2397–2408.

- Kallaia, J., Makany, T., Karadia, K. & Jacobs, W. J. (2005). Spatial orientation strategies in Morris-type virtual water task for humans. *Behavioural Brain Research*, *159*, 187–196.
- Keehner, M., Guerin, S. A., Miller, M. B., Turk, D. J. & Hegarty, M. (2006). Modulation of neural activity by angle of rotation during imagined spatial transformations. *NeuroImage*, *33*, 391–398.
- Keidel, W. D. (1970). *Kurzgefaßtes Lehrbuch der Physiologie* (2. Aufl.). Stuttgart: G. Thieme Verlag.
- Kerkhoff, G. (2002). Räumlich-perzeptive, räumlich-kognitive, räumlich-konstruktive und räumlich-topografische Störungen. In W. Sturm, M. Hermann & C. W. Wallesch (Hrsg.), *Lehrbuch der Klinischen Neuropsychologie* (S. 411–429). Lisse (NL): Swets & Zeitlinger Publishers.
- Kerkman, D. D., Wise, J. C. & Harwood, E. A. (2000). Impossible mental rotation problems a mismeasure of women's spatial abilities? *Learning and Individual Differences*, *12*, 253–269.
- Kimura, D. (1999). *Sex and Cognition*. Cambridge: MIT Press.
- Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W. & Fry, P. A. (1990). Acquisition of route and survey knowledge in absence of vision. *Journal of Motor Behavior*, *22*(1), 19–43.
- Köller, O., Rost, J. & Köller, M. (1994). Individuelle Unterschiede beim Lösen von Raumvorstellungsaufgaben aus dem IST- bzw. IST-70-Untertest „Würfelaufgaben“. *Zeitschrift für Psychologie*, *202*(1), 65–85.
- Laurance, H. E., Learmonth, A. E., Nadel, L. & Jacobs, W. J. (2003). Maturation of Spatial Navigation Strategies: Convergent Findings from Computerized Spatial Environments and Self-Report. *Journal of Cognition and Development*, *4*(2), 211–238.
- Laurence, H. E., Thomas, K. G. F., Newman, M. C., Kaszniak, A. W., Nadel, L. & Jacobs, W. J. (2002). Older Adults Map Novel Environments but Do Not Place Learn: Findings From Computerized Spatial Task. *Aging, Neuropsychology, and Cognition*, *9*(2), 85–97.
- Lawton, C. A. & Morrin, K. A. (1999). Gender differences in pointing accuracy in computer-simulated 3D mazes. *Sex Roles*, *40*, 73–92.
- Lee, D. D. & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*, 788–791.
- Lee, D. D. & Seung, H. S. (2001). Algorithms for non-negative matrix factorization. *Advances in Neural Information Processing Systems: Proceedings of the 2000 Conference*, *13*, 556–562.
- Lehmann, W. (2000, Februar). *Group differences in mental rotation*. Magdeburger Arbeiten zur Psychologie, Band 2, Heft 1. Magdeburg: Otto-von-Guericke-Universität, Institut für Psychologie.

- Lehning, M., Haaland, V., Pohl, J. & Lepow, B. (2001). Compass- versus Finger-Pointing Tasks: The Influence of Different Methods of Assessment in Age-Related Orientation Performance. *Journal of Environmental Psychology*, 21, 283–289.
- Lehrl, S., Triebig, G. & Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, 91(5), 335–345.
- Lemay, M., Bertram, C. P. & Stelmach, G. E. (2004). Pointing to an Allocentric and Egocentric Remembered Target in Younger and Older Adults. *Experimental Aging Research*, 30(4), 391–406.
- Lepow, B., Höll, D., Zeng, L. & Mehdorn, M. (2000). Investigation of Age and Sex Effects in Spatial Cognitions as Assessed in a Locomotor Maze and in a 2-D Computer Maze. In C. Freska, C. Habel & K. Wender (eds), *Spatial Cognition II - Integrating Abstract Theorie, Empirical Studies, Formal Methods, and Practical Applications* (vol. 1849, 399–418). Berlin: Springer Verlag.
- Lezak, M. D. (1995). *Neuropsychological Assessment* (3rd ed.). New York: Oxford University Press.
- Linn, M. C. & Petersen, A. C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development*, 56(6), 1479–1498.
- Loomis, J. M., Blascovich, J. J. & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments & Computers*, 31(4), 557–564.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G. & Philbeck, J. W. (1999). Human Navigation by Path Integration. In R. G. Golledge (Hrsg.), *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*. Johns Hopkins University Press.
- Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D. & O'Keefe, J. (1998). Knowing Where and Getting There: A Human Navigation Network. *Science*, 280, 921–924.
- Maier, P. H. (1996). Geschlechtsspezifische Differenzen im räumlichen Vorstellungsvermögen. *Psychologie in Erziehung und Unterricht*, 43(4), 245–265.
- Masters, M. S. & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics*, 23(4), 337–341.
- McNaughton, B. L., Battaglia, F., Jensen, O., Moser, E. I. & Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7, 663–678.
- Messick, S. (1984). The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychologist*, 19, 59–74.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.

- Miller, K. K., Grieco, K. A. & Klibanski, A. (2005). Testosterone Administration in Women with Anorexia Nervosa. *The Journal of Clinical Endocrinology & Metabolism*, 90(3), 1428–1433.
- Moffat, S. D., Kennedy, K. M., Rodrigue, K. M. & Raz, N. (2007). Extrahippocampal Contributions to Age Differences in Human Spatial Navigation. *Cerebral Cortex*, 17, 1274–1282.
- Moffat, S. D. & Resnick, S. M. (2002). Effects of Age on Virtual Environment Place Navigation and Allocentric Cognitive Mapping. *Behavioral Neuroscience*, 116(5), 851–859.
- Montello, D. R. & Pick, H. L. (1993). Integrating knowledge of vertically aligned large-scale spaces. *Environment and Behavior*, 25, 457–484.
- Monti, S., Tamayo, P., Mesirov, J. & Golub, T. (2003). Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning*, 52, 91–118.
- Moore, E. H. (1920). On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society*, 26, 394–395.
- Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods*, 11(1), 47–60.
- Morris, R. G. M., Garrud, P., Rawlins, J. N. & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, 297, 681–683.
- Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A. & Rosenbaum, R. S. (2006). The cognitive neuroscience of remote episodic, semantic and spatial memory. *Current Opinion in Neurobiology*, 16, 1–12.
- Naus, M. J. & Ornstein, P. A. (1983). Development of memory strategies: Analysis, questions, and issues. In M. T. H. Chi (ed), *Trends in memory development research* (pp. 1–30). Basel: S. Karger AG.
- Neumaier, A. (1998). Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Review*, 40, 636–666.
- Newman, M. & Kasznai, A. (2000). Spatial Memory and Aging: Performance on a Human Analog of the Morris Water Maze. *Aging, Neuropsychology, and Cognition*, 7(2), 86–93.
- O'Keefe, J. & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: University Press.
- Osterrieth, P. A. (1944). *Le test de copie d'une figure complexe*. Neuchâtel: Delachaux et Niestlé.
- Penrose, R. (1955). A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, 51, 406–413.
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyoung, R. & Richardson, C. (1995). A Redrawn Vandenberg and Kuse Mental Rotations Test: Different Versions and Factors That Affect Performance. *Brain & Cognition*, 28, 39–58.

- Phillips, K. & Silverman, I. (1997). Differences in the Relationship of Menstrual Cycle Phase to Spatial Performance on Two- and Three-Dimensional Tasks. *Hormones and Behavior*, 32, 167–175.
- Putz-Osterloh, W. (1977). Über Problemlöseprozesse bei dem Test Würfelaufgaben aus dem Intelligenztest IST und IST-70 von Amthauer. *Diagnostica*, 23, 252–265.
- Putz-Osterloh, W. & Lüer, G. (1979). Wann produzieren Probanden räumliche Vorstellungen beim Lösen von Raumvorstellungsaufgaben? *Zeitschrift für experimentelle und angewandte Psychologie*, 26(1), 138–156.
- Quaiser-Pohl, C., Geiser, C. & Lehmann, W. (2005). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences*, 40(3), 609–619.
- Rey, A. (1941). *L'examen psychologique dans les cas d'encéphalopathie traumatique*. In A. Rey (1969). *Psychologie clinique et neurologie*. Neuchâtel: Delachaux et Niestlé.
- Rönicke, J. (2003). *Gender differences in spatial abilities: Contribution of academic program and computer game experience in cross cultural perspective*. Unveröffentlichte Diplomarbeit, Institut für Psychologie I, Otto-von-Guericke-Universität Magdeburg.
- Roche, R., Mangaoang, M., Commins, S. & O'Mara, S. (2005). Hippocampal Contributions to Neurocognitive Mapping in Humans: A New Model. *Hippocampus*, 15, 622–641.
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Verlag Hans Huber.
- Samsonovich, A. & McNaughton, B. L. (1997). Path Integration and Cognitive Mapping in a Continuous Attractor neural network model. *The Journal of Neuroscience*, 17(15), 5900–5920.
- Sandstrom, N. L., Kaufman, J. & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research*, 6(4), 351–360.
- Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B., Witter, M. P., Moser, E. I. & Moser, M. B. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, 312, 758–762.
- Schönfeld, R., Lehmann, W., Ungewiß, J. & Leplow, B. (2007). *Leistungsdifferenzen in der Computerversion des Mental Rotations Test*. 18. Tagung der Fachgruppe Entwicklungspsychologie, Heidelberg, Deutschland.
- Schubert, T., Friedmann, F. & Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators & Virtual Environments*, 10(3), 266–281.
- Schuemie, M. & Mast, C. van der. (2001). VR Testbed Configuration For Phobia Treatment Research. In M.E.Domingo, J. Cebollada & C.P.Salvador (eds), *Proceedings of the Euromedia'2001 Conference* (pp.S 200–204). Valencia, Spain.

- Schultz, K. (1991). The Contribution of Solution Strategy to Spatial Performance. *Canadian Journal of Psychology*, 45(4), 474–491.
- Shahnaz, F. (2004). *Clustering method based on nonnegative matrix factorization for text mining*. Unpublished master thesis, Department of Computer Science, University of Tennessee, Knoxville.
- Shelton, A. L. & McNamara, T. P. (2004). Spatial memory and perspective taking. *Memory & Cognition*, 32(3), 416–426.
- Shepard, R. N. & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171, 701–703.
- Silverman, I., Choi, J., Mackewn, A., Fisher, M., Moro, J. & Olshansky, E. (2000). Evolved mechanisms underlying wayfinding: further studies on the hunter-gatherer theory of spatial sex differences. *Evolution and Human Behavior*, 21, 201–213.
- Silverman, I. & Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and data. In J. H. Barkow, L. Cosmides & J. Tooby (eds), *The adapted evolutionary psychology and the generation of culture* (pp. 533–553). New York: University Press.
- Skelton, R. W., Bukach, C. H., Laurance, H. E., Thomas, K. G. F. & Jacobs, W. J. (2000). Humans with traumatic brain injuries show place-learning deficits in computer-generated virtual space. *Journal of Clinical and Experimental Neuropsychology*, 22, 157–175.
- Skelton, R. W., Ross, S. P., Nerad, L. & Livingstone, S. A. (2006). Human spatial navigation deficits after traumatic brain injury shown in the arena maze, a virtual morris water maze. *Brain Injury*, 20(2), 189–203.
- Spiers, H. J. & Maguire, E. A. (2004). A 'landmark' study on the neural basis of navigation. *Nature Neuroscience*, 7(6), 572–574.
- Squire, L. S. (1986). Mechanisms of Memory. *Science*, 232, 1612–1619.
- Sturm, W., Willmes, K. & Horn, W. (1993). *Leistungsprüfsystem für 50–90jährige (LPS 50+)*. Göttingen: Hogrefe.
- Thurstone, L. L. (1950). Some Primary Abilities in Visual Thinking. *Proceedings of the American Philosophical Society*, 94(6), 517–521.
- Thurstone, L. L. & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Psychometric Monographs, Chicago: University of Chicago Press.
- Tiedemann, J. (1988). Zur Diagnostik kognitiver Stile. *Diagnostica*, 34, 289–300.
- Tikhonov, A. N. & Arsenin, V. A. (1977). *Solution of Ill-posed Problems*. Winston & Sons, Washington.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (eds), *Organization of memory* (pp. 381–403). New York: Academic Press.

