



universität
wien

Diplomarbeit

SPACE 2d/3d:

Entwicklung eines Rasch-homogenen
Raumvorstellungstests

Verfasserin

Christiane Heider, bakk.

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag^a. rer. nat.)

Wien, im Oktober 2010

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: Ao. Univ.-Prof. Dr. Georg Gittler

*,,Wissenschaft ist wie Sex.
Manchmal kommt etwas Sinnvolles dabei raus,
das ist aber nicht der Grund,
warum wir es tun.“*

(Richard P. Feynman)

Danksagung

Jeder, der in seinem Leben bereits eine Diplomarbeit geschrieben hat, weiß, dass dies nicht ohne die Hilfe von anderen zu bewerkstelligen ist. Aus diesem Grunde möchte ich mich an dieser Stelle bei einigen Menschen, die maßgeblich an der Entstehung dieser Arbeit beteiligt waren bedanken.

Mein Dank gilt in erster Linie Ao.-Univ. Prof. Dr. Gittler für die gute Betreuung, sowohl bei der Erstellung des, in dieser Diplomarbeit erstmals zum Einsatz gekommenen, Tests, als auch bei der anschließenden Auswertung der Daten.

Mein besonderer Dank ergeht auch an Frau Mag^a. Eva-Maria Adlmann, die mir stets mit Rat und Tat zur Seite stand, wann immer ich diesen benötigt habe.

Diese Arbeit wäre jedoch ohne die finanzielle Unterstützung meiner Eltern, Dr. Manfred und Dr. Ingeborg Heider, niemals zustande gekommen, da sie es mir ermöglicht haben, trotz einiger Bedenken, das Studium meiner Wahl zu absolvieren, weshalb ich auch ihnen zu Dank verpflichtet bin.

Weiters möchte ich mich auch bei meinem Bruder Michael dafür bedanken, dass er mir in schwierigen Phasen zur Seite gestanden hat und mir Mut zugesprochen hat, wann immer es notwendig war.

Nicht vergessen möchte ich an dieser Stelle meinen Freunden Andrea, Anna, Marie-Therese und Tadeh zu danken, die mich in stressigen Situationen immer wieder in die Fremde entführt haben, was es mir ermöglichte anschließend wieder mit neuem Elan an die Arbeit zu gehen.

Zuletzt möchte ich auch den zahlreichen TestteilnehmerInnen meinen Dank aussprechen, da diese Arbeit ohne sie nicht möglich gewesen wäre.

Danke

Inhaltsverzeichnis

1	EINLEITUNG.....	9
THEORETISCHER TEIL.....		11
2	RAUMVORSTELLUNG.....	13
2.1	DEFINITION VON RÄUMLICHEM VORSTELLUNGSVERMÖGEN	13
2.2	GESCHICHTLICHE ENTWICKLUNG VON RAUMVORSTELLUNG	14
3	DIFFERENTIALPSYCHOLOGISCHE BEFUNDE ZUM KONZEPT „MENTALE ROTATION“	18
3.1	NEUROPSYCHOLOGIE.....	19
3.2	HORMONELLER EINFLUSS.....	22
3.2.1	NATÜRLICHE FLUKTUATION IM ÖSTROGENHAUSHALT	23
3.2.2	NATÜRLICHE FLUKTUATION IM ANDROGENHAUSHALT.....	23
3.2.3	EXPERIMENTE ZUR HORMONHYPOTHESE.....	24
3.3	GESCHLECHTSROLLENIDENTITÄT.....	25
3.4	SOZIALISATIONSEFFEKTE.....	27
3.5	HOLISTISCHE VERSUS ANALYTISCHE BEARBEITUNGSSTRATEGIEN.....	28
3.6	PSYCHOBIOSOZIALES MODELL BEI KOGNITIVEN GESCHLECHTSUNTERSCHIEDEN ...	29
4	DER MENTAL ROTATION TEST (MRT)	31
4.1	VON DEN ANFÄNGEN DES MENTAL ROTATION TESTS	31
4.2	ITEMANALYSE.....	32
4.3	MODIFIZIERUNG DES MRT FÜR DIESE ARBEIT	34
EMPIRISCHER TEIL		39
5	DATENERHEBUNG	41
5.1	DURCHFÜHRUNG	41
5.2	VERSUCHSPLAN	43
5.3	BESCHREIBUNG DER STICHPROBE	45
6	ZIELSETZUNG UND FORSCHUNGSFRAGEN	47
6.1	ZIELSETZUNG	47
6.2	FORSCHUNGSFRAGEN	47
7	ERGEBNISSE.....	49
7.1	MODELLTEST NACH RASCH	49

7.2 AUSWERTUNG MITTELS KLASSISCHER TESTTHEORIE.....	54
8 ZUSAMMENFASSUNG UND KRITIK.....	63
8.1 ZUSAMMENFASSUNG.....	63
8.2 KRITIK UND ANREGUNG FÜR KÜNFTIGE FORSCHUNG.....	64
9 ABSTRACT – DEUTSCH.....	66
10 ABSTRACT – ENGLISH.....	67
11 LITERATURVERZEICHNIS	68
12 ABBILDUNGSVERZEICHNIS	74
13 TABELLENVERZEICHNIS	75
14 ANHANG	76
14.1 BILDERDOKUMENTATION DES „SPACE 2D/3D“ VORLAUFS.....	76
14.2 RASCH-MODELL-TEST ÜBER BEIDE VERSIONEN.....	85
14.3 MARTIN-LÖF – TEST FÜR DIE 2D UND 3D PRÄSENTIERTEN AUFGABEN	86
14.4 WLE-PARAMETER NACH WARM (1989)	87
14.5 CURRICULUM VITAE	88

1 Einleitung

Wofür benötigt man im alltäglichen Leben die Fähigkeit, Gegenstände im Gedanken zu manipulieren? Dies ist eine berechtigte Frage, zu Beginn einer wissenschaftlichen Arbeit, die sich genau diesem Thema widmet. Heutzutage gibt es eine Reihe von technischen Berufen, beziehungsweise von Aufnahmetests an diversen Hochschulen (z.B. Medizinische Universität Wien) bei denen diese Fähigkeit eine Voraussetzung für berufliches Weiterkommen darstellt. Da es bewiesenermaßen den Männern leichter fällt sich im Raum zu orientieren, stellen diese Aufnahmebedingungen eine Chancengleichheit gegenüber Frauen dar. Die zentrale Hypothese, die dieser Arbeit nun zugrunde liegt, ist jene, dass Frauen im Prinzip genauso gut in ihrer mentalen Rotationsfähigkeit sein könnten, es für sie aber möglicher Weise schwieriger ist, Gegenstände die zweidimensional präsentiert werden, gedanklich zu einem dreidimensionalen Objekt zu transformieren, dieses mental zu manipulieren, um in einem letzten Schritt das Endergebnis wiederum in eine zweidimensionale Antwort umzuwandeln.

Um diese erschwerenden gedanklichen Transformationsschritte zu verkürzen wurde deshalb ein Test entwickelt, der Aufgaben enthält, die bereits 3-dimensional präsentiert werden. Ob dieser Test die an ihn gestellten Anforderungen erfüllen kann und ob der vermutete erleichternde Effekt für Frauen tatsächlich eintritt, soll im Rahmen dieser Arbeit geklärt werden.

Im theoretischen Teil der Arbeit wird anfänglich auf die Geschichte der Testung der Fähigkeitsdimension „Raumvorstellung“ eingegangen, um anschließend die differentialpsychologischen Aspekte bezüglich des Geschlechtsunterschiedes speziell auf dem Gebiet der Mentalen Rotationsfähigkeit näher zu beleuchten. Neuropsychologische Befunde sollen hierbei genauso zur Aufklärung beitragen wie einige Ergebnisse zur Geschlechtsrollenidentität und zu Sozialisationseffekten.

Im vierten Kapitel wird der interessierte Leser schließlich über die Entstehungsgeschichte des ersten Mental Rotation Tests (MRT) nach Shepard und Metzler (1971) informiert, der seine Weiterentwicklung Vandenberg und Kuse (1978) verdankte. Danach

wird daran die neue Testkonstruktion skizziert, die auf einige, bisher wenig beachtete, Probleme klassischer Aufgabenstellungen, eingeht.

Im empirischen Teil dieser Arbeit, soll geklärt werden, ob das Ziel – die Neukonstruktion eines Rasch-homogenen Tests – erfüllt wurde. Sollte dies der Fall sein, bedeutet dies, dass alle konstruierten Aufgaben dieselbe latente Dimension erfassen. Zusätzlich wird der Frage nachgegangen, ob die unterschiedlichen Darstellungsmodi in „traditionell zweidimensional“ oder „stereoskopisch dreidimensional“ zu unterschiedlichen Leistungen bzw. bei 3d zu einem Schrumpfen des Geschlechtsunterschiedes führen.

Abschließend werden die wichtigsten Ergebnisse in einem eigenen Kapitel nochmals zusammengefasst und ein Ausblick auf zukünftige Forschung, die an die vorliegenden Resultate anknüpfen könnte, gegeben.

Es wird der besseren Lesbarkeit wegen bei allgemeinen Aussagen nur die männliche Form des Substantivs verwendet. Bei der Wahl der männlichen Form ist zugleich auch immer die weibliche Darstellungsform gemeint.

THEORETISCHER TEIL

2 Raumvorstellung

2.1 *Definition von räumlichem Vorstellungsvermögen*

Der Terminus räumliches Vorstellungsvermögen wird trotz seiner längeren wissenschaftlichen Erforschung immer noch aus vielen unterschiedlichen Perspektiven betrachtet, und selbst bei den untersuchenden Forschern herrscht bei der Definition von „spatial ability“ immer noch Uneinigkeit. Treffend formulierten Eliot und Smith folgende Aussage bezüglich der räumlichen Fähigkeit

„spatial ability has been defined in such a variety of different ways that it is often difficult to be precise about the meanings which we ascribe to the term“ (Eliot & Smith, 1983:1).

Aus diesem Grund soll an dieser Stelle ein kurzer Überblick über diverse Definitions- vorschläge einzelner Autoren skizziert werden, um daraus eine Arbeitsdefinition für die vorliegende wissenschaftliche Abhandlung festlegen zu können.

Caplan, MacPherson und Tobin (1985) machten darauf aufmerksam, dass es bei der Definition von räumlichem Vorstellungsvermögen wichtig sei, sich in einem ersten Schritt darüber klar zu werden, ob es sich bei „spatial ability“ um einen einzigen Faktor oder um mehrere Komponenten, die unterschiedliche Aspekte von Raumvorstellung erfassen, handelt.