- Ungerleider, L. & Mishkin, M. (1982). Two cortical visual systems. In M. A. Ingle, M. M. A. Goodale & R. Mansfield (eds), *Analysis of visual behaviour* (pp. 549–586). Cambridge: MIT Press.
- Van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia*, *42*, 1394–1413.
- Vandenberg, S. & Kuse, A. R. (1978). Mental Rotations: a group test of three-dimensional spatial visualization. *Perceptual Motor Skills*, *47*, 599–604.
- Voyer, D. & Hou, J. (2006). Type of Items and the Magnitude of Gender Differences on the Mental rotations test. *Canadian Journal of Experimental Psychology*, *60*(2), 91–100.
- Voyer, D. & Saunders, K. A. (2004). Gender differences on the mental rotations test: a factor analysis. *Acta Psychologica*, *117*(1), 79–94.
- Voyer, D., Voyer, S. & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A metaanalysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250–270.
- Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology*, *6*(4), 307–321.
- Waller, D. & Hodgson, E. (2006). Transient and Enduring Spatial Representations Under Disorientation and Self-Rotation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*(4), 867–882.
- Wang, R. F. & Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in cognitive sciences*, *6*(6), 376–382.
- Warren, R. & Wertheim, A. (eds). (1990). *Perception and control of self-motion*. Hillsdale: Lawrence Erlbaum Associates.
- Webb, A. (2002). *Statistical Pattern Recognition*. New York: John Wiley & Sons.
- Whishaw, I. Q. (1998). Place Learning in Hippocampal Rats and the Path Integration Hypothesis. *Neuroscience & Biobehavioral Reviews*, *22*(2), 209–220.
- Wilkniss, S. M., Jones, M. G., Korol, D. L., Gold, P. E. & Manning, C. A. (1997). Age-related differences in an ecologically based study of route learning. *Psychology of Aging*, *12*, 372–375.
- Wirtz, M. & Nachtigall, C. (2006). *Deskriptive Statistik - Statistische Methoden für Psychologen Teil 1* (4., überarbeitete Aufl.). Weinheim und München: Juventa.
- Witkin, H. A., Lewis, H. B., Hertzmann, M., Machover, K., Meissner, P. B. & Karp, S. A. (1954). *Personality through perception*. New York: Harper.
- Witmer, B. G. & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence*, *7*(3), 225–240.

- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46, 441–517.
- Zerssen, D. v. (1976). *Die Befindlichkeitsskala (Bf-S) - Manual*. Weinheim: Beltz Test.
- Zimmerman, M. E., Brickman, A. M., Paul, R. H., Grieve, S. M., Tate, D. F., Gunstad, J., Cohen, R. A., Aloia, M. S., Williams, L. M., Clark, R., Whitford, T. J. & Gordon, E. (2006). The relationship between frontal gray matter volume and cognition varies across the healthy adult lifespan. *American Journal of Geriatric Psychiatry*, 14(10), 823–833.

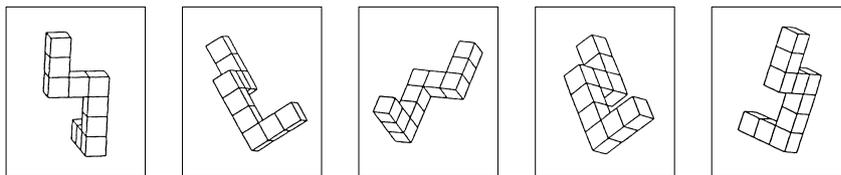
14 Tests und Fragebögen

14.1 Mental Rotations Test (revidierte Testform B)

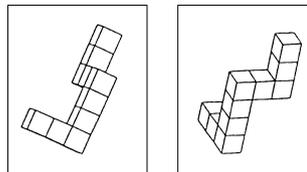
Mental Rotations Test (MRT- B (R))

Michael Peters, PhD

Sieh Dir bitte diese fünf Abbildungen an.



Es handelt sich dabei um dasselbe Objekt aus verschiedenen Perspektiven. Stell Dir nun vor, wie Du das Objekt drehen müsstest bzw. wie Du dich sich selbst um das Objekt bewegen müsstest, um jeweils die nächste Abbildung zu erzeugen.

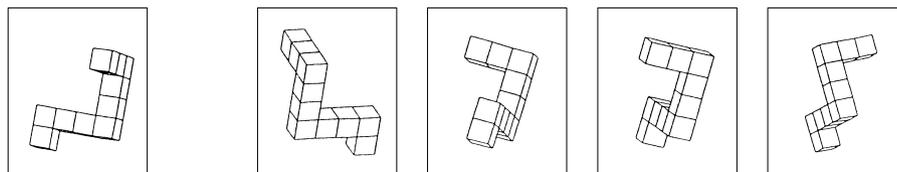


Hier siehst Du jetzt zwei Zeichnungen eines weiteren Objekts. Überzeuge Dich davon, dass dieses Objekt tatsächlich anders ist und sich nicht durch „Rotieren“ in das Objekt überführen läßt, das in den ersten fünf Zeichnungen zu sehen ist.

Sieh Dir jetzt dieses Objekt an:

Zwei dieser vier Abbildungen zeigen dasselbe Objekt. Kannst Du herausfinden welche? Kennzeichne diese mit einem großen Kreuz.

1.



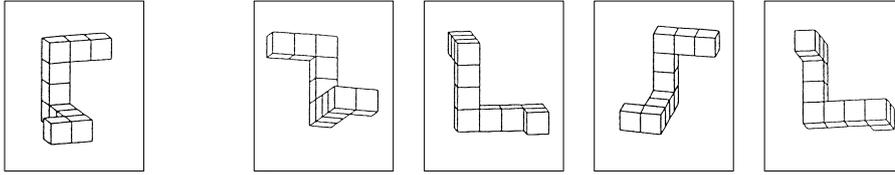
Wenn Du die erste und die dritte Zeichnung angekreuzt hast, hast Du dich richtig entschieden.

A

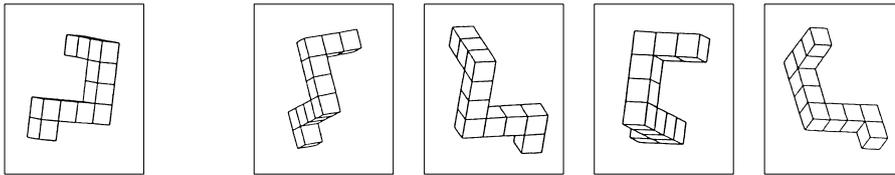
2

Hier siehst Du drei weitere Aufgaben. Wähle wiederum unter den vier Alternativen die beiden richtigen aus.

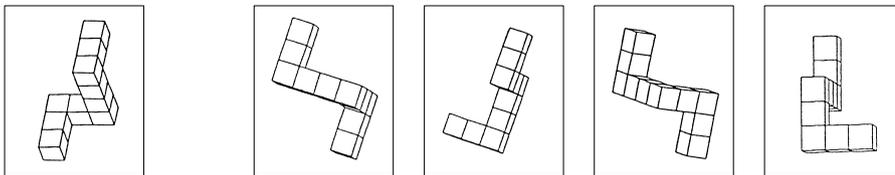
2.



3.



4.



Richtige Lösungen:

2: zwei und drei

3: eins und vier

4: eins und drei

Wenn Du den Test ausfüllst, denke immer daran: Bei jeder Aufgabe stimmen genau zwei Abbildungen mit der Vergleichsabbildung überein.

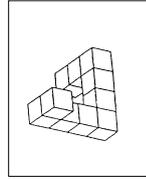
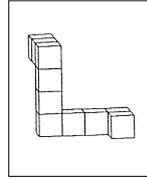
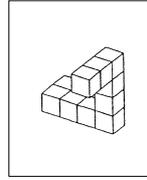
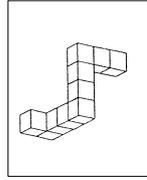
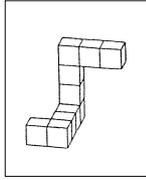
Du bekommst nur dann einen Punkt für die richtige Lösung, wenn Du die beiden richtigen Abbildungen angekreuzt hast. Wurde nur eine richtige angekreuzt, gibt es keinen Punkt.

B

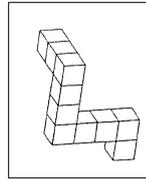
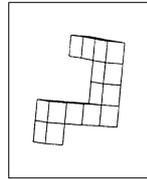
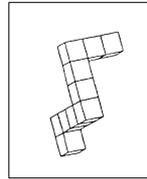
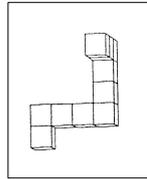
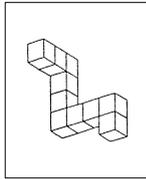
Teil 1 - Anfang

5

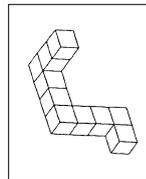
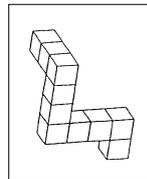
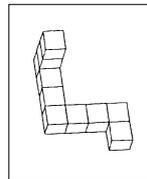
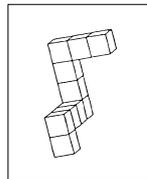
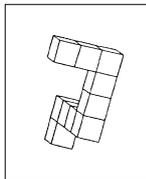
1.



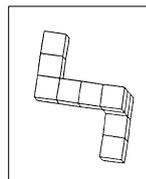
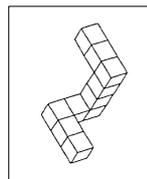
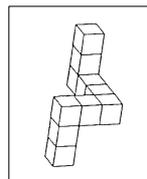
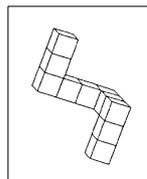
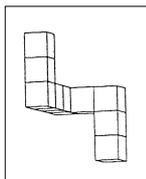
2.



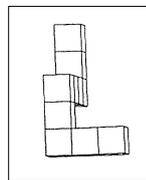
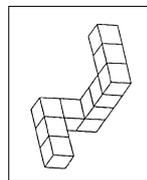
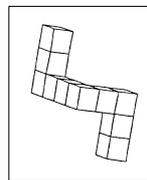
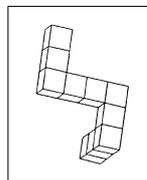
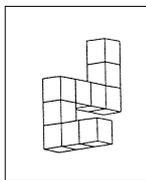
3.



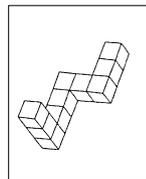
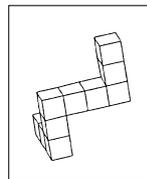
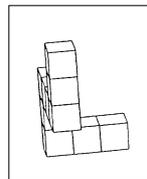
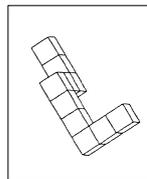
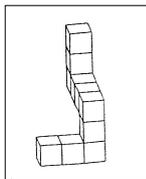
4.



5.



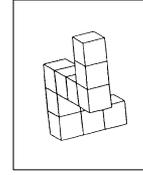
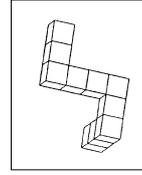
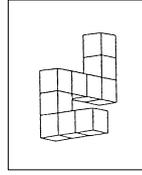
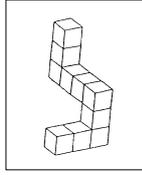
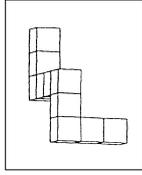
6.



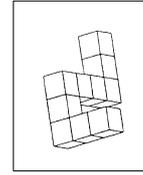
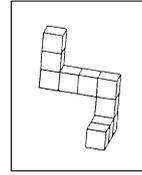
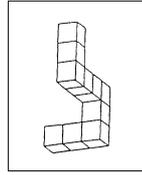
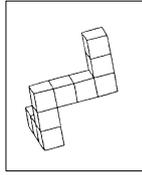
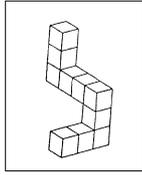
B

5

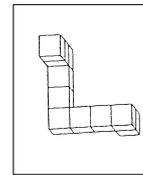
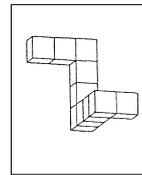
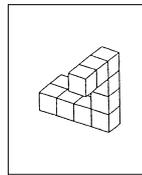
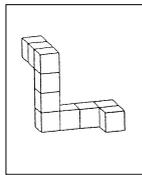
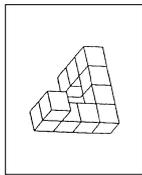
7.



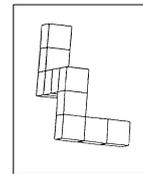
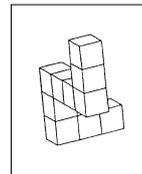
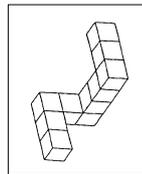
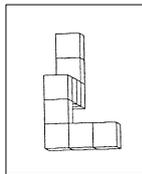
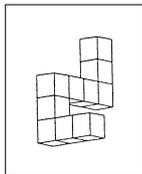
8.



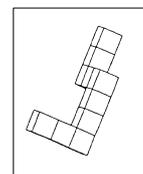
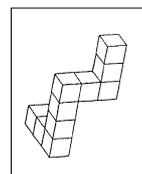
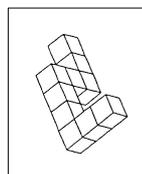
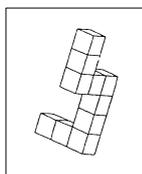
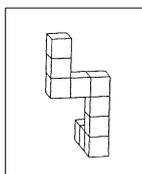
9.



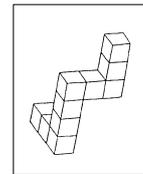
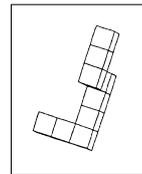
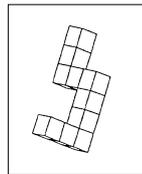
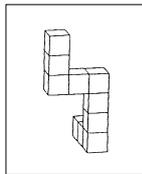
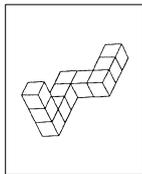
10.



11.



12.



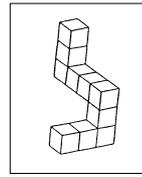
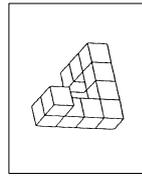
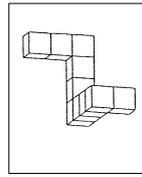
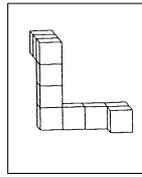
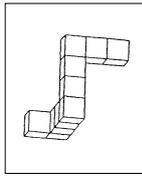
Teil 1 - Ende

B

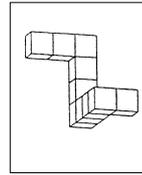
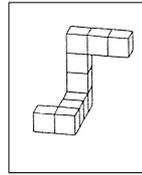
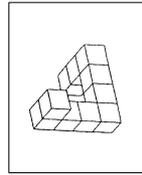
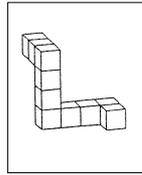
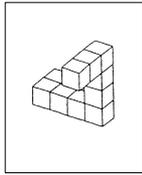
Teil 2 - Anfang

5

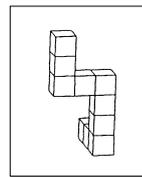
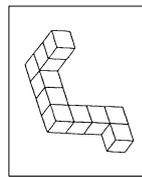
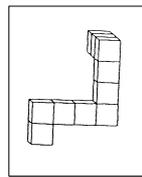
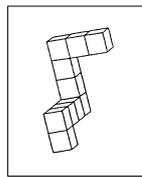
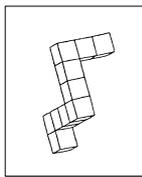
13.



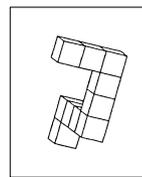
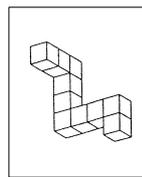
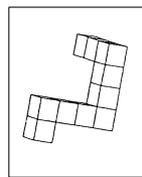
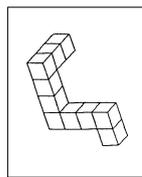
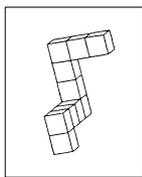
14.



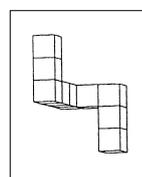
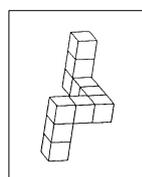
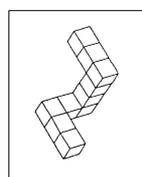
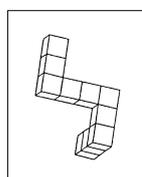
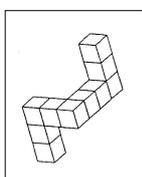
15.



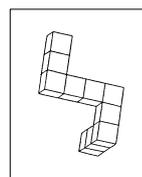
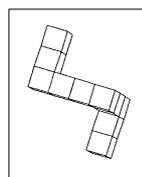
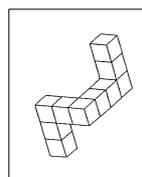
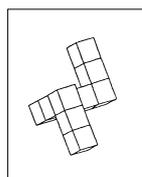
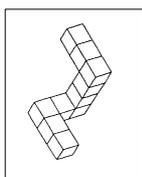
16.



17.



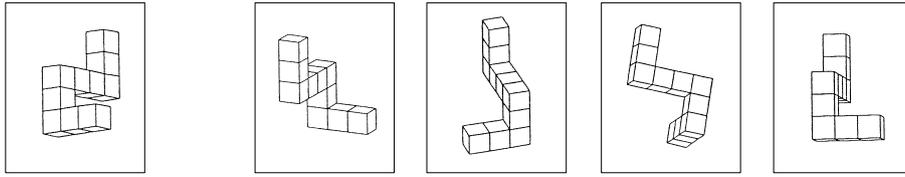
18.



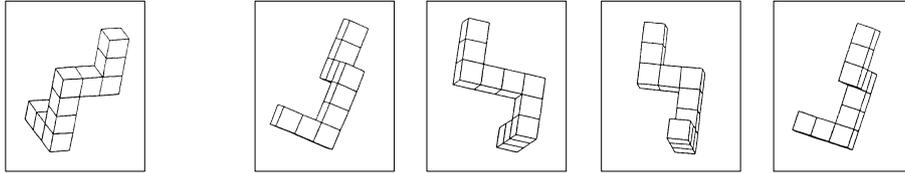
B

6

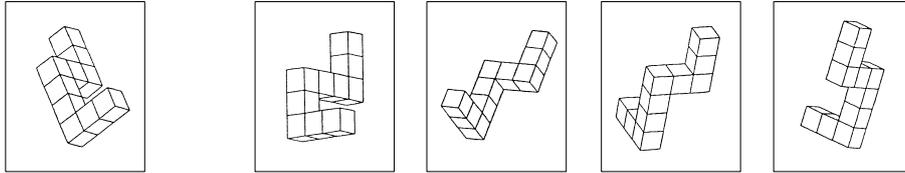
19.



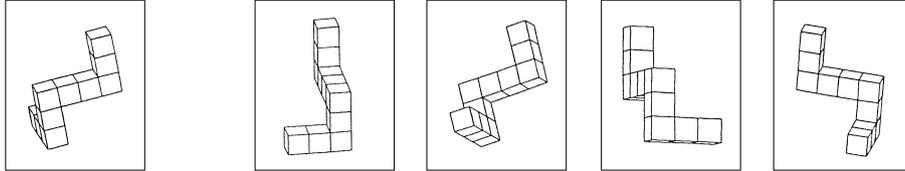
20.



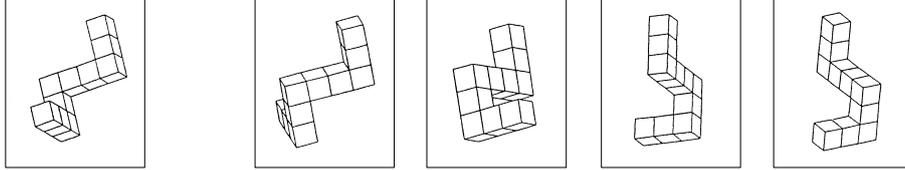
21.



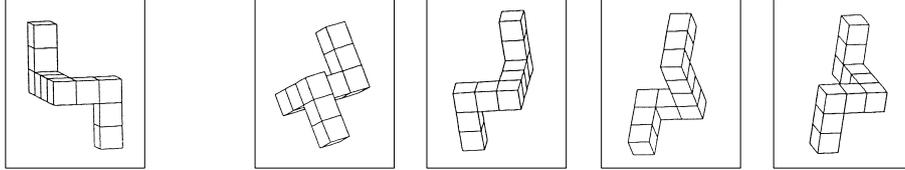
22.



23.



24.



Teil 2 - Ende

14.2 Test- und Aufgabeninstruktionen

14.2.1 Mental Rotations Test

Instruktionen zum Mental Rotations Test

Es ist zu beachten, dass der Test sehr sensitive bezüglich der Instruktion ist. Deshalb sollen den Probanden keine weiteren Informationen gegeben werden, außer denen, die auf den ersten beiden Seiten zur Verfügung stehen.

Instruktion der Beispiele

„Bitte nehmt Euch die erste Seite des Testbogens und seht euch die oberen fünf Abbildungen an. Es handelt sich dabei um dasselbe Objekt aus verschiedenen Perspektiven. Stellt euch nun vor, wie ihr das Objekt drehen müsstet bzw. wie ihr euch selbst um das Objekt bewegen müsstet, um jeweils die nächste Abbildung zu erzeugen.“



Bitte die rechte Hand zur Demonstration der Rotation verwenden!

Dauer: ca. 1 Minute

„Bitte seht euch nun die nächsten beiden Abbildungen an. Hier seht ihr jetzt zwei Zeichnungen eines weiteren Objekts. Überzeugt euch davon, dass dieses Objekt tatsächlich anders ist und sich nicht durch „Rotieren“ in das Objekt überführen lässt, das in den ersten fünf Zeichnungen zu sehen ist.“

Dauer: ca. ½ Minute

Instruktion des Problems

„Kommen wir nun zur Beschreibung der eigentlichen Aufgaben. Seht euch nun die erste Aufgabe unten auf der Seite an. Links seht ihr das Ausgangsobjekt. Auf der rechten Seite seht ihr vier weitere Abbildungen, von denen genau zwei das Ausgangsobjekt aus einer anderen Perspektive zeigen. Die beiden anderen Abbildungen zeigen Objekte, die nicht durch Rotation mit dem Ausgangsobjekt in Übereinstimmung gebracht werden können. Versucht in der ersten Aufgabe herauszufinden, welches die zwei richtigen Abbildungen sind. Kennzeichnet diese mit einem großen Kreuz. Die Antworten findet ihr darunter.“

Dauer: ca. 1 Minute

„Die erste und die dritte Abbildung stimmen mit der Ausgangsfigur überein. Ihr müsst beide richtigen Antworten finden, um einen Punkt zu bekommen. Nur eine richtige Antwort, oder eine richtige und eine falsche Antwort zählen nicht.“

Dauer: bis alle das Antwortprinzip verstanden haben

„Versucht nun die drei Aufgaben auf der zweiten Seite zu lösen. Die richtigen Antworten stehen wieder unter den Aufgaben. Deckt diese zunächst ab und vergleicht sie erst dann mit euren Lösungen, wenn ihr alle drei Aufgaben bearbeitet habt. Wer fertig ist, wartet – bitte NICHT weiter blättern!“

Dauer: ca. 5 Minute

Wenn alle Fragen zu den Übungsaufgaben besprochen wurden kann mit dem eigentlichen Test begonnen werden.

Instruktion des Tests

„Jetzt sind wir bereit mit dem Test zu beginnen. Denkt immer daran, wenn ihr den Test ausfüllt: Bei jeder Aufgabe stimmen genau zwei Abbildungen überein. Ihr bekommt nur dann einen Punkt, wenn ihr beide richtigen Abbildungen angekreuzt habt. Sonst gibt es keinen Punkt.“

Wir beginnen mit dem Test, wenn ich sage: „Bitte jetzt beginnen!“. Legt bitte eure Schreibgeräte aus der Hand wenn ich „Stopp“ sage – auch wenn ihr mitten in der Bearbeitung einer Aufgabe seid.“

Zuerst lösen wir die Aufgaben auf der Seite drei und vier. Es sind 12 Aufgaben und ihr habt dafür 3 Minuten Zeit zur Verfügung. Wer ehr fertig ist, legt ebenfalls seinen Stift aus der Hand und wartet. Bitte NICHT weiterblättern!

Danach machen wir 2 Minuten Pause, (in denen jeder auf seinem Platz bleibt!). Wir beginnen dann gemeinsam mit weiteren 12 Aufgaben auf den Seite 5 und 6. Dafür habt ihr wiederum 3 Minuten Zeit zur Verfügung.“

Teil 1: „Bitte jetzt beginnen - Stopp“

Dauer: 3 Minuten

- Pause -

Dauer: 2 Minuten

Teil 2: „Bitte jetzt beginnen - Stopp“

Dauer: 3 Minuten

14.2.2 Kompass-Zeigeaufgabe

Instruktion Kompass-Zeigeaufgabe

Es wird eine Parklandschaft mit Bäumen, Hügeln und Wegen dargeboten. Versuchen Sie sich in die Person hineinzusetzen, die durch den Park läuft. Die Person läuft selbständig und hält an einer Stelle an. Sie hat sich leider verlaufen und möchte wissen, wo sie gestartet ist. Ihre Aufgabe ist es nun, einen Kompass so einzustellen, dass die Nadel den kürzesten Weg zurück zum Ausgangspunkt anzeigt. Danach beginnt ein neuer Durchgang. Insgesamt wird es acht Durchgänge geben. Zum Anfang werden Sie zwei Testdurchgänge durchlaufen. Dann werden Ihnen drei verschiedene Wege zunächst in einer langsamen Version, anschließend in einer schnellen Version dargeboten.

Haben Sie noch Fragen? Stellen Sie diese bitte!

14.2.3 Virtuelles Training

Instruktion zum virtuellen Training

Zunächst bekommen Sie die Gelegenheit, sich mit der Bedienung der Steuerkonsole vertraut zu machen. Es werden insgesamt 7 Kisten vorhanden sein, die mit den Zahlen 1 bis 7 versehen sind. Ihre Aufgabe ist es, die Kisten in aufsteigender Reihenfolge zu öffnen. Beginnen Sie also mit der Kiste Nr. 1 und enden mit der Kiste Nr. 7. Wenn Sie gegen eine Kiste laufen, wird sich der Deckel öffnen und Sie können hineinschauen. Wenn Sie sich dann rückwärts bewegen, wird sich der Deckel wieder schließen und Sie können die nächste Kiste aufsuchen. Versuchen Sie es einfach mal!