Lohman, der sich bereits 1979 mit dieser Problematik beschäftigte, war darum bemüht eine klare Definition dieser kognitiven Fähigkeit zu finden. Er wählte einen eher allgemeinen und weitgegriffenen Erklärungsansatz des Konstrukts Raumvorstellung.

„Spatial ability may be defined as the ability to generate, retain, and manipulate abstract visual images. At the most basic level, spatial thinking requires the ability to encode, remember, transform, and match spatial stimuli.“ (Lohman, 1979 zit. n. Carroll, 1993:305).

Er unterteilte Raumvorstellung in folgende drei Komponenten:

1. *Spatial Relation*: Findet ihre Anwendung in Aufgaben bei denen Objekte mental rotiert werden müssen.

2. *Spatial Orientation*: Stellt die Fähigkeit dar, sich vorstellen zu können wie ein Objekt aus einer anderen Perspektive aussehen würde. Als problematisch erweist sich diese Komponente jedoch dadurch, dass sie auch durch mentale Rotation gelöst werden kann, ohne sich selbst im Raum neu auszurichten.
3. *Visualisation*: Hierbei handelt es sich um eine komplexe Komponente, bei der man sich Objekte vorstellen muss, wie sie z.B. auseinandergekommen aussehen würden (vgl. Paper-Folding-Test).

Linn und Petersen kamen 1985 im Zuge ihrer Meta-Analyse zu einer ähnlichen Unterteilung des menschlichen Raumvorstellungsvermögens, wobei sie lediglich die Faktoren umbenannten (*Spatial Perception*, *Mental Rotation*, *Spatial Visualisation*). Voyer, Voyer und Bryden (1995) kritisierten diesen Ansatz, da er ihnen als zu unreflektiert und zu simplifiziert vorkam, obwohl sie sich der Meinung anschlossen, dass es sich bei „spatial ability“ nicht um einen einzigen Faktor handeln würde, sondern um ein komplexeres Konstrukt, das einer weiteren Spezifizierung bedürfe.

Dieser Arbeit liegt nun die Hypothese zu Grunde, dass es sich bei Raumvorstellung, ähnlich wie bereits Lohman (1979) feststellte, um einen Faktor handelt, der sich in mindestens drei weitere Unterkategorien unterteilen lässt, wobei eine dieser Kategorien die mentale Rotation ist.

Mentale Rotation wird wiederum im *Wörterbuch der Psychologie* als eine „experimentelle Vorgehensweise,“ definiert „um mentale Prozesse bzw. Bilder und mentale Modelle zu analysieren“ (Fröhlich, 2005:318), während Voyer, Voyer und Bryden (1995) Mentale Rotation als die Fähigkeit verstanden, zwei- und dreidimensionale Gegenstände schnell und richtig in Gedanken drehen zu können.

2.2 Geschichtliche Entwicklung von Raumvorstellung

Geschichtlich betrachtet lassen sich drei unterschiedliche Abschnitte in Bezug auf die Erforschung von Raumvorstellung unterscheiden. Zum einen gab es die philosophische Periode, die als „vorwissenschaftlich“ bezeichnet werden kann, gefolgt von der Quantifizierungsperiode („measurement period“), die wiederum von der kognitiven und neuropsychologischen Periode abgelöst wurde.

Während der philosophischen Periode gingen die wichtigsten griechischen Philosophen davon aus, dass Raumvorstellung ein wesentlicher Bestandteil der menschlichen Psyche und von Gedanken an sich sei. Zu den namhaftesten Vertretern gehören zweifelsohne Aristoteles und Platon. Eine ähnliche Ansicht vertraten im 16. und 17. auch die englischen Empiristen John Locke, David Hume und George Berkeley.

Die quantitative Erfassung von Raumvorstellung und damit auch die Abgrenzung zur vorwissenschaftlichen Periode, geht maßgeblich auf Sir Francis Galton [1822 – 1911] zurück. Er erkannte als Erster, dass objektives Messen nur dann möglich sei, wenn er Vergleichswerte von mehreren Personen vorliegen habe, weshalb er beispielsweise eine Umfrage in Bezug auf das Erinnerungsvermögen an das morgendliche Frühstück durchführte. Die Befragten mussten sich dabei möglichst genau an den Frühstückstisch und an alles was sich auf ihm befand, erinnern. Auf diese Art und Weise entwickelte er ein Verfahren mit dem er Geschlechts-, Alters- und andere individuelle Unterschiede messen konnte. Allerdings scheiterten seine Versuche, Intelligenz mit psychometrischen Verfahren zu erheben, da er hierfür, bis auf die Reaktionszeit, nur ungeeignete Maße (Körpergröße, Sehschärfe und Körperumfang) überprüfte (Carroll 1993:31).

Angeregt durch Sir Francis Galton, suchten die Wissenschaftler in der Zeit von 1904 – 1938 nach Beweisen für bzw. gegen die Existenz eines räumlichen Faktors (Eliot & Smith, 1983). Wichtigen Einfluss nahm in diesem Zusammenhang die im Jahre 1904 veröffentlichte Generalfaktortheorie von Charles E. Spearman [1863 – 1945], in der er postulierte, dass jeder Intelligenztest im Wesentlichen auf zwei Faktoren beruhe, nämlich einen für „general intelligence“ (g) und einen für jeden Test spezifischen (s). Der g-Faktor liegt laut Spearman allen Tests zugrunde, während der Faktor „s“ spezielle Fähigkeiten (bsp. mentale Rotation) erfasst. Damit kognitive Inhalte überhaupt zustande kommen können, bedarf es nach Spearman (1927) drei kognitiver Fähigkeiten: erstens die Fähigkeit, sich Erfahrungen aneignen zu können („apprehension of experience“), zweitens die Fähigkeit Relationen zwischen diesen Erkenntnissen ableiten zu können („eduction of relations“) und drittens die Fähigkeit, Zusammenhänge zwischen den gewonnenen Reizen herstellen zu können („eduction of correlation“).

Als Ergänzung zu Spearmans Generalfaktortheorie entwickelte sein Schüler Karl Holzinger [1892 – 1954] das Bi-Faktor-Modell, das davon ausgeht, dass es einerseits einen Generalfaktor gäbe, der in allen Variablen hoch lädt, aber dass es auf der anderen Seite auch sogenannte Gruppenfaktoren gäbe, die lediglich auf einem Teil der Variablen hoch laden würden, auf den anderen dafür gar nicht (Holzinger & Swineford, 1937). Mittels der Bi-Faktor-Methode extrahierte Holzinger einen Generalfaktor und fünf Gruppenfaktoren (zit. n. Carroll, 1993:54 – *spatial relations, verbal, perceptual speed, recognition, associative memory*). Und auch El Koussy (1935) erweiterte das g-Faktor-Modell von Spearman mittels der Bi-Faktoren-Methode um einen „k-Faktor“. Am Rande erwähnt seien ähnliche Ansätze, die von einem Generalfaktor ausgehen. Diese stammen unter anderem von McFarlane („practical ability“), Hargreaves („faculty of imagination“) oder Cox („mechanical aptitude“) (zit. n. Carroll 1993:54).

Die anschließenden Jahre 1938 bis 1961 sind vor allem durch den Methodenwechsel gekennzeichnet, der von Louis Leon Thurstone [1887 – 1955] bereits 1931 ausgelöst wurde. Im Vordergrund dieser Periode steht die klare Identifizierung eines Raumvorstellungsfaktors S (Space). Besonders zu Beginn des Zweiten Weltkriegs wurden jene Testentwicklungen vorangetrieben, die es ermöglichen Soldaten zu selektieren, die für den Flugeinsatz im Rahmen der United States Army Air Force geeignet waren. Im Zuge dieser faktorenanalytisch geleiteten Ansätze bestätigte sich die in Kapitel 2.1 bereits erwähnte Einteilung des Faktors Raumvorstellung in die drei Unterkategorien „*spatial relation*“, „*spatial visualization*“ und „*mental rotation*“ (Eliot & Smith, 1983).

Allerdings sank das Interesse an Raumvorstellung genauso schnell, wie es entstanden war, da durch den Behaviorismus jegliche Form von Introspektion abgelehnt wurde. Weil aber räumliche Vorstellung eine alltägliche kognitive Fähigkeit ist, ist es wenig verwunderlich, dass es in den späten 1960er Jahren in zwei Schüben zu einem wieder-auflebenden Interesse an der Erforschung von räumlichen Fähigkeiten kam.

Der erste Impuls entstand durch die erneute quantitative Erfassung von Raumvorstellung durch Wissenschaftler wie Sheehan (1967). Bugelski (1970) und Paivio (1969) beschäftigten sich zwar auch mit der Messung von Raumvorstellung aber auf einer etwas theoretischeren Basis als Sheehan. Den zweiten Schub zurück zur Erfassung von

Raumvorstellung realisierten die Wissenschaftler Shepard (1975) und Shepard und Metzler (1971) auf kognitiver Ebene, auf dem Gebiet der Neurokognitionen unter anderem Farah (1988) und Kosslyn (1988) (zit. nach Solso 2001: 292 ff).

Shepard und Metzler (1971), die Pioniere auf dem Gebiet der wissenschaftlichen Erforschung von mentaler Rotation waren, wurden vor allem von Joy Paul Guilford [1897 – 1987] und seinem „Structure of Intellect“ Modell beeinflusst (Guilford, 1967). Dieses Modell geht davon aus, dass jede Intelligenzleistung im Allgemeinen auf eine Beteiligung von fünf „*kognitiven Operationen*“, vier „*figuralen Inhalten*“ und sechs „*transformativen Produkten*“ zurückzuführen sei. Kombiniert man diese einzelnen Anteile, erhält man ein System von insgesamt 120 theoretisch postulierten Intelligenzfaktoren. Die Kategorie der Block Rotation Tests (und somit auch des Mental Rotation Test von Shepard und Metzler) beruht auf eben diesen drei genannten Faktoren (Eliot & Smith, 1983).

Da hier lediglich ein kurzer Überblick über die geschichtliche Entwicklung von Raumvorstellungsforschung und –testung gegeben werden konnte, sei der interessierte Leser zur Vertiefung auf die Werke von Eliot und Smith (1983) sowie von Carroll (1993) hingewiesen.