14.2.4 Water-Maze-Aufgabe

Instruktion zur virtuellen Water-Maze-Aufgabe

Stellen Sie sich vor, Sie wären auf einer einsamen Insel gestrandet. Irgendwo auf dieser Insel ist eine Schatzkiste versteckt, die nützliche Dinge enthält. Ihre Aufgabe ist es, diese Kiste zu finden. Im ersten Durchgang werden Sie blau-weiße Bälle sehen, die Ihnen helfen sollen, den Weg zur Kiste zu finden. Nachdem Sie die Kiste gefunden haben, bleibt Ihnen noch etwas Zeit, um sich auf der Insel umzuschauen. In den folgenden Durchgängen werden diese Bälle nicht mehr vorhanden sein und Sie erhalten keine zusätzliche Zeit mehr. Versuchen Sie sich die Position der Kiste bereits im ersten Durchgang zu merken, denn Sie sollen die Kiste mehrmals finden. Die Kiste wird sich stets am gleichen Ort befinden. Nur auf der Insel wird sich gegen Ende des Versuchs etwas verändern. Schauen Sie sich also gut auf der Insel um!

Haben Sie noch Fragen? Falls ja, stellen Sie diese Fragen bitte!

14.3 Eigene Fragebögen

14.3.1 Demographische Daten

14.3.1.1 Experiment 1

Fragebogen 1

Untersuchung zur Raumvorstellung

Wir bitten Dich zunächst, einige Angabe zu machen. Alle Daten werden selbstverständlich anonym und vertraulich behandelt und dienen ausschließlich wissenschaftlichen Zwecken.

Tag der Untersuchung (Datum):

Alter: Klassenstufe:

Geschlecht: weiblich ?
männlich ?

Muttersprache:

Ich bin Rechtshänder(in) ?
Linkshänder(in) ?

Mathematikzensur:

Ich habe schon einmal an einem Raumvorstellungstest teilgenommen?

Ja ?

Nein ?

Wenn ja, welche(r)?

Zur anonymen Verarbeitung deiner Testergebnisse dient ein **Codewort**.

Das Codewort ist aus individuellen Daten zusammengesetzt, und zwar

1. den ersten beiden Buchstaben des Vornamens Deiner Mutter

2. den Tag Deines Geburtstages (z.B. am 5. April geboren: **05**)

3. den ersten beiden Buchstaben Deines Geburtsortes:

Schreibe bitte das **gesamte** Codewort in eine Zeile.
Du wirst es noch einige Male angeben müssen.

14.3.2 Computererfahrung und Vorerfahrung Raumvorstellungstest

14.3.2.1 Experiment 1

Fragebogen 2

Fragebogen zur Computererfahrung

Bei der Beantwortung dieses Fragebogens gibt es keine richtigen und falschen Antworten. Ihr solltet beachten, dass immer das anzukreuzen ist, was am Ehesten zutrifft.

Zuerst bitten wir Dich um die Angabe deines Codeworts:

Beantworte bitte nun die folgenden Fragen!

Ich arbeite am Computer mit

- ? Anwendungsprogrammen
- ? Computerspielen
- ? beiden
- ? keinem von beiden

Um welche Anwendungsprogramme handelt es sich?

.....

.....

Wie oft arbeitest Du insgesamt mit den genannten Anwendungsprogrammen?

- ? täglich
- ? weniger als 1 Stunde
- ? mehr als 1 Stunde
- ? mehr als 2 Stunden
- ? mehr als 3 Stunden
- ? öfter, aber nicht jeden Tag
- ? selten

Um welche Spiele handelt es sich?

.....

.....

Wie oft spielst Du insgesamt am Computer?

- ? täglich
- ? weniger als 1 Stunde
- ? mehr als 1 Stunde
- ? mehr als 2 Stunden
- ? mehr als 3 Stunden
- ? öfter, aber nicht jeden Tag
- ? selten

14.3.2.2 Experiment 3

Erfahrungsfragebogen

Schätzen Sie bitte anhand der folgenden Aussagen Ihre Erfahrungen im Umgang mit räumlichen Anforderungen am PC und in der realen Welt ein! Es gibt keine richtigen oder falschen Antworten. Kreuzen Sie bitte immer das an, was am ehesten auf Sie zutrifft!

- Wie oft arbeiten Sie am Computer?
 - täglich
 - wöchentlich
 - monatlich
 - selten
 - gar nicht

 - Wie oft spielen Sie dreidimensionale Computerspiele?
 - täglich
 - wöchentlich
 - monatlich
 - selten
 - gar nicht

 - Wenn ja, welche Spiele?.....

 - Wie schätzen Sie Ihre Fähigkeit ein, sich in neuen, also fremden Umgebungen zurechtzufinden?

eher gut	eher schlecht
----------	---------------

 - Wie schätzen Sie Ihre Fähigkeit ein, sich anhand von Landkarten (Stadtplan, Autoatlas etc.) zu orientieren?

eher gut	eher schlecht
----------	---------------
-
- Haben Sie vorher schon Computeraufgaben dieser Art gelöst?

ja	nein
----	------

 - Wie schwierig fanden Sie die Computeraufgaben?

eher leicht	eher schwer
-------------	-------------

14.3.3 Strategiefragebögen

14.3.3.1 Mental Rotations Test

Fragebogen 1

Fragebogen zur räumlichen Vorstellung

Bei der Beantwortung dieses Fragebogens gibt es keine richtigen und falschen Antworten. Ihr solltet beachten, dass immer das anzukreuzen ist, was am Ehesten zutrifft. Bei jeder Frage kann ja oder nein angekreuzt werden, unabhängig von der Antwort der vorhergehenden Frage.

Zuerst bitten wir Dich um die Angabe deines Codeworts:
Beantworte bitte nun die folgenden Fragen!

1. Ich hatte Spaß an der Lösung der Aufgaben.
? ja ? nein

2. Die Bearbeitung der Aufgaben fiel mir schwer.
? ja ? nein

Nun zu Euren Lösungswegen im Mentalen Rotationstest!

3. Ich schließe zuerst unpassende Figuren aus, um zur Lösung der Aufgabe zu gelangen.
? ja ? nein

4. Ich ging systematisch vor und prüfte die vier Lösungsfiguren einer Aufgabe in der angebotenen Reihenfolge.
? ja ? nein

5. Ich überprüfte, welche der vier Lösungsfiguren zusammenpassen könnten und traf dann meine Entscheidung.
? ja ? nein

6. Ich verglich jede Lösungsfigur einzeln mit der Ausgangsfigur, auch wenn ich schon zwei mögliche Lösungen gefunden hatte.
? ja ? nein

7. Ich verglich zur Sicherheit die zunächst gefundene Lösungsfigur sowohl mit den anderen Lösungsfiguren als auch mit der Zielfigur.
? ja ? nein

8. Ich suchte zuerst nach einer passenden Lösungsfigur und verglich diese dann mit den drei anderen angebotenen Lösungsfiguren.
? ja ? nein

9. Beim Vergleich der Figuren habe ich die ganze Figur in Gedanken kreisen lassen, ohne dabei eine bestimmte Teilfläche im Auge zu behalten.
? ja ? nein

10. Ich habe beim Vergleich der Figuren die ganze Figur in Gedanken kreisen lassen, wobei ich die Augen immer auf eine bestimmte Teilfläche gerichtet hatte.
? ja ? nein

1-2

11. Beim Vergleich der Figuren habe ich nur einen Teil der Figur in Gedanken kreisen lassen, ohne eine bestimmte Teilfläche im Auge zu behalten.
 ? ja ? nein
12. Ich habe beim Vergleich der Figuren nur einen Teil der Figur in Gedanken kreisen lassen, wobei ich immer eine fixierte Teilfläche im Auge behielt.
 ? ja ? nein
13. Ich habe mir die Schritte in Worten durchdacht. (z.B. 3 Würfel hoch und 2 Würfel nach rechts).
 ? ja ? nein
14. Ich habe mir die Figuren nur als Bild im Kopf vorgestellt und die Schritte nicht in Worten durchdacht.
 ? ja ? nein
15. Ich nutzte meine Hände zur Simulation der Figuren.
 ? ja ? nein
16. Ich benutzte Gegenstände (z.B. Papier, Lineal, ...) zur Simulation der Figuren.
 ? ja ? nein
17. Ich habe zur Lösung der Aufgaben keine Hilfsmittel benutzt.
 ? ja ? nein
18. Ich entwickelte nach den ersten Aufgaben einen Lösungsplan und wendete ihn bei den nachfolgenden Aufgaben konsequent an.
 ? ja ? nein
19. Ich veränderte im Verlauf des Tests meinen anfänglichen Lösungsweg.
 ? ja ? nein
20. Ich habe mehrere Strategien zur Lösung der Aufgaben entwickelt und wechselte meine Strategien von Aufgabe zu Aufgabe.
 ? ja ? nein
21. Ich legte den Schwerpunkt auf die richtige Lösung der Aufgaben, nicht auf das Zeitlimit.
 ? ja ? nein
22. Ich legte den Schwerpunkt darauf, die Aufgaben im Zeitlimit zu schaffen und achtete nicht so sehr auf die richtige Lösung.
 ? ja ? nein
23. Ich war mir meiner Antwort sicher, bevor ich zur nächsten Aufgabe überging.
 ? ja ? nein
24. Ich habe bei einigen Antworten geraten.
 ? ja ? nein

14.3.3.2 Kompass-Zeigeraufgabe

Strategiefragebogen - Heimkehreraufgabe

- Wie haben Sie versucht, Ihre Richtungsangaben zu bestimmen?
- Haben Sie sich den Park von oben, also aus der Vogelperspektive vorgestellt?

ja	nein
----	------
- Sind Sie den Weg, den die Person gelaufen ist, in Gedanken rückwärts gelaufen?

ja	nein
----	------
- Haben Sie sich bei jeder Richtungsänderung überlegt, wie sich die Richtung zum Startpunkt ändert?

ja	nein
----	------
- Haben Sie immer die aktuelle Position mit der Ausgangsposition verglichen?

ja	nein
----	------
- Haben Sie sich bei Ihren Richtungsangaben auf Ihr Gefühl verlassen?

ja	nein
----	------
- Sind Ihnen die schnelleren Durchgänge leichter gefallen?

ja	nein
----	------

14.3.3.3 Water-Maze-Aufgabe

Strategiefragebogen - Navigationsaufgabe

- Wie haben Sie versucht, sich auf der Insel zu orientieren?
- Waren Ihnen die Bälle im ersten Durchgang eine Hilfe, um die Schatzkiste zu finden?
ja nein
- War Ihnen die Zeit zum Umschauen im ersten Durchgang eine Hilfe, um sich die Position der Schatzkiste zu merken?
ja nein
- Haben Sie in den folgenden Durchgängen eine bestimmte Strategie verwendet, um die Schatzkiste zu finden?
ja nein
- Falls ja, welche Strategie?
- Haben Sie sich an den vorhandenen Gegenständen (Segelboot, Leuchtturm, Windrad, Heißluftballon) orientiert, um die Schatzkiste zu finden?
ja nein
- An welchen Gegenständen haben Sie sich orientiert?
Segelboot
Leuchtturm
Windrad
Heißluftballon
- Falls Sie sich nur am Windrad und/oder Leuchtturm orientiert haben, warum?
- Falls Sie sich nicht an den vorhandenen Gegenständen orientiert haben, wie sind Sie dann vorgegangen?
- Wie sind Sie vorgegangen, als nur noch der Leuchtturm und das Windrad auf der Insel vorhanden waren, um die Kiste zu

finden?

- Hatten Sie in diesem Durchgang große Schwierigkeiten, die Kiste zu finden?

ja nein

- Wie sind Sie vorgegangen, als keiner der Gegenstände mehr auf der Insel vorhanden war, um die Kiste trotzdem zu finden?

- Hatten Sie in diesem Durchgang große Schwierigkeiten, die Kiste zu finden?

ja nein

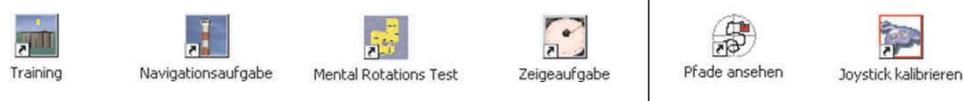
15 Technische Details - Softwareentwicklung

15.1 Hardwarevoraussetzungen und Installation

Technische Voraussetzungen. Die VR-Experimente können auf jedem IBM-kompatiblen Computersystem mit einem 32-Bit-Prozessor ab einer Taktrate von mindestens 700 MHz (z. B. Intel® Pentium-III® oder vergleichbare) und mindestens 128 MB RAM eingesetzt werden. Es wird die Verwendung einer hochwertigen 3D-Grafikkarte mit Hardwarebeschleunigung und mindestens 32 MB VRAM empfohlen. Die besten Ergebnisse in Qualität und Leistung wurden in den bisherigen Anwendungen mit ATI-Grafiksystemen (z. B. ATI Radeon™-Serie) erreicht. Als Bildschirm sind hochauflösende CRT-Displays mit hohen Bildwiederholraten (ca. 90 Hz) oder TFT-Displays zu verwenden. Die sichtbare Bildschirmdiagonale sollte wenigstens 40 cm (ca. 15 Zoll) betragen.

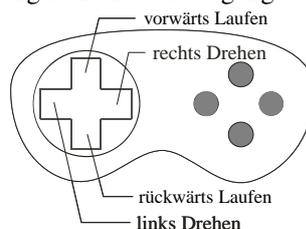
Installation. Zur Installation wird das auf der CD befindliche VisualBasic®-Skript *setup.vbs* gestartet. Es werden alle Programme und zusätzlichen Dateien auf die Festplatte des Computers in das Verzeichnis C:\VR kopiert. Danach werden die Verknüpfungen von Abbildung 15.1 auf dem Desktop angelegt. Mit diesen Verknüpfungen können die entsprechenden VR-Tests und Zusatzprogramme gestartet werden.