3 Differentialpsychologische Befunde zum Konzept „Mentale Rotation“

Die mentale Rotationsleistung gilt in der psychologischen Forschung als eine der geschlechtssensitivsten kognitiven Domänen, bei denen Männer durchschnittlich um 1-2 Standardabweichungen besser abschneiden als Frauen (Voyer, Voyer & Bryden, 1995; Linn & Peterson, 1985; McGee, 1979). Allerdings gilt es hierbei einzelne Faktoren genauer unter die Lupe zu nehmen, da noch immer nicht restlos geklärt zu sein scheint, was die genaue Ursache für diese Unterschiede ist. Wiederholt taucht deshalb in diesem Zusammenhang die Diskussion um die Anlage-Umwelt-Problematik auf (McGee, 1979). Zum einen werden biologische Faktoren wie beispielsweise Gene, Lateralisierung im Gehirn oder spezifische Geschlechtshormone als Erklärungsansatz herangezogen (Roberts & Bell, 2003; Kimura, 2002; Hausmann, Slabbekoorn, Van Goozen, Cohen-Kettenis & Güntürkün, 2000; McGee, 1979), auf der anderen Seite dürften auch Sozialisation bzw. Erfahrungen im frühen Kindesalter (Baenninger & Newcombe, 1989), geschlechtsspezifische Spiele (Quaiser-Pohl, Geiser & Lehmann, 2006; Bjorklund & Brown, 1998; Newcombe, Bandura & Taylor 1983) sowie Kontrollüberzeugung und Selbstkonzept (Cooke-Simpson & Voyer, 2007) auf die räumliche Vorstellungsfähigkeit einen gewissen Einfluss ausüben.

Das Alter betreffend berichten Linn und Petersen (1985), dass sich der geschlechtsspezifische Unterschied in der räumlichen Vorstellung erst mit Beginn der Pubertät zeige, wobei diese Aussage durch die Tatsache relativiert werden muss, dass vor allem der Mental Rotation Test nach Vandenberg und Kuse (V/K MRT, 1978) nur Personen vorgegeben wurde, die mindestens 13 Jahre alt waren. Für die Fähigkeit in mentaler Rotation, wie sie beispielsweise im Card Rotation Test erhoben wird, ermittelten sie eine gewichtete Effektstärke von $d = .73$ (bei einem 95% Konfidenzintervall .50 – .96) und speziell für den V/K MRT erhielten sie $d = .94$ (.77 – 1.12). Zehn Jahre nach dieser Metaanalyse untersuchten Voyer, Voyer und Bryden (1995) erneut den aktuellen Forschungsstand in Bezug auf Geschlechtsunterschiede in der räumlichen Vorstellung und kamen zu ähnlichen Ergebnissen. Allerdings ermittelten sie lediglich eine gewichtete Effektstärke von $d = .56$ während die Effektstärke für den V/K MRT mit $d = .94$ gleich

blieb. Um den Einfluss des Alters näher zu ergründen widmeten sich Geiser, Lehmann und Eid (2008) speziell dieser Forschungsfrage und kamen zu dem Ergebnis, dass auch schon neunjährige Buben den gleichaltrigen Mädchen überlegen waren, obgleich die Fähigkeit zur mentalen Rotation bei beiden Geschlechtern anstieg, aber auch die Differenz in der Leistung nahm mit Anstieg des Alters zu.

In den nachfolgenden Kapiteln 3.1 bis 3.6 wird näher auf die einzelnen Komponenten eingegangen, die einen möglichen Einfluss auf die unterschiedlichen Raumvorstellung leistungen zwischen Männern und Frauen bedingen. Es wird versucht an viele Themen der psychologischen Forschung, angefangen bei der Neuropsychologie bis hin zu Geschlechtsrollenidentität und Sozialisationseffekte anzuknüpfen, um ein breites Spektrum an Erklärungsansätzen zu eröffnen.

3.1 Neuropsychologie

Die Gehirnstruktur von Männern und Frauen ist in Bezug auf die Verarbeitung kognitiver Inhalte unterschiedlich organisiert. Eine Theorie diesbezüglich geht auf die Jäger- und Sammlervergangenheit der Menschen zurück, in der der Mann dafür verantwortlich war, auch in großer Entfernung auf die Jagd zu gehen, während die Frau sich um die Kindererziehung sowie um die Nahrungssuche in der Nähe der Höhle kümmerte. Infolge dessen entwickelte sich bei Männern der räumliche Sinn besser als bei Frauen, wohingegen Frauen einen Vorteil in der Wahrnehmungsgeschwindigkeit gewannen – vor allem wenn es darum geht, Ähnlichkeiten zwischen zwei Objekten zu identifizieren (Silverman, Choi & Peters, 2007; Kimura, 2002). Aber auch die Sexualhormone Estradiol und Testosteron stehen im Verdacht, wesentlich am Aufbau des Gehirns beteiligt zu sein und so zu alternierenden Neuronenstrukturen zu führen (de Vries, 2004).

Andererseits gilt es in der Scientific Community als hinreichend gesichert, dass bei beiden Geschlechtern die rechte Gehirnhälfte vornehmlich für die Verarbeitung von analytischen und räumlich-orientierenden Informationen zuständig ist, während die linke Hemisphäre eher für die verbale Informationsverarbeitung verantwortlich zu sein scheint. Inwiefern diese Lateralisierung bei Männern und Frauen bei der Bearbeitung von räumlichen Vorstellungsaufgaben eine Rolle spielt, wurde bereits mehrfach unter-

sucht, allerdings mit teils widersprüchlichen Ergebnissen (Thomsen et al., 2000; Dietrich et al., 2001).

Jordan, Wüstenberg, Heinze, Peters und Jäncke (2002) gehen in ihrer Studie davon aus, dass es einen Geschlechtsunterschied in der neuronalen Aktivierung bei Männern und Frauen während der Bearbeitung von mentalen Rotationsaufgaben gibt. Um zu überprüfen, ob dieser genetisch bedingt, und nicht ein Resultat von Leistungsunterschieden ist, haben sie eine fMRI-Studie an 22 Freiwilligen durchgeführt, wobei sie die Leistungen der Männer und Frauen konstant hielten. Jeder Proband nahm an insgesamt vier Versuchsbedingungen teil: (1) klassischer Mental Rotation Test (MRT) nach Shepard und Metzler (1971), (2) mentale Rotation von abstrakten Figuren, (3) mentale Rotation von Buchstaben, (4) gleich-ungleich Entscheidung von nicht rotierten Aufgaben der vorangegangen Aufgabentypen. Jede Aufgabe aus den vier genannten Samples wurde für 3 Sekunden dargeboten. Die Versuchspersonen wurden aufgefordert so schnell und so richtig als möglich auf das Stimulusmaterial zu reagieren, in dem sie einen von zwei Knöpfen betätigten.

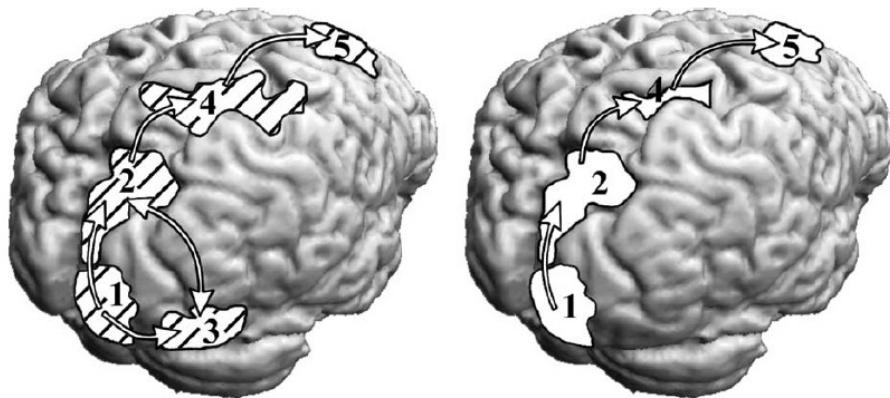


Abbildung 1 Hypothetische Verarbeitungsprozesse zur Lösung von mentalen Rotationsaufgaben bei Frauen (li.) und Männern (re.): (1) primär visuelles Areal, (2) extrastriater Kortex, (3) temporal inferiore /occipital- inferiore Region, (4) parietale Region (intraparietaler Sulcus), (5) prämotorischer Kortex (Jordan et al., 2002:2406)

Bezüglich des Mental Rotation Tests ergab die Studie von Jordan et al. (2002), dass sich Männer und Frauen bei gleicher Leistung signifikant in ihrer Gehirnaktivität unterschieden. Frauen zeigten signifikant bilaterale Aktivitäten während der Bearbeitung der

MRT Aufgaben im superioren Parietallappen¹, mit Beteiligung des intraparietalen Sulcus. Zusätzlich wurden signifikante Aktivitäten bilateral im hinteren Bereich des dorsalen prämotorischen Areals, im prä-supplementären motorischen Areal², im Gyrus temporalis inferior³ sowie im rechten Gyrus frontalis inferior⁴ gefunden.

Bei den Männern herrschten signifikante Aktivitäten im superioren und mittleren occipitalen Gyrus sowie dem parieto-occipitalen Sulcus. Ähnlich wie bei den Frauen waren außerdem der superiore Parietallappen, das prämotorische Areal und das pre-supplementäre motorische Areal bei der Bearbeitung beteiligt (vgl. Abbildung 1).

Als möglichen Erklärungsansatz für die gefundenen Differenzen in der Gehirnaktivität nennen Jordan et al. (2002) folgende drei Hypothesen:

1. Männer und Frauen wenden unterschiedliche Lösungsstrategien für mentale Rotationsaufgaben an.
2. In Ergänzung zu Punkt 1 legen Männer und Frauen einen unterschiedlichen Fokus auf die zu bearbeitenden Aufgaben.
3. Das neuronale Netzwerk für MRT-Aufgaben ist bei Männern und Frauen geschlechtsspezifisch organisiert.

Roberts und Bell (2003) knüpften an die Ergebnissen von Jordan et al. (2002) an und stellten die Hypothese auf, dass die oben genannte Lateralisierung nur auftritt, wenn das Itemmaterial komplex gestaltet ist und die mentale Rotation nicht nur anhand einer Achse stattfindet, da bei leichten Aufgaben auch der linke Parietallappen an der Verarbeitung beteiligt sei. Mittels einer EEG-Studie konnten sie beweisen, dass Männer bei leichten, zwei-dimensionalen mentalen Rotationsaufgaben vermehrt auf die linke Gehirnhälfte zurückgriffen und bei komplexeren, dreidimensionalen Rotationsaufgaben (wie sie der MRT darstellt) ausschließlich Aktivitäten in der rechten parietalen Region aufwiesen. Frauen hingegen zeigten bei leichten und schweren Aufgaben vermehrte

¹ **Superiorer Parietallappen:** Dient dem Erkennen visueller Objekte, der Synthese und dem Vergleich visueller Objekte, der Linienorientierung, dem Nachzeichnen sowie dem Erkennen von Objekten durch Berührung (Seidentopf, 2010).