Abbildung 15.1: Desktopverknüpfungen zum Starten der VR-Experimente (linke Seite) und der Zusatzprogramme (rechte Seite).



Eingabegeräte. Im MRT und im VPT werden alle Eingaben über die Computermaus vorgenommen. Zur Navigation im VWM wird ein Joystick oder ein Gamepad benötigt.

Abbildung 15.2: Tastenbelegung des D-Pad.



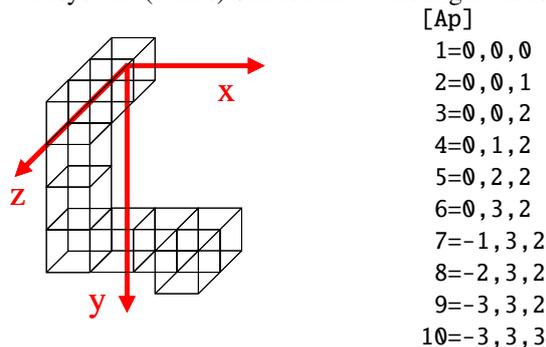
Gamepads lassen sich leichter bedienen und werden i. R. besser von den Versuchspersonen akzeptiert als Joysticks. Besonders ungeübte Versuchsteilnehmer erlernen sehr schnell den Umgang mit einem Gamepad. Zur Steuerung dient das D-Pad auf der linken Seite der Steuerkonsole (s. Abb. 15.2). Die Abbildung zeigt außerdem die Belegung der Richtungstasten.

15.2 Computer Mental Rotations Test

15.2.1 Konfiguration

Konstruktion der Shepard-Metzler-Figuren. Grundlegend für die Implementierung des MRT ist der Entwurf der 3d-Modelle der Shepard-Metzler-Figuren. Diese sind im Computertest jeweils aus 10 einzelnen Würfeln der Kantenlänge 1 zusammengesetzte virtuelle Objekte. Die Objekte werden während der Laufzeit aus den Angaben aus einer Konfigurationsdatei, die sich im Verzeichnis `c:\VR\env\MRT` befinden muss, konstruiert. Abbildung 15.3 zeigt das Prinzip dieser Konstruktion.

Abbildung 15.3: Zusammenbau der Shepard-Metzler-Figuren. (links) Skizze der Konstruktion im Koordinatensystem. (rechts) Sektion in der Konfigurationsdatei.



In der Konfigurationsdatei können beliebig Sektionen angelegt werden, deren Name in eckige Klammern geschrieben wird. Der Name dient später zum Aufruf der jeweiligen Figuren innerhalb des Experiments und muss mit einem Buchstaben beginnen. Zu einer so definierten Sektion gehören die Schlüsselwörter 1 bis 10 (s. Abb. 15.3 rechts). Jedes dieser Schlüsselwörter steht für einen der 10 Basiswürfel. Parameter sind die XYZ-Koordinaten eines Würfels im kartesischen Koordinatensystem (s. Abb. 15.3 links).

Für den MRT sind auf diese Weise 10 Grundfiguren (5 Basisfiguren und 5 zu den Basisfiguren gespiegelte Figuren) generiert worden. Abbildung 15.4 zeigt alle 10 Figuren als Screenshot aus dem Computertest. Ausschließlich diese 10 Grundfiguren wurden in den Items des MRT verwendet. Den Aufbau der einzelnen Items (inkl. Übungsitams) aus dem MRT der Testformen A und B fassen die Tabellen 15.1 und 15.2 zusammen. Diese Tabellen dienen der strukturellen Analyse des ursprünglichen Testmaterials und der originalgetreuen Implementierung der einzelnen Items in den Computerversionen.

Abbildung 15.4: 10 Grundfiguren des Computertests.

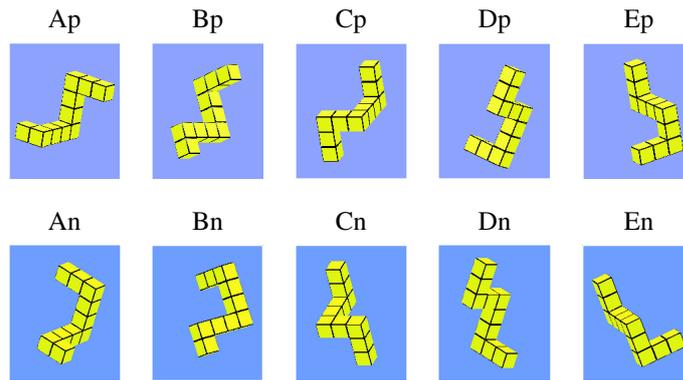


Tabelle 15.1: Aufbau der Items des MRT-A aus den 10 Grundfiguren nach Shepard&Metzler.

Item	Ziel	Alternativen				Typ	Okklusion
Übung 1	Bp	Bp	Bn	Bp	Bn	I	
Übung 2	Ap	An	Ap	Ap	An	I	
Übung 3	Bn	Bn	Bp	An	Bn	III	
Übung 4	Cn	Cn	Dp	Cn	En	II	
1	Ap	Ap	An	Ap	An	I	
2	An	An	Ap	Ap	An	I	
3	Ap	Bp	Ap	Bn	Ap	II	
4	An	Bn	An	An	Bp	II	
5	Bn	Bn	Bp	Bn	Bp	I	
6	Bp	Bp	Bn	Bn	Bp	I	
7	Bn	An	Bn	Ap	Bn	II	
8	Bp	Ap	Bp	Bp	An	II	
9	Cp	Cn	<u>Cp</u>	Cn	Cp	I	<u>1</u>
10	Cn	<u>Cn</u>	Cp	<u>Cp</u>	Cn	I	<u>2</u>
11	En	Ep	Ep	En	<u>En</u>	I	<u>1</u>
12	Ap	An	Ap	Ap	An	I	
13	Cp	Cp	Cp	Dp	Ep	II	
14	Cn	Ep	Cn	Dn	<u>Cn</u>	II	<u>1</u>
15	Dn	Dp	<u>Dn</u>	Dn	Dp	I	<u>1</u>
16	Dp	Dp	Dn	Dn	Dp	I	
17	Dn	Cn	<u>Dn</u>	Ep	Dn	II	<u>1</u>
18	Dp	<u>Ep</u>	Dp	Dp	Cp	II	<u>1</u>
19	Ep	Ep	<u>En</u>	Ep	En	I	<u>1</u>
20	En	En	Ep	Ep	En	I	
21	Ep	Cp	Ep	Dn	Ep	II	
22	En	Dn	En	En	Cp	II	
23	Bn	Bn	Bp	An	Bn	III	
24	Cn	Cn	Dp	Cn	En	II	

Tabelle 15.2: Aufbau der Items des MRT-B aus den 10 Grundfiguren nach Shepard&Metzler.

Item	Ziel	Alternativen				Typ	Okklusion
Übung 1	Bp	Bp	Bn	Bp	Bn	I	
Übung 2	Ap	An	Ap	Ap	An	I	
Übung 3	Bn	Bn	Bp	An	Bn	III	
Übung 4	Cn	Cn	Dp	Cn	En	II	
1	Ap	An	Ap	Ap	An	I	
2	Bn	Bp	Bn	Bn	Bp	I	
3	Bp	Bp	Bn	Bp	Bn	I	
4	Cp	Cp	Cp	Cn	Cn	I	
5	En	En	Cn	Cn	En	II	
6	Ep	Dp	Ep	Ep	Dp	II	
7	En	Ep	En	En	Ep	I	
8	Ep	Ep	En	Ep	En	I	
9	An	Ap	Ap	An	An	I	
10	En	En	Cn	Ep	En	III	
11	Dp	Dp	Dp	Dn	Dn	I	
12	Dp	Dp	Dn	Dp	Dn	I	
13	An	Ap	An	An	Ep	III	
14	Ap	Ap	An	Ap	An	I	
15	Bn	Bp	Bn	Bn	Dp	III	
16	Bp	Bp	Bn	Bn	Bp	I	
17	Cp	En	Cn	Cp	Cp	III	
18	Cn	<u>Cn</u>	Cp	Cn	En	III	<u>1</u>
19	En	Ep	Ep	En	En	I	
20	Dn	Dn	En	Dn	Dp	III	
21	Dp	En	Dp	Dn	Dp	III	
22	Ep	Ep	En	En	Ep	I	
23	En	Ep	En	En	Ep	I	
24*	Cn	<u>Cn</u>	Cp	Cn	Cp	I	<u>1</u>

*im Original fehlerhaftes Item

Konfiguration des Testablaufs. Das Layout und die Funktionalität jeder Bildschirmseite im Computer-MRT wird innerhalb der Konfigurationsdatei durch jeweils eine Sektion festgelegt. Die Sektionen müssen mit Zahlen von 1 bis N fortlaufend nummeriert werden. Innerhalb einer Sektion können die in Tabelle 15.3 zusammengefassten Schlüsselwörter und Parameter verwendet werden. Im wesentlichen sind dabei die *Viewports*, in denen die 3d-Objekte der Shepard-Metzler-Figuren dargestellt werden, Textfenster und Buttons zu platzieren und dimensionieren.

Die aktuellen Parameterdateien für den Computer-MRT befindet sich im Verzeichnis $C:\backslash VR\env\MRT\backslash$. Es sind 8 Konfigurationen vordefiniert, die beim Start des Experiments über das Menü *Testform* ausgewählt werden können. Dabei handelt es sich um die Tests MRT-A und MRT-B mit Instruktionen jeweils für Kinder bzw. Erwachsenen mit 2×3 min bzw. 2×5 min Zeitlimit. Die Konfigurationen können an Ihren Dateinamen eindeutig identifiziert werden

Tabelle 15.3: Schlüsselwörter der Sektionen [1] bis [N] im Computer-MRT.

Schlüsselwort	Syntax	Bedeutung
Title-1	x, y, Breite, Höhe, Text	Textfenster mit 3d-Text und animierten 3d-Figuren
Text-1..4	x, y, Breite, Höhe, Text	Textfenster mit normalem 12pt-Text
ButtonNext	x, y, Breite, Text	Bildschirmseite vor
ButtonBack	x, y, Breite, Text	Bildschirmseite zurück
ButtonCancel	x, y, Breite, Text	Programmabbruch
Port-1..7	x, y, Breite, Auswählbar, 3d-Figur, Pitch, Roll, Yaw, Interaktiv, Vorausgewählt	Viewport einer 3d-Figur, Auswählbar: kann angekreuzt werden, Interaktiv: kann aktiv gedreht werden, Vorausgewählt: ist bereits angekreuzt
TimeLimit	Zeit, SektionsNr.	Zeitlimit in Sekunden, Sprung zu Sektions-Nr. bei Ablauf
Score	1	Item in den Summenscore aufnehmen

(z.B. MRT-A Kinder 3min.ini). Eines Beispiel, wie eine Konfigurationsdatei zu gestalten ist, gibt 15.5 anhand eines Items aus dem MRT.

Abbildung 15.5: Beispiel für eine Konfigurationsdatei des Computer-MRT: Definition der Grundfiguren An und Ap (nur angedeutet) und deren Verwendung in einem Item.

```
[An]
1=0,0,0
2=...
[Ap]
...

[1]
Port-1= 17,175,170,0, Ap, 15, -74, 0, 0
Port-2=272,175,170,1, Ap, 31, 33, 0, 0
Port-3=459,175,170,1, An, -20, -133, 0, 0
Port-4=646,175,170,1, Ap, 16, 120, 0, 0
Port-5=833,175,170,1, An, -13, 11, 0, 0

ButtonNext=906,520,96, Weiter
Text-1=7,140, 20, 20, 1.
Score=1
TimeLimit=300, 26
```

15.2.2 Datenformat

Aus dem Computer-MRT werden die Antworten und Summenscores pro Versuchsperson im Format einer EXCEL[®]-Tabelle geschrieben. Tabelle 15.4 zeigte eine solche Protokolldatei. Die erste Zeile enthält die Antworten für jedes Item, codiert mit Nummer des Items und den Buchstaben a bis d für die vier Antwortalternativen. In der zweiten Zeile werden die Punkte für jede Item, die Summenscores für die beiden Testhälften und der Gesamtscore nach dem *strengen Punktesystem* protokolliert.

Die zweite Zeile enthält die Bewertung nach dem *liberalen Punktesystem*. Diese Dateien befinden sich nach Durchführung eines Test im Verzeichnis C:\VR\env\MRT\DATA\

Tabelle 15.4: Format der Protokolldatei aus dem Computer-MRT.

1ac	2bc	3ac	4ab	...	23bc	24	Teil1	Teil2	Gesamt
0	1	1	1	...	1	0	10	8	18
1	2	2	2	...	2	0	22	19	41

15.3 Kompass-Zeigeaufgabe

15.3.1 Konfiguration

Verwendung vordefinierter Pfade. Zur Verwendung im Kompass-Zeigeaufgabe können verschieden Pfade in beliebiger Reihenfolge zu einem Experiment zusammengestellt werden. Es stehen derzeit insgesamt 42 bereits berechnete Pfade zur Auswahl (Tab. 15.7). Diese Pfade variieren in der Anzahl der gelaufenen Abzweigungen (2 bis 4 Turns) und den Winkeln der Abzweigungen. Außerdem bestehen Unterschiede bei ansonsten identischen Pfaden in einzelnen Wegsegmenten, so dass der auf dem Kompass einzustellende Winkel variiert.