² **Supplementäres Motorisches Areal:** Kontrolliert den Ablauf und die Planung komplexer Bewegungen – auch aktiv, wenn man sich eine Bewegung vorstellt (Seidentopf, 2010).

³ **Gyrus temporalis inferior:** Dient der Analyse von Farbe und Form visueller Stimuli, die über cortikokortikale Projektionen zugeführt werden (Seidentopf, 2010).

⁴ **Gyrus frontalis inferior:** Zuständig für das Halten von Informationen im Kurzzeitgedächtnis. (Seidentopf, 2010)

Aktivitäten im rechten Parietallappen zeigten, obgleich sich die Leistungen von Männern und Frauen in den leichteren zwei-dimensionalen Aufgaben statistisch nicht von einander unterschieden.

Dietrich et al. (2001) wählten eine andere Herangehensweise, in dem sie den fluktuierenden Östrogenspiegel der Frauen mit in ihre Studie aufnahmen. Sie testeten 14 Frauen mit dem Mental Rotation Test zu zwei Zeitpunkten. Der erste Zeitpunkt fand während der Menstruation der Frauen statt, der zweite um den 11. Tag des Zyklus (Eisprung). Aufgrund der ausgewerteten Daten kamen sie zu dem Schluss, dass Frauen während der Menstruation (= niedriger Östrogenspiegel) ähnliche Gehirnaktivitäten zeigten wie die männliche Kontrollgruppe, während sie in der Phase des Eisprungs (= hoher Östrogenspiegel) vor allem Aktivitäten im Gyrus angularis / occipito-temporalis und im superioren Parietallappen zeigten.

Als interessanter Zusatz zu den bereits berichteten Studien zeigten Vogel, Bowers und Vogel (2003) in ihrer Meta-Analyse, dass vor allem schlechte Raumvorsteller zu starken Aktivitäten in der rechten Gehirnhälfte tendieren, während gute Raumvorsteller zu keiner bestimmten Gehirnhälfte neigen.

3.2 Hormoneller Einfluss

Der hormonelle Einfluss auf die kognitiven Unterschiede zwischen Männern und Frauen stützt sich vor allem auf die Theorie, dass die Geschlechtshormone maßgeblich zur Organisation des neurokognitiven Systems während eines frühen Stadiums der embryonalen Entwicklung beitragen. Hormone, die von den Geschlechtsdrüsen (Gonaden) ausgesondert werden, stehen schon seit längerer Zeit in Verdacht, das Verhalten sowie die kognitiven Fähigkeiten des Menschen zu beeinflussen. Die wichtigsten Sexualhormone in diesem Zusammenhang sind Testosteron und Östrogen. Genauso wie der weibliche Zyklus auf den Östrogenspiegel einwirkt, variiert der Testosteronspiegel beim Mann in Abhängigkeit von der Jahres- sowie von der Tageszeit (Kimura, 1997; Moffat & Hampson, 1996; Kimura & Hampson, 1994). Die Rolle des Testosterons bei der Erbringung von guten Leistungen bei mentalen Rotationsaufgaben ist jedoch noch nicht vollends geklärt.

Auch das weibliche Geschlechtshormon Östrogen soll auf die räumlichen Fähigkeiten bei Frauen Einfluss nehmen. Einige Studien berichten in diesem Zusammenhang unterschiedliche Leistungen im Laufe des Menstruationszyklus (Hausmann et al., 2000; Phillips & Silverman, 1997; Kimura & Hampson, 1994).

3.2.1 Natürliche Fluktuation im Östrogenhaushalt

Bei Frauen mit regelmäßigen Menstruationszyklus schütten die Eierstöcke mehr als 95% des im Körper zirkulierenden Östradiols aus (Aloisi, 2007). Zu Beginn der Menstruation befindet sich das Östrogen auf seinem niedrigsten Level, um zwischen dem 11.-14. Tag auf sein Maximum anzusteigen, da sich die Frau in dieser Zeit in ihrer fruchtbaren Phase befindet und es zum Eisprung kommt (vgl. Abbildung 2). Sobald sich die Frau in der Menopause befindet, stellen die Eierstöcke die Produktion von Östrogen ein bzw. verlagern sie auf extragonadale Stellen. Der Östrogenspiegel bei der Frau variiert jedoch nicht nur in Abhängigkeit zum Menstruationszyklus und der Menopause (Simpson, 2003).

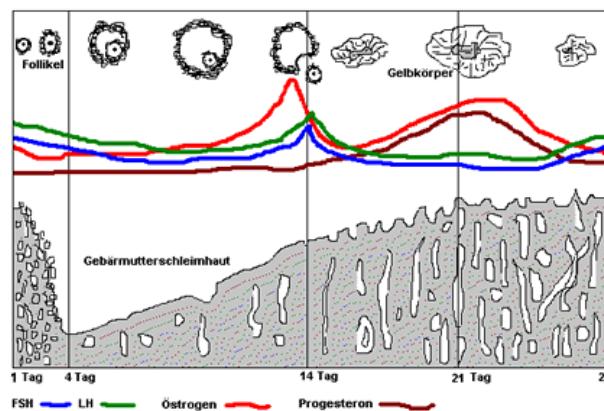


Abbildung 2 Hormonelle Schwankungen während eines Menstruationszyklus

Die Frau in der Menopause befindet, stellen die Eierstöcke die Produktion von Östrogen ein bzw. verlagern sie auf extragonadale Stellen. Der Östrogenspiegel bei der Frau variiert jedoch nicht nur in Abhängigkeit zum Menstruationszyklus und der Menopause (Simpson, 2003).

3.2.2 Natürliche Fluktuation im Androgenhaushalt

Interessanter Weise und nebenbei erwähnt fanden Kimura und Hampson (1994) in ihrer Studie heraus, dass der Testosteronspiegel bei Männern systematisch in Abhängigkeit von der Jahres- und Tageszeit variiert. Da das Testosteron bei Männern hauptsächlich in den Hoden produziert wird, zeigen europäische und nordamerikanische Männer im Herbst ein höheres Testosteronlevel als im Frühling, was mit der erhöhten sexuellen Aktivität in diesem Zeitraum zusammenhängen könnte. Dies würde auf die evolutionäre Vergangenheit verweisen, in der Kinder, die in den Wintermonaten geboren wurden, eine schlechtere Überlebenschance hatten, als andere Kinder. In Anbetracht dessen, dass Männer im Frühling einen niedrigeren Testosteronlevel haben, erzielen sie in dieser Jahreszeit auch bessere Ergebnisse bei Raumvorstellungstests (Kimura & Hampson, 1994).

Zusätzlich zu den saisonalen Veränderungen unterliegt der Testosteronspiegel auch zirkadianen Schwankungen. In den frühen Morgenstunden befindet sich der Testosteronlevel bei Männern im Vergleich zur restlichen Tageszeit auf seinem Höchststand, wobei die räumliche Fähigkeit zu dieser Zeit am schlechtesten ist (Kimura 2002; Moffat & Hampson, 1996; Dabbs, 1990). Von dieser Schwankung im Testosteronhaushalt sind jedoch nur solche Tests betroffen, die räumliche Fähigkeiten erfassen, während Leistungen im mathematischen Schlussfolgern in keiner Weise beeinflusst werden.

Diese Ergebnisse sind jedoch sehr kritisch zu betrachten, da in neueren, bisher unpublizierten, Forschungsarbeiten gegenteilige Resultate berichtet werden (Spohn & Buse, 2010)

3.2.3 Experimente zur Hormonhypothese

Studien an Patienten mit angeborener Nebennierenrindenhyperplasie (CAH – congenital adrenal hyperplasia), die bereits pränatal einen erhöhten Nebennieren-Androgenlevel aufweisen, kommen bezüglich der räumlichen Fähigkeit bei dieser Patientengruppe zu divergenten Ergebnissen (Hines, Fane, Pasterski, Mathews Conway & Brook, 2003; Helleday, Bartfai, Ritzen & Forsman, 1994; Resnick, Berenbaum, Gottesman & Bouchard, 1986). Frauen mit CAH zeigen bereits in der Kindheit mehr männliche Verhaltenseigenschaften, indem sie sich eher mit typisch männlichen Spielzeugen oder männlichen Freunden befassen. Obwohl es naheliegend schien, dass diese Frauen auch bessere Werte bei Aufgaben zur mentalen Rotation erbringen, konnten in der Studie von Hines et al. (2003) keine signifikanten Unterschiede zu der weiblichen Kontrollgruppe gefunden werden. In derselben Studie schnitten Männer mit CAH signifikant schlechter ab als die Probanden der männlichen Kontrollgruppe. Dies deutet darauf hin, dass es einen optimalen Testosteronlevel gibt, bei dem man bei bestimmten räumlichen Aufgaben am besten abschneidet.

Den gravierenden Einfluss von Sexualhormonen wiesen unter anderem auch Studien an Frau-zu-Mann-Transsexuellen nach. Hormonbehandelte Frauen zeigten bereits nach einer dreimonatigen Einnahme von hohen Dosen Testosteron einen drastischen Anstieg in ihren räumlichen Fähigkeiten, während ihre verbalen Leistungen erheblich sanken (Van Goozen, Cohen-Kettenis, Gooren, Frijda, & Van De Poll, 1995).

Bezüglich der Schwankung der Leistungsfähigkeit bei Frauen im Verlauf des Menstruationszyklus, fanden Heider, Adlmann und Gittler (2010a) Hinweise darauf, dass jene Frauen, welche die konstante Einnahme von hormoneller Kontrazeptiva berichten, während der Menstruation signifikant bessere Ergebnisse erzielten, als um den 11. – 14. Tag im Zyklus. Die Bearbeitungsgeschwindigkeit der Probanden variierte ebenfalls in Abhängigkeit zum Menstruationszyklus. Frauen mit niedrigem Östrogenlevel benötigten signifikant mehr Zeit zum Bearbeiten mentaler Rotationsaufgaben, als jene die sich in ihren fruchtbaren Tagen befanden (Heider, Adlmann, & Gittler, 2010b).

3.3 Geschlechtsrollenidentität

Es gibt zahlreiche Beispiele in denen Stereotype und individuelle Überzeugungen auf das biologische System, wie beispielsweise die Hormonsekretion, Atemfrequenz, Verdauung und vieles mehr, einwirken können. Einen ähnlichen Einfluss kann die Geschlechtsrollenidentität auch auf die kognitive Leistungsfähigkeit von Männern und Frauen haben (Halpern & Tan, 2001).