Die für das Experiment ausgewählten Pfade werden unter Angabe des entsprechende Namens aus Tabelle 15.7 in die Datei C:\VR\env\VPT\paths.cfg eingetragen. Getrennte durch ein Leerzeichen steht hinter jedem Namen die gewünschte Abspielgeschwindigkeit, die nur in ganzzahligen Vielfachen erhöht werden kann. Abbildung 15.6 zeigt als Beispiel, die in der empirischen Untersuchung verwendete Konfiguration.

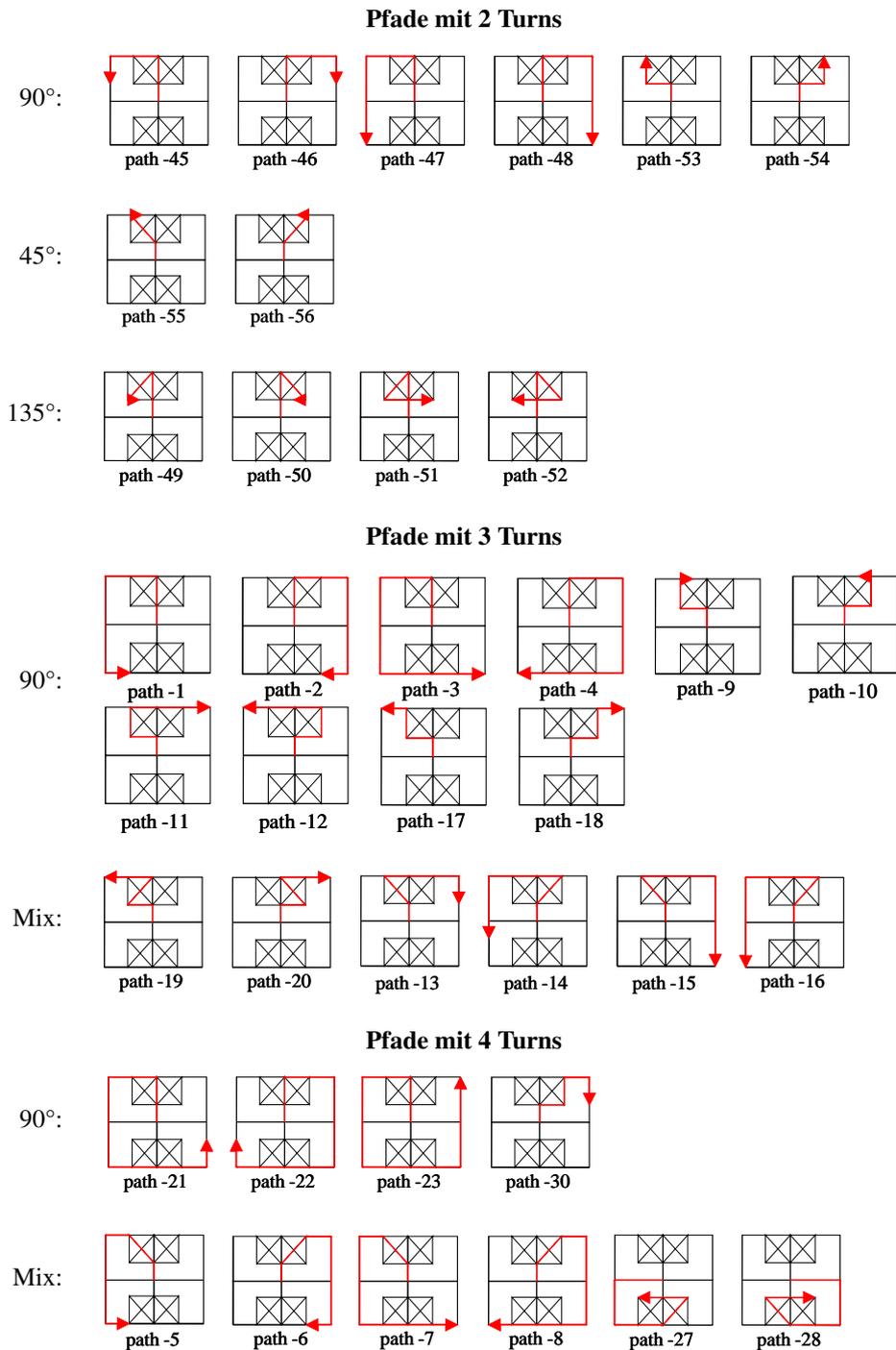
Abbildung 15.6: Beispiel der Konfigurationsdatei paths.cfg mit 6 Durchgängen (3 Pfade, einfache und doppelte Abspielgeschwindigkeit).

```
"path-45" 1
"path-18" 1
"path-49" 1
"path-45" 2
"path-18" 2
"path-49" 2
```

Berechnung neuer Pfade. Sollten die vorbereitenden Pfade für die Gestaltung eines Experiments nicht ausreichen, können auf einfache Weise eigenen Pfade generiert werden. Dazu steht das in der Programmiersprache Perl (The Perl Foundation¹) erstellte Skript C:\VR\env\VPT\paths\bin\paths.pl zur Verfügung.

¹www.perl.org

Abbildung 15.7: Vordefinierte Pfade für die Zeigeaufgabe im virtuellen Park. Der rote Pfeil markiert die Endposition und die Blickrichtung. Unter jedem Pfad steht der in der Konfigurationsdatei zu verwendete Name.

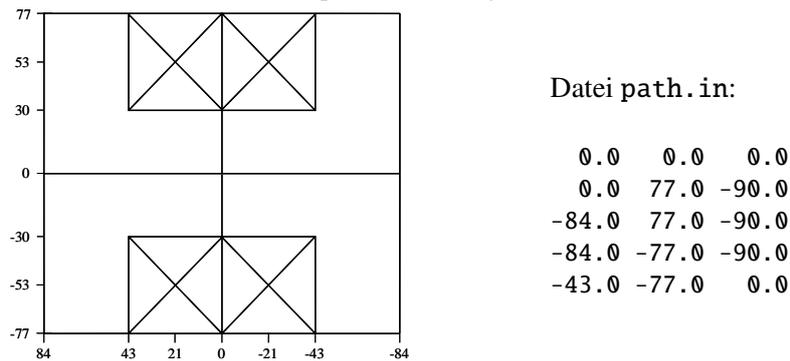


Aus einer Liste von Stützstellen (Koordinaten und Winkel), an denen eine Richtungsänderung in der virtuellen Umgebung stattfinden soll, interpoliert das Programm einen im Bewegungsverlauf geglätteten Pfad. Das Skript wird von der Kommandozeile wie folgt aufgerufen:

```
perl paths.pl path.in 0.25 >path-neu.pth.
```

h

Abbildung 15.8: Bestimmung der Stützstellen für die Interpolation weitere Pfade. (links) Grundriss mit x/y-Koordinaten. (rechts) Beispieldatei mit x, y, Drehwinkel



Der Parameter `path.in` ist der Name der Datei, die die Stützstellen enthält. Der Parameter `0.25` ist die Schrittweite für die Interpolation. Die Ausgabe des Programms ist eine Liste von Positions- und Rotationsdaten, die in die Datei `path-neu.pth` umgeleitet wird. Diese kann von der VR-Software als Pfad abgespielt werden.

Zur Bestimmung der Stützstellen werden die genauen Koordinaten der Weggabelungen innerhalb der virtuellen Parklandschaft benötigt. Im Grundriss in der Abbildung 15.8 können diese abgelesen und in eine Datei (z.B. `path.in`) zur Berechnung eingetragen werden. Das Format dieser Datei besteht aus den x/y-Koordinaten und dem Winkel der relativen Drehung an dieser Position. Alle drei Werte sind durch Leerzeichen voneinander getrennt.

15.3.2 Datenformat

Aus dem Experiment werden in einem Datensatz der tatsächliche Winkel, der geschätzte Winkel, die Differenz aus beiden geteilt durch 6 und die Latenz protokolliert. Es wird eine Protokolldatei pro Versuchsperson im Format einer EXCEL®-Tabelle angelegt (Tab. 15.5). Diese Werte werden für jeden Durchgang in eine einzelne Tabellenzeile geschrieben.

Tabelle 15.5: Format der Protokolldatei aus dem VPT.

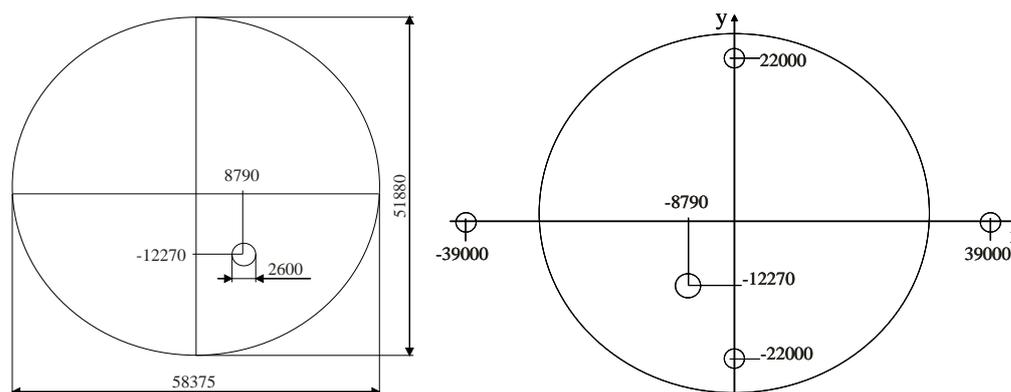
Pfad	Geschwindigkeit	Realer Winkel	Win- kel	Geschätzter Winkel	Schätz- fehler	Orientierungs- zeit [ms]
path-45	1	289		308	3	4500
path-18	1	137		117	-3	3530
...

15.4 Water-Maze-Aufgabe

15.4.1 Konfiguration

Der Grundriss der Insel für das virtuelle Training und für die virtuelle Water-Maze-Aufgabe sind identisch. Es unterscheiden sich lediglich die Positionen der Vertiefung in der die Ziellokation versteckt ist (s. Abb. 15.9).

Abbildung 15.9: Grundriss der VEs aus dem Training und der Water-Maze-Aufgabe.



Im Rahmen der Parametrisierung zur Softwareentwicklungen der virtuellen Umgebung der virtuellen Water-Maze-Aufgabe und dem virtuellen Training wurde eine einfache auf der Struktur der Microsoft Windows® INI-Dateien basierende Beschreibungssprache entwickelt. Damit können zahlreiche Version des virtuellen Water-Maze ohne Programmierkenntnisse erstellt werden.

Der Aufbau der Konfigurationsdateien für beide virtuelle Umgebungen ist im Prinzip gleich und gliedert sich in Sektionen und in bestimmte Schlüsselwörter. Eine Sektion beginnt mit einem Bezeichner in eckigen Klammern und endet mit dem Beginn der nächsten Sektion. Schlüsselwörter bezeichnen Eigenschaften von Objekte innerhalb der VE, die konfigurierbar sind.

Allgemeine Einstellungen. In Sektion [general] werden allgemeine Einstellungen zur Sensibilität des verwendeten Eingabegeräts vorgenommen. Außerdem kann ein allgemeiner Text formuliert werden, der bei Überschreitung des Zeitlimits in einem Durchgang erscheinen soll. Tabelle 15.6 fasst alle Schlüsselwörter dieser Sektion zusammen.

Tabelle 15.6: Schlüsselwörter der Sektion [general].

Schlüsselwort	Syntax	Bedeutung
Sensetivity	$1 \dots N$	Schrittweite Translation
AngularRate	$0.0001 \dots \pi$	Schrittweite Drehung
TimeLimitText	Text	Hinweistext bei Überschreitung des Zeitlimits
Text	Text	Hinweistext nach Abschluss des letzten Durchgangs

Virtuelles Training. Für die Konfiguration des virtuellen Trainings gibt es nur noch eine weitere mögliche Sektion [1]. Hier können die Positionen der Kisten, die Startposition, das Zeitlimit, eine Bezeichnung und der Instruktionstext für die Versuchsbedingung festgelegt werden (Tab. 15.7).

Die aktuelle Parameter Datei, die im virtuellen Training verwendet wurde befindet sich unter `C:\VR\env\VT\training.ini`. Die in den Kisten platzierten Objekte können beliebig aus folgender Liste ausgewählt werden und in der Konfigurationsdatei mit der entsprechenden Nummer eingetragen werden: (1) Mobiltelefon, (2) Feuerzeug, (3) Karte, (4) Schnorchel, (5) Taucherflossen, (6) Cola-dose und (7) Kaffeebecher.

Tabelle 15.7: Schlüsselwörter der Sektionen im VT.

Schlüsselwort	Syntax	Bedeutung
Time	Zeit	Zeitlimit in Sekunden
Start	x, y, Richtung	Startposition und initiale Blickrichtung
Crate-1..7	x, y, Richtung, Inhalt	Position und Ausrichtung der Kiste, Inhalt: Objekte 1 bis 7
Name	Text	Bezeichnung des Durchgangs
Text	Text	Hinweistext bei Beginn eines Durchgangs

Water-Maze-Aufgabe. Im Water-Maze-Experiment folgen mehrere Durchgänge direkt aufeinander. Die Konfiguration eines Durchgangs erfolgt jeweils in einer Sektion. Die Bezeichner der Sektion (in eckigen Klammern) dürfen nur aus aufeinander folgenden Zahlen von 1 bis zur maximalen Anzahl der geplanten Durchgänge bestehen. Die Sektion [1] steht also für den ersten Durchgang, die Sektion [2] für den zweiten usw. Damit ist das Water-Maze-Experiment um Versuchsbedingungen beliebig erweiterbar. Die in Tabelle 15.8 angegebenen Schlüsselwörter und Parameter sind zur Konfiguration des Experiments verfügbar.

Tabelle 15.8: Schlüsselwörter der Sektionen im VWM.

Schlüsselwort	Syntax	Bedeutung
Time	Zeit	Zeitlimit in Sekunden, 0: ohne Zeitlimit
Range	Radius	Erkundungsradius um das Ziel nach Erreichen
Start	x, y, Richtung	Startposition und initiale Blickrichtung
Land-1..4	x, y, Richtung	Position und Ausrichtung einer Landmarke
Prox-1..3	x, y, Richtung	Position und Ausrichtung eines proximalen Cues
TextDuration	Zeit	Dauer der Textanzeige in Sekunden
Name	Text	Bezeichnung des Durchgangs
Text	Text	Hinweistext bei Beginn eines Durchgangs

Die aktuelle Parameter Datei, die im virtuellen Water-Maze verwendet wurde befindet sich unter `\VR\env\VWM\mittel.ini`. Es können beliebig viele Konfigurationen vordefiniert werden, die

beim Start des Experiments über das Menü *Konfiguration* ausgewählt werden können. Die Schlüsselwörter für die Landmarken wurden wie folgt vergeben: (Land-1) Ballon, (Land-2) Leuchtturm, (Land-3) Segelyacht und (Land-4) Windrad. Die verfügbaren proximalen Cues (Prox-1 bis 3) sind drei identische blau-weiße Bälle. Eine Konfiguration eines Experiments mit 2 Durchgängen im VWM könnte beispielsweise wie in Abbildung 15.10 aussehen.

Abbildung 15.10: Beispielkonfiguration einer Water-Maze-Aufgabe.

```
[general]
TimeLimitText = Sie haben das Zeitlimit überschritten ...
Text = Sie haben alle Aufgaben erfüllt ...
Sensitivity = 10
AngularRate = 0.005
```

```
[1]
Name = Lernen-2
Time = 0
Range = 5000
Start = 18900, 18900, -135
Prox-1 = 13000, 11000, 0
Land-1 = 39000, 0, 0
Text = Sie sind auf einer Insel gestrandet...
TextDuration = 30
```

```
[2]
Name = Lernen-3
Time = 240
Range = 500
Start = 18900, -18900, -45
```

15.4.2 Datenformat

Aus dem VT und dem VWM werden hauptsächlich Rohdaten protokolliert, aus denen sich der exakte räumliche und zeitliche Verlauf der Bewegungen der Versuchsperson in der virtuellen Umgebung rekonstruieren lässt. Dazu wird von jedem Durchgang eine ASCII-Dateien unter dem Namen `<VP-Code><Nr. des Durchgangs>.prc` angelegt. Diese Datei enthält eine Tabelle, die bei jeder Positionsänderung der Versuchsperson um eine Zeile verlängert wurde. Die Zeilen bestehen aus Tupeln der Form:

$$(t_i, x_i, y_i, d_i, l_i), \quad 1 \leq i \leq n, n \in \mathbb{N}.$$

t_i ist die aktuelle Zeitangabe seit Start des Durchgangs (Zeitstempel), x_i und y_i sind die augenblickliche Position, d_i der momentane Abstand zur Ziellokation und l_i die kumulierte Pfadlänge bis zum Zeitpunkt t_i . Die Dateien aus dem VWM mit diesem Format können mit dem *Dataviewer* betrachtet und weiterverarbeitet werden.

Bereits während des Experimentes werden die Merkmale *Pfadlänge* und *Navigationszeit* aus den Rohdaten aggregiert und in einer weiteren Datei über alle Durchgänge hinweg zusammengefasst. Im VT hat diese Datei das in Tabelle 15.9 dargestellte Tabellenformat.

Tabelle 15.9: Format der Protokolldatei aus dem VT.

Nr	Bedingung	Lokation	Zeit[sec]	Pfadlänge[px]
01	Training	Crate-1	35	5017
01	Training	Crate-2	135	4321
.

Aus dieser Datei kann für weitere statistische Auswertungen direkt die Zeit und der zurückgelegte Weg zwischen dem öffnen zweier Kisten abgelesen werden. Diese Datei befindet sich im Verzeichnis C:\VR\env\VT unter dem Namen <VP-Code>.txt. Im VWM hat diese Datei das in Tabelle 15.10 dargestellte Format.

Tabelle 15.10: Format der Protokolldatei aus dem VWM.

Nr	Bedingung	Zeit[sec]	Pfadlänge[px]
01	Lernen-2	1955	801805
02	Lernen-3	145	74628
.

Aus dieser Datei sind unter den Spalten Nr und Bedingung die Nummer und die Bezeichnung des jeweilige Durchgangs protokolliert. Es können Latenz und Pfadlänge aus jedem Durchgang abgelesen werden. Die Datei befindet sich im Verzeichnis C:\VR\env\VWM unter dem Namen <VP-Code>.txt.

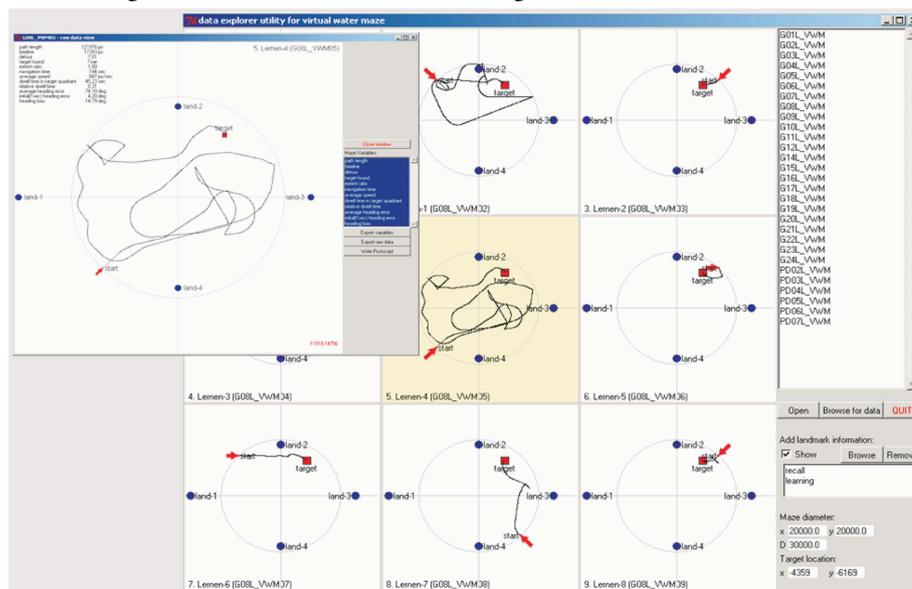
15.4.3 Dataviewer

Obwohl Pfadlänge und Latenz als wesentliche Merkmale der VWM-Leistung bereits während des Experiments berechnet werden, sind für eine bessere Charakterisierung der Verhaltensantwort im VWM oft weitere Variablen zu extrahieren. Außerdem ist es sinnvoll die von den Versuchspersonen zurückgelegten Wege nach einem Experiment erst einmal zu Visualisieren. Dazu steht das eigens dafür entwickelte Hilfsprogramm *Dataviewer* zur Verfügung, dessen Funktionsweise kurz beschrieben wird.

Datensatz öffnen. Üblicherweise befinden sich alle Datensätze aus einem Experiment im Verzeichnis C:\VR\env\VWM\DATA. Nach der Installation der VR-Software ist dieses Verzeichnis auch für den *Dataviewer* als Arbeitsverzeichnis eingestellt. Auf der rechten Seite des Programmfensters (s. Abb. 15.11) erscheint deshalb eine Liste mit allen in diesem Verzeichnis vorhandenen Datensätzen. Sollte das nicht der Fall sein, muss mit dem Button **Browse for data** zunächst das richtige Arbeitsverzeichnis angegeben werden.

Ein Datensatz wird geöffnet indem man mit einem Doppelklick den entsprechenden Eintrag in der Liste anklickt. Im linken größten Bereich des Fensters werden alle Durchgänge aus dem VWM, die

Abbildung 15.11: Der *Dataviewer* mit einem geöffneten Datensatz aus dem VWM.



diese Versuchsperson absolviert hat dargestellt. Es werden jeweils der Grundriss des Maze mit einer Einteilung der 4 Quadranten, die Ziellokation, die Startposition und der zurückgelegte Weg dargestellt. Das Öffnen des gesamten Datensatzes kann u. U. längere Zeit in Anspruch nehmen, da die auszulesenden PRC-Dateien sehr lang sein können. Das gilt insbesondere für Explorationsdurchgänge, in denen die Versuchspersonen oft sehr lange Wege zurückgelegt haben.

Um zusätzliche Information über die Position von Landmarken und proximalen Hinweisreizen anzuzeigen, benötigt das Programm Angabe aus der im Experiment verwendeten Konfigurationsdatei, d. h. die INI-Dateien. Mit dem Button **Browse** kann diese Datei im System aufgesucht werden und in eine Liste von verfügbaren Konfigurationen eingefügt werden. Nach der Installation der VR-Software ist diese Angabe im *Dataviewer* automatisch eingetragen. Sollte dies nicht der Fall sein, befindet sich diese Datei im Verzeichnis C:\VR\env\VWM. Zur endgültigen Anzeige der Landmarkeninformation muss dann noch das Kontrollkästchen **Show** ausgewählt werden.

Im rechten unteren Bereich des Fensters befinden sich Angaben zu den Abmaßen des VWM und zur Position der Ziellokation. Diese sind beim Start des Programms auf Default-Werte eingestellt, die den Maßen des verwendeten Maze entsprechen und normalerweise nicht verändert werden müssen.

Nach dem ein Datensatz geöffnet und die Landmarkeninformation gelesen wurde können aus den einzelnen Durchgängen weitere Information extrahiert werden. Um mit einem einzelnen Durchgang weiter zu arbeiten, klick man einmal auf die entsprechende Abbildung. Es erscheint ein zweites großes Fenster in dem nur der ausgewählte Durchgang dargestellt wird. Auf der rechten Seite befindet sich eine Liste mit möglichen Merkmalen, die aus diesem Durchgang aggregiert werden können und drei Button für die folgenden Funktionen.

1. Merkmale exportieren. Unter dem Button *Export variables* können zuvor in der Liste ausgewählte Merkmale wie z. B. Pfadlängen oder Verweildauern in eine Textdatei im CSV-Format² geschrieben werden. Zunächst wählt man in der Liste die gewünschten Merkmale aus, deren Werte berechnet und im rechten oberen Fensterbereich dargestellt werden. Die Auswahl wird zwischengespeichert und beim Öffnen eines weiteren Durchgangs oder einem anderen Datensatz wieder angewendet. Das ermöglicht eine komfortable Aggregation der gleichen Variablen über mehrere Durchgänge und Versuchspersonen hinweg, ohne das Variablen vergessen werden.

Für die Exportdatei wählt man beim ersten Export ein Verzeichnis und einen Namen mit dem Button *Browse*. Diese Datei wird als erweiterbare Tabelle organisiert, in die sukzessive mehrere Versuchspersonen mit jeweils mehreren Durchgängen geschrieben werden können. Das Programm verwaltet dabei Variablen eines Durchgangs in benachbarten Spalten und fügt weitere Variablen automatisch in der richtigen Reihenfolge ein. Die Merkmale der Versuchspersonen liegen dann in separaten Zeilen vor, so dass die Tabelle direkt als Datenmatrix in SPSS[®] exportiert werden kann.

Zur Berechnung bestimmter Merkmale sind Informationen aus der Konfigurationsdatei notwendig. Findet das Programm diese Datei nicht erscheint eine Warnung, dass diese Werte nicht berechnet werden konnten. Es ist dann zunächst Landmarkeninformationen zu laden, wie es unter dem Punkt **Datensatz öffnen** beschrieben wurde. Die meisten Merkmale werden jedoch aus den Tupeln der Bewegungsprotokolle (PRC-Dateien) berechnet. In der Tabelle 15.11 ist eine Übersicht aller im *Dataviewer* implementierten Aggregationsmethoden und deren Berechnungsvorschriften zusammengestellt.

2. Daten exportieren. Mit dieser Funktion kann der gesamte Datensatz einer Versuchsperson in das Format der Analysesoftware *Wintrack* exportiert werden. Diese Software bietet weitere Möglichkeiten der Berechnung von Merkmale zur Charakterisierung des Bewegungsverhaltens im *Water maze*, wie sie in den Tierexperimenten üblich sind.

3. Postscript schreiben. Praktisch für Publikationen und Präsentation ist die Darstellung einzelner Bewegungsmuster aus dem VWM in Form einer Abbildung. Dazu können aus dem *Dataviewer* einzelne Durchgänge im Postscript-Format (Level 2) exportiert werden. Diese können leicht in Textdokumente eingebunden, nachbearbeitet, beliebig skaliert und in weitere gebräuchliche Formate konvertiert werden.