Die Geschlechtsrollenidentität versteht sich in diesem Zusammenhang als das Ausmaß, in dem Individuen stereotypische Merkmale von Männern und Frauen in ihrer Persönlichkeit internalisiert haben. Aktuelle Forschungsansätze sehen daher das biologische Geschlecht („sex“) und das psychologische Geschlecht („gender“) als zwei voneinander relativ unabhängige Größen an. Das bedeutet, dass Männer und Frauen unabhängig von ihrem biologischen Geschlecht sowohl typisch männliche als auch typisch weibliche Persönlichkeitseigenschaften in ihre Persönlichkeit integrieren, und dadurch sowohl in typisch männlicher als auch in typisch weiblicher Art und Weise agieren können (Saucier, McCreary & Saxberg, 2002).

Der Aspekt der sexuellen Orientierung stellte für Rahman und Wilson (2003) ein interessantes Forschungsgebiet dar, da sie davon ausgingen, dass Homosexuelle sich zu meist entgegen ihres biologischen Geschlechts verhalten würden. Sie hypothetisierten, dass heterosexuelle Männer aufgrund ihres psychologischen Geschlechts („männlich“) im Mental Rotation Test besser abschneiden würden, als homosexuelle Männer, während es sich bei den Frauen genau anderes herum verhalten sollte. Tatsächlich erzielten heterosexuelle Männer in ihrer Studie signifikant bessere Ergebnisse ($F(1, 119) = 13,44$;

$p < ,0001$) als homosexuelle. Bei den Frauen fanden sie ihre Theorie ebenfalls bestätigt, da homosexuelle Frauen ebenfalls signifikant höhere Werte im MRT erzielten ($F(1,119) = 6,672$; $p = ,01$), als heterosexuelle Frauen.

Massa, Mayer und Bohon (2005) untersuchten in ihrer Studie, inwiefern die Geschlechtsrollenidentität von Männern und Frauen deren Leistungen in einem Raumvorstellungstest beeinflussen können. Sie gehen jedoch nicht mehr, wie noch in der klassischen Sichtweise verankert, davon aus, dass die kognitive Leistung in einem entsprechenden Test nur von der Fähigkeit der Person alleine abhängt, sondern auch davon wie die Testinstruktion gestaltet ist, und ob der Test eine Fähigkeit anspricht, von der die Probanden annehmen, dass es eine Fähigkeit sei, die ihnen im Speziellen und ihrem Geschlecht im Allgemeinen liegt (vgl. Abbildung 3).

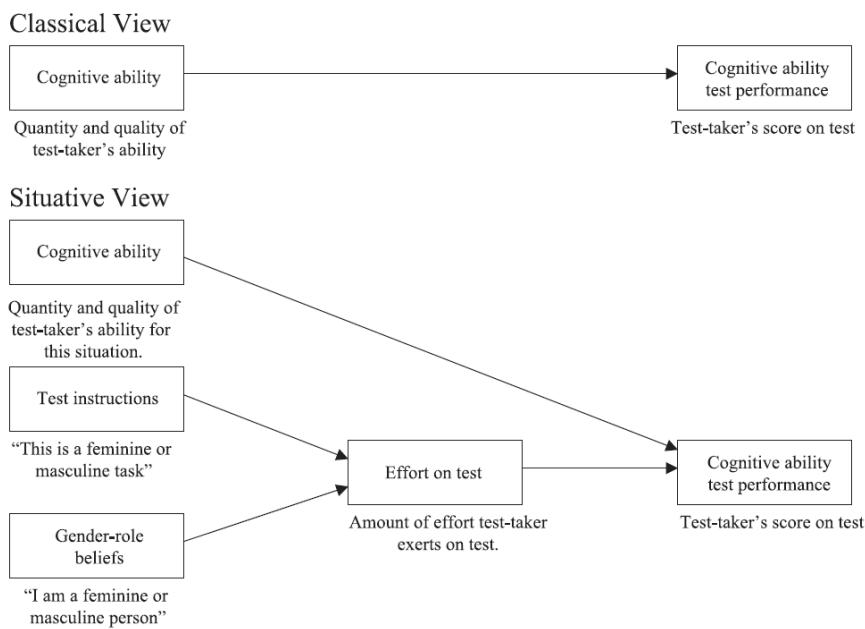


Abbildung 3 Darstellung der klassischen und modernen Sichtweise beim Testen kognitiver Fähigkeiten (Massa et al., 2005:101)

Für die weiblichen Probanden fanden Massa, Mayer und Bohon (2005) ihre Hypothese bestätigt, dass Frauen, mit einer weiblichen Geschlechtsrollenidentität beim MRT besser abschnitten, wenn die Instruktion die empathische Fähigkeit in diesem Test hervorhob, während Frauen mit männlicher Geschlechtsrollenidentität bessere Leistungen erbrachten, wenn die Instruktion die räumliche Vorstellungsfähigkeit in den Mittelpunkt

stellte. Ungeachtet dessen konnten sie keine Rückschlüsse auf die Männer machen, da diese scheinbar für derartige Manipulationen unempfänglich waren.

Cooke-Simpson und Voyer (2007) gingen, in Ergänzung zur Geschlechtsrollenidentität, der Frage nach, ob Selbstvertrauen in die eigenen (räumlichen) Fähigkeiten einen Einfluss auf die Leistung im Mental Rotation Test haben würde. Hierfür testeten sie 80 Studenten ($m=40$, $w=40$) und batn sie nach jeder Aufgabe auf einer Ratingscale von 1 (überhaupt nicht sicher) bis 7 (extrem sicher) anzugeben, wie stark sie der Überzeugung seien, das vorangegangene Item richtig bearbeitet zu haben. Sie fanden heraus, dass das bloße Vorhandensein einer Ratingscale negativ mit der gesamten Testleistung korrelierte, aber auch, dass Männer signifikant eher dazu neigen, sich ihrer Antworten und ihrer Fähigkeiten sicher zu sein, als Frauen, obwohl sich Frauen realistischer einschätzen konnten. Weiters wiesen Cooke-Simpson und Voyer (2007) nach, dass Frauen grundsätzlich weniger Vertrauen in die eigenen Fähigkeiten setzten, als ihre männlichen Kollegen.

3.4 Sozialisationseffekte

Im Gegensatz zu den Erkenntnissen aus der Neurobiologie, gehen die neueren For-

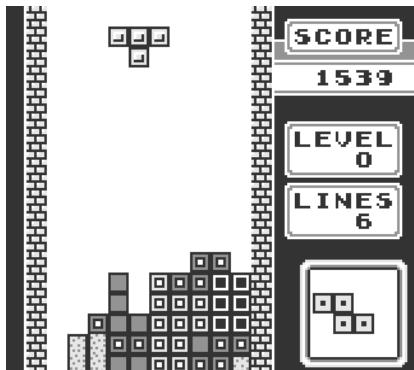


Abbildung 4 "Tetris®"

schungsansätze davon aus, dass die frühe Sozialisation einen großen Einfluss auf die Entwicklung kognitiver Fähigkeiten hat. Bereits 1983 beschäftigten sich Newcombe, Bandura und Taylor mit der Frage, ob Geschlechtsunterschiede in der mentalen Rotationsfähigkeit möglicherweise auf die (räumlichen) Aktivitäten in der Jugend und im Erwachsenenalter zurück zu führen seien. Sie konnten jedoch keine signifikanten Zusammenhänge berichten. Während sich die untersuchten Aktivitäten bei Newcombe et al. noch auf Baseball (typisch männlich) bzw. Ballett (typisch weiblich) und ähnliches konzentrierten, beschäftigen sich die aktuellere Studien zeitgemäß unter anderem mit dem Einfluss von Computerspielen (wie beispielsweise von Schieß- oder Rollenspielen) auf das räumliche Vorstellungsvermögen (Quaiser-Pohl, Geiser & Lehmann, 2006). Terlecki, Newcombe und Little (2008) fanden zum Beispiel heraus, dass das wiederholte Spielen von „Tetris®“ (Abbildung 4) die Leistungen im Mental Rotati-

onstest verbessern kann. Eine weitere Studie von Newcombe et al. (2009) bestätigte, dass das wiederholte Spielen von „Tetris®“ die Leistungen im Mental Rotation Test verbessert. Die Autoren schließen daraus, dass die wiederholte Spielaktivität die räumlichen Fähigkeiten der Spieler verbessert. Dies ist ein interessanter Befund, da es sich um ein einfaches Tetris-Spiel handelt, das nicht speziell auf räumliche Fähigkeiten ausgerichtet ist. Es ist möglich, dass das Spiel die räumlichen Fähigkeiten der Spieler durch die ständige Orientierung und Planung beim Fallen von Blöcken verbessert.

on Test nach Vandenberg und Kuse (1978) signifikant und über einen längeren Zeitraum hinweg, verbesserten. Interessant an ihrer Studie war, dass sie nachweisen konnten, dass sich der Trainingseffekt durch das Computerspielen auch auf andere Tests generalisierte ließ.

Der erleichternde Effekt von Computerspielen auf das räumliche Vorstellungsvermögen wurde unter anderem durch die Studie von Cherney (2008) repliziert, in der vor allem Frauen signifikant von einem Training mittels Tetris© bzw. Autorennspiels⁵ profitierten ($t(39) = -3,6; p = ,001$).

3.5 Holistische versus analytische Bearbeitungsstrategien

Ein weiterer Erklärungsansatz für die Überlegenheit der Männer bei mentalen Rotationsaufgaben besagt, dass diese eine effektivere Bearbeitungsstrategie anwenden als Frauen. Männer, so die Hypothese, sehen die Objekte ganzheitlich („holistisch“) und erkennen dadurch schneller, ob es sich um eine gedrehte oder eine gespiegelte Figur handle. Frauen hingegen neigen dazu, diesen Aufgabentypus in Gedanken zu verbalisieren oder in seine Einzelteile zu zerlegen, was nicht nur längere Reaktionszeiten bei komplexeren Aufgaben nach sich zieht, sondern auch zu einer erhöhten Fehleranfälligkeit führt (Heil und Jansen-Osmann, 2008; Jordan, Wüstenberg, Heinze, Peters und Jäncke, 2002). Dies könnte auch erklären, warum gerade beim Mental Rotation Test von Vandenberg und Kuse (1978) sowie in seiner revidierten Fassung (Peters, Laeng, Latham, Jackson, Zaiyouna und Richardson, 1995) der Geschlechtsunterschied besonders resistent ist. Peters (2005) untersuchte in einer groß angelegten Studie den Zeitfaktor und seine Bedeutung für den Geschlechtsunterschied und fand dabei heraus, dass sowohl Männer als auch Frauen, von einer verlängerten Testphase profitierten, der Geschlechtsunterschied jedoch signifikant bestehen blieb. Er führte dies hauptsächlich darauf zurück, dass Frauen länger brauchen, um sich sicher zu sein, dass eine als „falsch“ verifizierte Antwort – also gespiegelte – tatsächlich nicht in die Ausgangsfigur gedreht werden kann. Männer hingegen halten sich nicht allzu lange mit solchen Überlegungen auf. Hirnstein, Bayer und Hausmann (2009) unterscheiden, zusätzlich zu den beiden bereits genannten Bearbeitungsstilen, den konservativen vom sprunghaften Lö-

⁵ Cherney (2008) verwendete für ihre Studie sowohl ein Spiel namens „Tetrus“, das dem Nintendo-Spiel Tetris © nachempfunden wurde, sowie ein 3d-Autorennspiel, dass mit den Figuren aus dem Film „Antz – Was krabbelt da“ animiert wurde.

sungsweg. Beim konservativen Weg, der ihrer Hypothese nach vor allem von Frauen angewendet wird, wird sowohl die als richtig identifizierte Antwort nochmals überprüft, als auch die als falsch eingestufte Antwortmöglichkeit wiederholt mit der Ausgangsfigur abgeglichen. Männer dürften hingegen eher den sprunghaften Lösungsweg präferieren, bei dem sie sich nur auf die zwei richtigen Antwortmöglichkeiten konzentrieren, und, sobald sie glauben, diese gefunden zu haben, ohne weitere Überprüfung zur nächsten Aufgabe weiter gehen. Um ihre Hypothese abzusichern, gaben sie ihren Versuchspersonen jeweils zwei Versionen des Mental-Rotation-Tests vor. Die eine Fassung stellte den klassischen MRT nach Vandenberg und Kuse dar, bei dem immer genau zwei Antwortmöglichkeiten richtig waren, die zweite Fassung jedoch ließ alle Antwortmöglichkeiten von null bis vier richtigen Antworten zu. Für die klassische Version fanden sie ihre Hypothese bestätigt, allerdings war der sprunghafte Bearbeitungsstil in der modifizierten Fassung nicht möglich, zumal man alle Antwortfiguren mit der Ausgangsfigur vergleichen musste. Die Bearbeitungsstile von Männern und Frauen nähern sich einander an, sobald es keine fixe Vorgabe bezüglich der Anzahl richtiger Antworten gibt.

3.6 Psychobiosoziales Modell bei kognitiven Geschlechtsunterschieden

In der klassischen Forschung zu kognitiven Geschlechtsunterschieden, gab es eine große Debatte zur Anlage-Umwelt-Problematik, bei der vor allem Neurobiologen annahmen, dass sämtliche Ähnlichkeiten und Unterschiede zwischen den Geschlechtern auf ihr Erbgut zurück zu führen seien. Die heutige Forschung verfolgt eher die Idee, dass die meisten Unterschiede durch Sozialisation entstehen. Versucht man jedoch diese beiden Faktoren nicht als einander ausschließende Erklärungsansätze zu betrachten, sondern vielmehr als voneinander unabhängige Variable, die miteinander interagieren können, erhöht sich der Anteil der erklärbaren Varianz für mentale Rotation merklich. Halpern und Tan (2001) schlagen deshalb vor die beiden Faktoren Anlage und Umwelt nicht mehr auf einem Kontinuum anzusiedeln, sondern sie in eine zirkuläre Beziehung zueinander zu setzen (Abbildung 5).

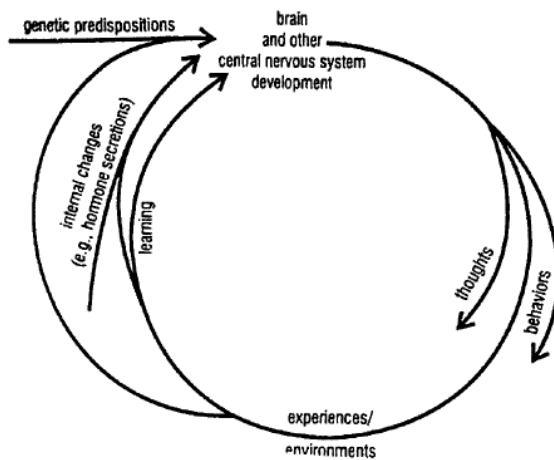


Abbildung 5 Schematische Darstellung des psychobiosozialen Models bei dem Anlage und Umwelt stetig und untrennbar sind (Halpern und Tan, 2001:396)

Halpern und Tan (2001) postulieren, dass die in den vorangegangenen Kapiteln diskutierten Faktoren, wie die geschlechtsspezifische Gehirnstruktur, Sexualhormone, Geschlechtsrollenidentität und Sozialisation nicht getrennt für sich betrachtet werden dürfen, sondern gemeinsam auf die kognitiven Fähigkeiten eines Individuums einwirken, und erst das Zusammenspiel dieser Faktoren zu der im Endeffekt erbrachten Leistung führt.

In Bezug auf das psychobiosoziale Modell wurde trotz alledem nur der gemeinsame Einfluss der Variablen „Geschlechtsstereotyp“ und „Sexualhormon“ auf die kognitiven Fähigkeiten untersucht (Hausmann, Schoofs, Rosenthal & Jordan, 2009; Halpern & Tan, 2001). Hierfür wurde einen Self-Rating Fragebogen, bei dem Geschlechtsstereotypen für Männer und Frauen aktiviert wurden verwendet. Für die Dimension der mentalen Rotation fanden sie einen signifikanten Zusammenhang zwischen der Aktivierung von Geschlechtsstereotypen, die vor allem die Männer zu besseren Leistungen antrieben, während der Leistungsabfall bei den Frauen, im Vergleich zur Kontrollgruppe nicht signifikant ausfiel. Zusätzlich ließen sich erhöhte Testosteronwerte in der geprägten Männergruppe finden.

4 Der Mental Rotation Test (MRT)

4.1 Von den Anfängen des Mental Rotation Tests

Shepard und Metzler veröffentlichten erstmals 1971 eine chronometrische Studie zur räumlich mentalen Vorstellungskraft. In ihrer Studie verwendeten sie Aufgaben, die aus zehn gleichartigen Würfeln bestanden, die jeweils über drei rechtwinklige Abzweigungen

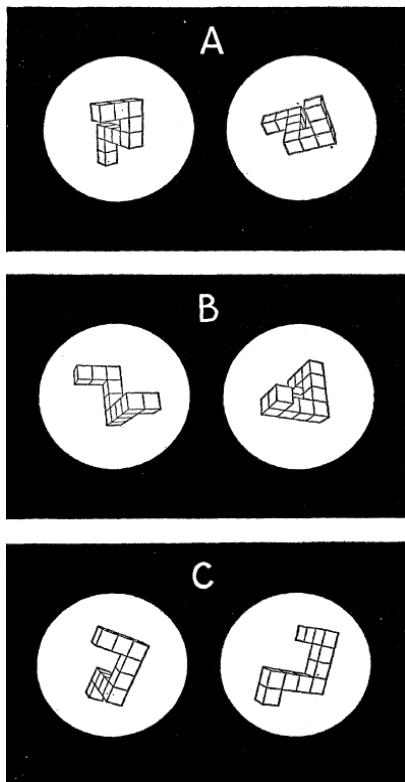


Abbildung 6 Die drei unterschiedlichen Aufgabentypen von Shepard und Metzler (1971)

gen und einer Größe von 4 bis 5 cm verfügten (siehe Abbildung 6). Die Aufgaben wurden so konstruiert, dass sie entweder in der Ebene (A) oder entlang der horizontalen Achse (B) rotiert wurden, aber immer noch in die Ausgangsfigur (links) überführt werden konnten, oder aber eine spiegelverkehrte, rotierte Variante der Ausgangsfigur darstellten (C). Die Rotationen bestanden aus Drehungen in 20° Schritten, wobei von den 18 resultierenden Ansichten jeweils nur sieben ausgewählt wurden, um sehr ähnliche Antwortalternativen auszuscheiden. Obwohl die Figuren am Computer erzeugt wurden, handelte es sich hierbei noch um einen Papier und Bleistift Test, bei dem 1600 Karten mit den erwähnten drei Unterkategorien hergestellt wurden. Für das Antwortformat entschieden sich Shepard und Metzler

(1971) für einen paarweisen Vergleich. Die ersten Ergebnisse wiesen darauf hin, dass eine nahezu lineare Korrelation zwischen dem Rotationswinkel der beiden Vergleichsfiguren und deren Bearbeitungszeit, vorlag.

1978 entwickelten Vandenberg und Kuse diesen Test weiter, indem sie den Probanden nicht nur eine Vergleichsfigur vorgaben, sondern das Antwortformat auf vier Alternativen ausweiteten. Allerdings legten sie fest, dass stets zwei Antworten als richtig markiert werden müssen, wohingegen zwei Antwortalternativen immer falsch seien und unmarkiert bleiben sollten. Die Distraktoren bestanden entweder aus zwei rotierten und

gespiegelten Figuren der Ausgangsfigur, oder aus zwei Figuren die andere Ausgangsfiguren darstellten. Abbildung 7 veranschaulicht diese beiden Aufgabentypen.

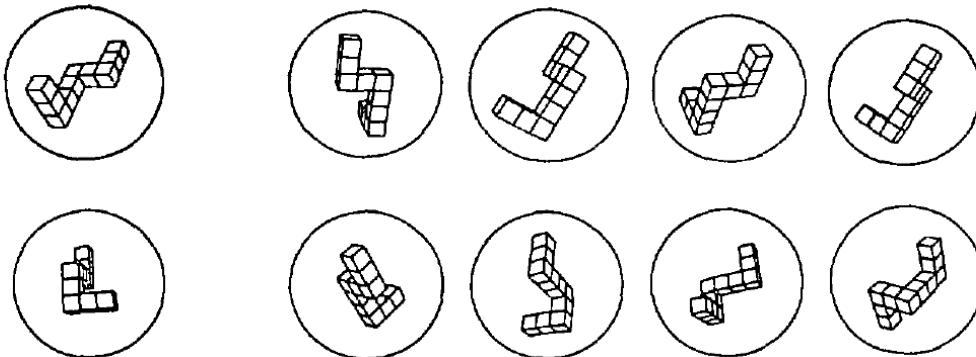


Abbildung 7 Mental Rotation Test nach Vandenberg und Kuse (1978).

Links befinden sich jeweils die Ausgangsfiguren, rechts daneben die vier Antwortalternativen bei denen jeweils genau zwei Antworten (bis auf die Rotation) ident mit der Ausgangsfigur sind. In der ersten Reihe handelt es sich bei Figur eins und vier um korrekte Antworten, zwei und drei sind gespiegelte Antwortalternativen. In Reihe zwei sind die Figuren zwei und drei die richtigen Alternativen, bei den Distraktoren handelt es sich um zwei andere Ausgangsfiguren.