²Colon Separated Values: Einfaches Format für Tabellendokumente, dass mit Microsoft Excel und SPSS gelesen werden kann.

Tabelle 15.11: Im *Dataviewer* implementierte Merkmalsaggregationen aus den VWM-Bewegungsrohdaten.

Name	Abk.	Einheit	Berechnung
Pfadlänge	pthlen	px	$\sum_i \left\ \begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix} \right\ _2$
Luftlinie	bl	px	$\left\ \begin{pmatrix} x_{target} \\ y_{target} \end{pmatrix} - \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \right\ _2$
Latenz	nvtm	sec	$t_n - t_1$
Umweg	dt	–	Pfadlänge / Luftlinie
Durschn. Geschw.	avgs	px sec ⁻¹	Pfadlänge / Latenz
Ziel gefunden?	tf	–	1: Ziel gefunden, 0: sonst
Verweildauer im:			$\sum_i \Delta t_i$ mit:
I. Quadrant (Ziel)	dwtmq1	sec	$\Delta t_i = \{t_{i+1} - t_i: x_i < 0 \wedge y_i < 0, 0: \text{sonst}\}$
II. Quadrant	dwtmq2	sec	$\Delta t_i = \{t_{i+1} - t_i: x_i > 0 \wedge y_i < 0, 0: \text{sonst}\}$
III. Quadrant	dwtmq3	sec	$\Delta t_i = \{t_{i+1} - t_i: x_i > 0 \wedge y_i > 0, 0: \text{sonst}\}$
IV. Quadrant	dwtmq4	sec	$\Delta t_i = \{t_{i+1} - t_i: x_i < 0 \wedge y_i > 0, 0: \text{sonst}\}$
Durschn. Heading Error	avghderr	Grad	$\sum_i \left \angle \left[\begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix}, \begin{pmatrix} x_{target} - x_i \\ y_{target} - y_i \end{pmatrix} \right] \right $
Heading Bias	hdbias	Grad	$\sum_i \angle \left[\begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix}, \begin{pmatrix} x_{target} - x_i \\ y_{target} - y_i \end{pmatrix} \right]$
Initialer Heading Error nach 20 sec	inihderr	Grad	$\left \angle \left[\begin{pmatrix} x_i - x_1 \\ y_i - y_1 \end{pmatrix}, \begin{pmatrix} x_{target} - x_1 \\ y_{target} - y_1 \end{pmatrix} \right] \right $ mit $1 < i \leq n$, so dass $t_i - t_1 \geq 20$ sec
Anzahl Wendepunkte	ntps	–	$\begin{cases} \text{ntps} + 1 & \left \angle \left[\begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix}, \begin{pmatrix} x_i - x_{i-1} \\ y_i - y_{i-1} \end{pmatrix} \right] \right > 90^\circ \\ 0 & \text{sonst} \end{cases}$

16 Technische Details - Datenanalyse

16.1 Matlab-Skripte

16.1.1 Nicht-negative Matrix-Faktorisierung

```
%snmfTikhonov implement the least square Tikhonov regularization method,  
%input: A, the nonnegative matrix to decompose; k, the dimension desired;  
%lambda, control parameter for sparseness, larger lambda results in sparse  
%H and thus W. N the maximum allowed iteration. N, a few hundred seems to  
%be enough.  
%output: W and H, all positive.  
%Author: Yuan Gao, Ph.D.  
%Department of Genetics, Harvard Medical School, Boston, MA 02115  
%ygao@receptor.med.harvard.edu  
%2004-2005  
  
function [W, H] = snmfTikhonov(A, k, lambda, N)  
[m,n] = size(A);  
epsilon= 10^-10;  
WK = power(lambda,2)*eye(k);  
W = rand(m, k);  
KK=zeros(k,n);  
AK = [A;KK];  
H = zeros(k, n);  
for i = 1 : N  
    WWK =[W;WK];  
    for j = 1 : n %solve the reg. least square problem  
        H(:,j) = WWK\AK(:,j);  
    end  
    for i2=1:k %nonnegative constraint  
        for j2=1:n  
            if (H(i2,j2)<0)  
                H(i2,j2)=0;  
            end  
        end  
    end  
    W = W .* (A * H') ./ (W * (H * H') + epsilon); %Euklidian update rule  
end  
delta = norm(A-W*H);
```

16.1.2 Consensus-Clustering

```

% NMF-Anwendung auf Analyse von Verhaltensdaten
% Quelle: Brunet et al. 2004, Monti et al. 2003

%NMF parameter
lambda = 0.1;
N = 800;
kMin=2;
RUNS = 20;

%read datasets as a binary nxm matrix V
[n,m]=size(V);
kMax = ceil(n*m/(n+m));
if kMax > 10
    kMax=10;
end

dist = zeros(kMax,1);

for k=kMin:kMax
    C=zeros(m); % connectivity matrix
    for iter=1:RUNS
        [W,H] = snmfTikhonov ( V, k, lambda, N); %sparse NMF algorithm
        [y,c]=max(H);
        C=C+(repmat(c,m,1)==repmat(c',1,m));
    end
    C_=C/RUNS; %consensus matrix

    uvec=C_(1,2:end);
    for i=2:m-1;
        uvec=[uvec C_(i,i+1:end)]; %get upper diagonal elements of consensus
    end

    y=1-uvec;          % consensus are similarities, convert to distances
    z=linkage(y,'average'); % use average linkage
    coph(k)=cophenet(z,y);
    clust{k}=cluster(z,k); % get cluster id

    fig = figure('visible','off'); % turn off dendrogram plot
    [h,t,ord]=dendrogram(z,0); % get permutation vector
    close(fig)

    Consensus{k}=C_(ord,ord); % reorder consensus matrix for proper plots
end

```

16.1.3 Vorverarbeitung der Spurdaten der Water-Maze-Aufgabe

```

%load a single rawdate file (track date from VWM)
%built a single datavector
%author: R.Schoenfeld
%2006

function [B]=loadsinglefile (szFilename, mazeSize, scaleSize)

    Data = load (szFilename);

    %get correct coordinates from data vector (timestamp,x,z)
    Pos = [Data(:,3) -Data(:,2)];
    Time = [diff(Data(:,1)); 0];

    mazeSize = [mazeSize(2) mazeSize(1)];
    D = mazeSize/scaleSize;

    %shift to positive coordinates and down scale data
    Index = fix((Pos+repmat(mazeSize/2, size(Pos,1), 1))/D)+1;
    B = zeros(scaleSize);

    %build binary matrix B
    for i=1:size(Index,1)
        B(Index(i,1),Index(i,2)) = B(Index(i,1),Index(i,2)) + Time(i);
    end

    B(find(B>0))=1;

%use function above for all data files zo construct the full datamatrix V
%downscale raw date to a 31x31 binary matrix
%author: R.Schoenfeld
%2006

function [V]=loadfiles(szListName, szCondition, scaleSize)

    fList = importdata (szListName);

    for i=1:length(fList)

        B=loadsinglefile ([fList{i} szCondition '.prc'], [54000 60500], [31 31]);

        V(:,i)=reshape (B, prod(size(B)), 1);
    end
end

```

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation mit dem Thema

„Alters- und Geschlechtsdifferenzen in der Raumkognition - auch eine Frage der Lösungsstrategien“

selbstständig verfasst, nicht schon als Dissertation verwendet sowie die benutzten Hilfsmittel und Quellen vollständig angegeben habe.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr.rer.nat.) an anderen Einrichtungen eingereicht habe.

Halle (Saale), _____
(Datum)

(Unterschrift)

Lebenslauf

Name: Robby Schönfeld
Geburtstag: 03. Juli 1973
Geburtsort: Hohenmölsen, Deutschland (Sachsen-Anhalt)
Anschrift: Streiberstraße 2
06110 Halle (Saale)
Familienstand: verheiratet
Kinder: 2

Ausbildung und Berufserfahrung

10/1993 –09/1998 Diplom-Studiengang Informatik
Martin-Luther-Universität Halle-Wittenberg
Abschluss: Dipl.-Informatiker

10/1998 –03/2001 wissenschaftlicher Mitarbeiter
Institut für Informatik, Martin-Luther-Universität Halle-Wittenberg
DFG-Projekt „Genetische Algorithmen im digitalen Entwurf Integrierter Schaltkreise“

04/2001 –01/2003 wissenschaftlicher Mitarbeiter
Institut für Psychologie, Martin-Luther-Universität Halle-Wittenberg
DFG-Schwerpunktsprojekt „Spatial Cognition“

06/2003 –12/2004 wissenschaftlicher Mitarbeiter
Klinik für Diagnostische Radiologie, Universitätsklinikum der Martin-Luther-Universität Halle-Wittenberg
Projekt (Haushalt): Prä-/postoperative Hirnkartierung mit fMRT in der computerassistierten Neurochirurgie

11/2005 –08/2008 Kooperatives Promotionsverfahren
Institut für Psychologie, Martin-Luther-Universität Halle-Wittenberg
Institut für Psychologie I, Otto-von-Guericke-Universität Magdeburg

11/2008 wissenschaftlicher Mitarbeiter
Institut für Psychologie, Martin-Luther-Universität Halle-Wittenberg

Verzeichnis wissenschaftlicher Veröffentlichungen

Peer-reviewed Artikel

Schönfeld, R., Mönich, N., Müller, F.J., Lehmann, W., & Leplow, B. (in review). Search strategies in a human water maze analogue analyzed with a parts-based approach. *Behavioral Brain Research*.

Schönfeld, R., Lehmann, W., & Leplow, B. (in review) Specificity of age and sex differences in tasks of spatial navigation and mental rotation. A special issue of the *Journal of Individuell Differences*.

Höll, D., Leplow, B., Schönfeld, R., & Mehdorn, M. (2003). Is It Possible to Learn and Transfer Spatial Information from Virtual to Real Worlds? In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition III, Lecture Notes in Computer Science*, (pp. 143-156). Berlin: Springer-Verlag.

Schönfeld, R. & Molitor, P. (2002). What are the samples for learning efficient routing heuristics? [MCM routing]. In *Circuits and Systems, 2002 Asia-Pacific Conference on Circuits and Systems (APCCS'02)*, (pp. 267-272). New York: Wiley-IEEE Press.

Günther, W., Schönfeld, R., Becker, B., & Molitor, P. (2001). k-layer Straightline Crossing Minimization by Speeding Up Sifting. In J. Marks (Ed.), *Graph Drawing, 8th International Symposium, GD 2000 Colonial Williamsburg, VA, USA. Lecture Notes In Computer Science*, (pp. 253-258). Berlin: Springer-Verlag.

Matuszewski, C., Schönfeld, R., & Molitor, P. (1999). Using Sifting for k-Layer Straightline Crossing Minimization. In J. Kratochvil (Ed.), *Graph Drawing, 7th International Symposium, GD'99 Štirín Castle, Czech Republic. Lecture Notes In Computer Science*, (pp. 217-224). Berlin: Springer-Verlag.

Publizierte Abstracts

Schönfeld, R. & Leplow, B. (2008). Sex-specific solution strategies in spatial tasks. In C. Dalbert (Ed.), *XXIX International Congress of Psychology: Abstracts. Special Issues of the International Journal of Psychology*,. Hove, UK: Psychology Press.

Schönfeld, R., Drauschke, A., Röpke, F., & Leplow, B. (2007). Domain specific enhancements of spatial cognition by unphysiological estrogens and progesterone levels. In E. Wascher, M. Falkenstein, G. Rinkenauer, & M. Grosjean (Hrsg.), *Psychologie und Gehirn 2007, Tagung der Deutschen Gesellschaft für Psychophysiologie und ihre Anwendung*, (p.10). <http://hdl.handle.net/2003/24421>.

Schönfeld, R., Lehmann, W., & Leplow, B. (2006). Pfadintegration in virtuellen Umgebungen. In F. Lösel & D. Bender (Hrsg.), *45. Kongress der Deutschen Gesellschaft für Psychologie*, (pp. 333). Oberhaching-München: Dusterl-Verlag.

Berndt, U., Schönfeld, R., John, R., Thomssen, C., & Leplow, B. (2007). Kognitive Auswirkungen

der systemischen Therapie bei Mammakarzinompatientinnen. In B. Dümde (Hrsg.), *MedReview, 1. Jahrestagung der Mitteldeutschen Gesellschaft für Frauenheilkunde und Geburtshilfe*, (S. 13-14). Berlin: Blackwell Verlag

Leonhardt, G., Schönfeld, R., Knörger, M., Schillbach, H., Spielmann, R., & Lepow, B. (2005). Zerebrale Aktivierung beim Konditional Assoziativen Lernen. In *Aktuelle Neurologie, 78. Kongress der Deutschen Gesellschaft für Neurologie*, (S. 321). Stuttgart: Georg Thieme Verlag.

Beiträge auf wissenschaftlichen Fachtagungen

Schönfeld, R., Lehmann, W., Ungewiß, J., & Lepow, B. (2007). Leistungsdifferenzen in der Computerversion des Mental Rotations Test. Vortrag auf der 18. Tagung der Fachgruppe Entwicklungspsychologie vom 24.-27. September, Heidelberg.

Technische Berichte

Schmiedle, F., Matuszewski, C., & Schönfeld, R. (1999). Channel routing with genetic algorithms. Report No. 125, Institut für Informatik, Universität Freiburg.