Vandenberg und Kuse (1978) legten für die Auswertung ihrer Version des Mental Rotation Test fest, dass nur jene Aufgaben (Reihe) als richtig zu werten seien, bei denen die beiden korrekten Antworten gegeben wurden, um so die Ratewahrscheinlichkeit zu reduzieren. Wählt ein Proband jedoch eine falsche Antwortalternative, führt dies zu einem Abzug eines Punktes. Im Anschluss gaben sie 312 Probanden den Test, bestehend aus 2x 10 Aufgaben, vor und fanden signifikante Leistungsunterschiede zwischen den beiden Geschlechtern zugunsten der Männer.

4.2 Itemanalyse

Bereits im Jahre 1987 beschäftigten sich Johnson und Meade mit dem Mental Rotation Test und mit der Frage, welche Aufgabentypen schwieriger wären als andere. Sie schreiben diesbezüglich:

Items for which the two ends of the figure are parallel are more difficult than items for which the two ends are orthogonal. This item characteristic interacts with two

others: the extent of the rotation required before the two figures can be compared, and whether or not the figures match. Thus the most difficult item is one for which a large angular rotation is required, for which the ends of the figures are parallel, and for which the correct answer is “different”. (Johnson & Meade, 1987: 737)

Hornke (2002) wiederum beschreibt drei Itemeigenschaften, die aus seiner Sicht für die unterschiedlichen Schwierigkeitsstufen entscheidend sind:

1. Komplexität des Rasteraufbaus: Anzahl der Extremitäten (3,4,5 „Arme“)

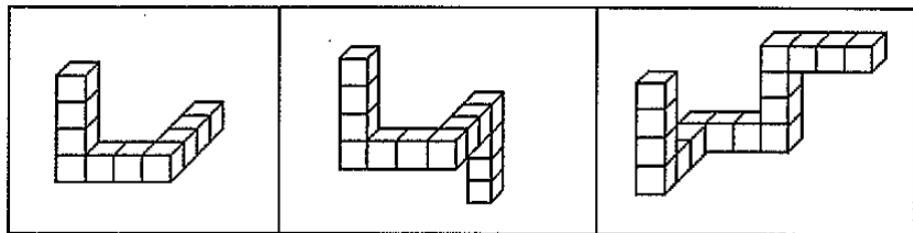


Abbildung 8 Unterschiedliche Komplexität des Rasteraufbaus

2. Anzahl an räumlichen Abzweigungen: 1 oder 2

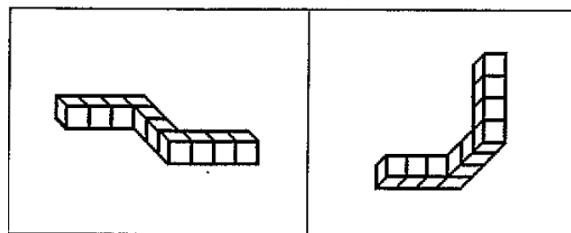


Abbildung 9 Figur mit 1 bzw. 2 Abzweigungen im Raum

3. Transformationskomplexität: Drehwinkel (z.B. 20°–Schritte,...), Anzahl der Rotationen (rechts-links und/oder vor-zurück)

Allerdings sind nicht alle aus diesem Regelwerk entstehenden Aufgaben für die Testkonstruktion verwertbar, da daraus auch Ansichten resultieren können, die uneindeutig sind, weil beispielsweise ein oder mehrere „Arme“ aufgrund der Komplexität nicht mehr zweifelsfrei zu identifizieren wären (Hornke, 2002).

Ein weiterer, nicht unwesentlicher Punkt, der auf die Bearbeitung dieser Aufgaben erschwerend bzw. erleichternd wirkt, besteht in der vorgegebenen Anzahl an Vergleichsfiguren. Peters und Battista (2008) berichten, dass die Geschlechtsunterschiede beim MRT vor allem in der Version von Vandenberg und Kuse (1978) signifikant werden. Sie führten dies unter anderem auf die erhöhte Augenbewegung zurück, die erforderlich

ist, um mehrere Figuren miteinander vergleichen zu können. In der ursprünglichen Version von Shepard und Metzler (1971) müssen die Probanden lediglich zwei Figuren auf ihre mögliche Gleichheit überprüfen, während die Aufgabenstellung bei Vandenberg und Kuse multiple Vergleiche erfordert und somit auch höhere Leistungen an das Kurzzeitgedächtnis stellt, um falsch-positive bzw. falsch-negative Antworten zu vermeiden.

4.3 Modifizierung des MRT für diese Arbeit

Da die klassischen Aufgaben im Mental-Rotation-Test strenggenommen nur 2-dimensionale Darstellungen von 3-dimensionalen Objekten repräsentieren, wurde in der vorliegenden modifizierten Fassung des MRT, mittels rot-blau Anaglyphen-Brille, eine 3-dimensionale Darstellung realisiert, die einen alltagsnahen Eindruck vermitteln soll.

Entsprechend dem Einwand von Peters und Battista (2008), die feststellten, dass jene Studien, die mit dem klassischen Shepard-Metzler Test arbeiteten (= paarweise Bearbeitung der Aufgaben) entweder nicht signifikante oder nur schwache Geschlechtsunterschiede finden konnten, während jene mit den Modifikationen von Vandenberg und Kuse zu signifikanten Unterschieden führten, wurde für die Konstruktion des vorliegenden Tests, eine Variante gewählt, bei der eine Ausgangsfigur mit vier Antwortfiguren verglichen werden musste, wobei alle möglichen Antwortmuster von null bis vier richtigen Antworten realisiert wurde.

Weiters wurden solche Aufgabentypen entwickelt, die entsprechend der Hypothese von Johnson und Meade (1987) sowohl gleichlange, als auch zwei unterschiedlich lange Enden besitzen. Die Gradanzahl sowie die Rotationsrichtung variieren entsprechend der These von Hornke (2002), um auf diese Weise unterschiedliche Itemschwierigkeiten erzeugen zu können.

Zusätzlich wurde das ursprünglich vorhandene Gitternetz bewusst aus den neuen Aufgaben ausgeblendet, da hier ein hoher Anteil an Reasoning, anstatt der reinen mentalen Rotationsleistung, vermutet wurde. Dies hätte eventuell die Eindimensionalität des Itemmaterials und damit die Rasch-Homogenität gefährdet.

Die neu entwickelten Aufgaben bestehen aus 6 bis 10 Würfeln pro Item, die mindestens zwei Abzweigungen im Raum machen (also auf allen Ebenen – x-y-z – mindestens einmal vertreten sind). Wäre diese Bedingung nicht erfüllt worden, wären die gespiegelten Aufgaben nicht eindeutig von den nicht-gespiegelten Aufgaben zu unterscheiden gewesen. Bei der 3-dimensionalen Version des MRT wurde ein bewusst starker räumlicher Eindruck eingesetzt um tatsächlich einen Effekt der Komponente „realitätsnahes Testen“ messen zu können.

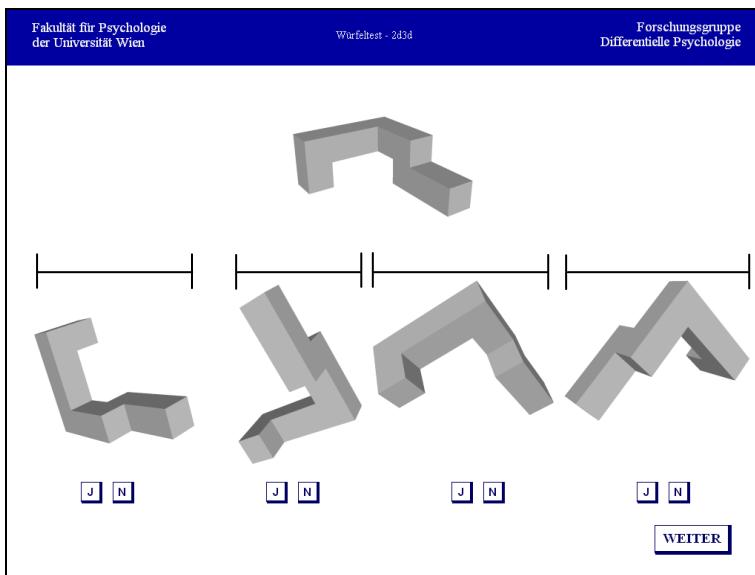


Abbildung 10 Beispiel für ein neugeneriertes Item

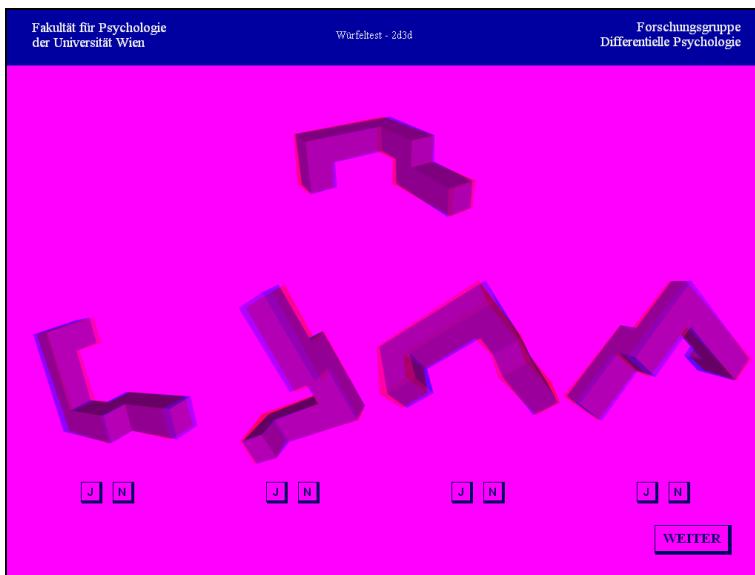


Abbildung 11 Beispiel für ein stereoskopisch generiertes Item (zu sehen mit Anaglyphenbrille)

Abbildung 10 und Abbildung 11 zeigen exemplarisch, wie die neu generierten Aufgaben in zweidimensionaler und dreidimensionaler Ansicht präsentiert wurden. In beiden Ver-

sionen wurde darauf geachtet, dass die vollständige Tiefeninformation vorhanden war. Das heißt kein Würfel wurde komplett durch einen anderen verdeckt und auch die Ausdehnung in der Breite in jeder Antwortmöglichkeit war etwa gleich.

Die neuen Aufgaben wurden mit dem Programm Visualization Toolkit (VTK) realisiert. VTK ist ein objekt-orientiertes open source Programm auf Basis von C++ (<http://www.vtk.org>), das von Will Schroeder, Ken Martin und Bill Lorensen (1998) entwickelt wurde und ursprünglich der Darstellung von Informationen aus bildgebenden Verfahren diente. Das Programm ermöglicht es, anhand der x-, y-, bzw. z-Achse beliebig stark gedrehte Perspektiven sowohl in zweidimensionaler als auch stereoskopisch dreidimensionaler Ansicht zu generieren. Die einzelnen Blöcke entsprechen jeweils einer Systemeinheitslänge und wurden in randomisierter Weise zusammengeführt. Die Grundfarbe dieser Würfelfiguren war schwarz (RGB 3, 3, 3), wobei durch Lichtstreuung und –spiegelung der in Abbildung 10 ersichtliche Grauton konstruiert wurde.

Standardisiert wurden die einzelnen Aufgaben über ein 900 x 900 Pixel großes Fenster ausgegeben und anschließend in das Bildbearbeitungsprogramm SnagIT 9.13 zur weiteren Bearbeitung übernommen.

In einem nächsten Schritt wurden die so erzeugten Würfelaufgaben einheitlich auf 270 x 270 Pixel skaliert, als Bitmap abgespeichert, und in die Testentwicklungssoftware Teach And Test (TNT; Brugger, 1993) eingebunden.

EMPIRISCHER TEIL

5 Datenerhebung

5.1 Durchführung

Nach einer fast einjährigen Entwicklungsphase, in der die einzelnen Items mittels dem Software Programm „Visualization Toolkit“ generiert und in einen eigens hierfür programmierten Test eingebunden wurden, erfolgte Anfang April 2010 im Rahmen des Forschungsseminars für Fortgeschrittene bei Prof. Dr. Gittler eine Vorstudie. Bis auf zwei Items, die in der Vorstudie von allen Teilnehmern gelöst wurden, wurde der Test als gut befunden.

Die beiden Aufgaben, die in der Vorstudie auffällig geworden waren, wurden als Instruktionsbeispiele eingesetzt, da sie leicht genug waren, die Anforderungen an die Probanden verständlich zu erklären. Weiters diente die Vorstudie der Klärung der Frage, ob einzelne Items möglicherweise zu schwer bzw. zu uneindeutig seien, und daher nicht gelöst werden könnten. Diese wären nämlich für die Testung ungeeignet gewesen. Ferner wurden mögliche Zusatzfragen in Bezug auf den erleichternden Effekt der stereoskopischen Darstellungsweise im Vergleich zur zweidimensionalen Ansicht besprochen.

Für die tatsächlichen Datenerhebung, die Ende Mai 2010 stattfand, wurden über einen Zeitraum von zwei Wochen Psychologiestudenten des 1. Abschnittes rekrutiert, die im Rahmen der Lehrveranstaltung Differentielle Psychologie einen Prüfungsbonus erwerben wollten, und hierfür noch zwei weitere, nicht Psychologie studierende Freunde, mitbringen mussten. Aufgrund des hohen Frauenanteils im Studienfach Psychologie, durften die Teilnehmer einer solchen Dreiergruppen nicht alle demselben Geschlecht angehören, um einen Überhang an weiblichen Testpersonen zu vermeiden.

Da es sich um eine reine Computertestung handelte, fand die Testdurchführung für alle Teilnehmenden im psychologischen Zentrallabor der Fakultät für Psychologie im Neuen Institutsgebäude (NIG, Universitätsstraße 7, 6. Stock) statt und dauerte durchschnittlich 30 – 45min.

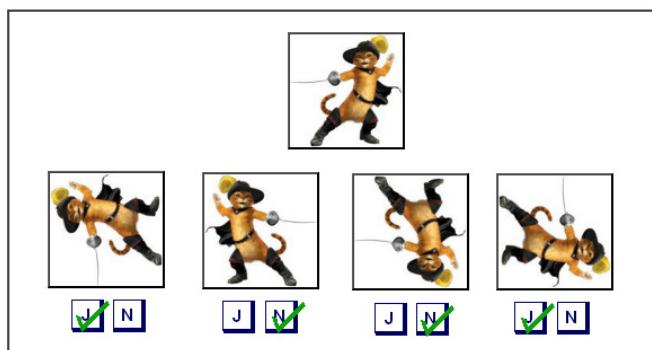
Vor Beginn der eigentlichen Testung wurden die Teilnehmer gebeten, Angaben zu ihren soziodemographischen Daten zu machen (Alter, Geschlecht, Bildungsgrad, etc. – siehe Anhang). Zusätzlich wurde die Frage nach einer möglichen Farbsehschwäche (rot-grün/rot-blau-Blindheit) gestellt, da dies einen störenden Effekt auf die Wahrnehmungsfähigkeit von stereoskopisch dargebotenem Itemmaterial haben könnte.

Nachdem die Probanden ihr Einverständnis zur Testung und zur anonymisierten Auswertung der Daten für wissenschaftliche Zwecke gegeben hatten, folgte auf mehreren Bildschirmseiten die Instruktion. Auf der ersten Seite erfuhren die Probanden folgendes:

In den nun folgenden Aufgaben geht es darum eine geometrische Ausgangsfigur mit vier dargebotenen Antwortfiguren zu vergleichen. Diese vier Antwortfiguren zeigen die Ausgangsfigur aus einer anderen Perspektive oder aber ein Spiegelbild der Ausgangsfigur ebenfalls aus einer anderen Perspektive. Ihre Aufgabe ist es für jede der vier Alternativen zu entscheiden, ob es sich um die Ausgangsfigur handelt oder um ein Spiegelbild.

Auf der nächsten Bildschirmseite folgte eine Beispielaufgabe mit weiteren Erklärungen:

Im unten dargestellten Fenster sehen Sie in der ersten Zeile die Ausgangsfigur, darunter befinden sich die vier Antwortfiguren. Ihre Aufgabe besteht darin, für jede der vier Antwortfiguren zu entscheiden, ob sie der Ausgangsfigur gleicht oder nicht. Im Beispiel entsprechen die Antworten eins und vier der Ausgangsfigur, deshalb ist „J“ für JA markiert. Die Antworten zwei und drei entsprechen nicht der Ausgangsfigur (Spiegelbilder), deshalb ist „N“ für NEIN markiert.



Um zu überprüfen, ob die Testteilnehmer diese Instruktion verstanden hatten, folgten zwei Beispielaufgaben, bei denen den Probanden sofort nach der Bearbeitung rückgemeldet wurde, ob ihre Antworten richtig waren, oder nicht. Entschied sich ein Teilnehmer zwei Mal hintereinander für die falsche Antwort, erschien automatisch am Bildschirm ein Hilfetext, der ihm die richtige Lösung und eine Erklärung hierfür präsentierte.

Wurden die beiden Beispielaufgaben richtig gelöst, folgte noch eine letzte Bildschirmsseite mit wichtigen Informationen vor Testbeginn:

Es folgen nun 21 Testaufgaben. Sie sollen möglichst genau arbeiten, und trotzdem so rasch wie es Ihnen möglich ist. (Eine Stoppuhr rechts oben im Bild soll Sie daran erinnern). Als richtige Lösungen können alle Antwortkombinationen in Frage kommen (z.B. 4xJ, 4xN, sowie alle weiteren Kombinationen). Im Laufe der Testbearbeitung werden Sie gebeten, eine so genannte 3D-Brille aufzusetzen. Mit dieser Brille erscheinen die Bilder räumlich-dreidimensional. (Die Brillenträger unter Ihnen setzen die 3D-Brille über die eigene Brille). Nutzen Sie die Pausen beim Brillenwechsel, um Ihre Augen auszuruhen. Danach sollen Sie wieder mit voller Konzentration weiterarbeiten. Im Test wird Ihnen nicht mehr rückgemeldet, ob Sie die Aufgabe richtig oder falsch bearbeitet haben.

5.2 Versuchsplan

Das Versuchsdesign sah eine randomisierte Zuweisung der Probanden zu den Testversionen A bzw. B vor, die sich im Mittelteil in ihrem Darbietungsmodus unterschieden. Das genaue Versuchsdesign wurde in Abbildung 12 veranschaulicht. Für den Fall, dass eine Versuchsperson angeben sollte, dass für sie stereoskopisches Sehen nicht möglich wäre, wurde noch eine dritte Version C vorbereitet, die alle Aufgaben in zweidimensionaler Weise präsentierte. Allerdings wählte keiner der Probanden diese Version, weshalb sie im Folgenden keine Erwähnung mehr findet.

1. Erhebung der soziodemographischen Daten
2. Ausschluss von Person, die eine rot-blau-Farbschwäche haben (diese wurden gebeten Version C zu bearbeiten)

Randomisierte Zuweisung

Item NR (mit Warming up)		Version A (n=116)	Version B (n=111)
0		2d	2d
1		3d	3d
2		3d	3d
3		2d	2d
4		2d	2d
5		2d	3d
6		2d	3d
7		2d	3d
8		2d	3d
9		2d	3d
10		3d	2d
11		3d	2d
12		3d	2d
13		3d	2d
14		3d	2d
15		2d	2d
16		2d	2d
17		2d	2d
18		3d	3d
19		3d	3d
20		3d	3d

Abbildung 12 Versuchsdesign des SPACE 2d/3d

Die ersten und die letzten fünf Aufgaben (exklusiv dem ersten Warming-up Item) im Test dienten als sogenannte Link-Item, während die mittleren 10 Items in Version A und B entweder in stereoskopischer oder zweidimensionaler Form dargeboten wurden, um den unterschiedlichen Einfluss dieser beiden Ansichten in den folgenden Analysen miteinander vergleichen und den Effekt der stereoskopischen Aufgaben abschätzen zu können.

5.3 Beschreibung der Stichprobe

Die Stichprobe mit $n = 227$ Personen setzte sich wie folgt zusammen:

- Die Geschlechterverteilung war annähernd gleich verteilt, da 100 männliche (44,1%) und 127 weibliche (55,9%) Probanden an der Testung teilnahmen.
- Für die beiden Testversionen ergab sich folgende Aufteilung nach dem Geschlecht.

Tabelle 1 Kreuztabelle für Geschlecht und Testversion

		Version		Gesamt
		A	B	
Sex	männlich	52	48	100
	weiblich	64	63	127
Gesamt		116	111	227

- Das Alter der Teilnehmer lag zwischen 18 und 30 Jahren ($\bar{x}_{\text{Alter}} = 21,45$; $SD = 2,57$), weshalb man davon ausgehen kann, dass die Probanden im Umgang mit dem Computer geübt waren, und dies nicht als Störvariable einen negativen Einfluss auf die Testung gehabt haben könnte.
- Bezüglich der Händigkeit gaben 203 Personen an, Rechtshänder zu sein. Weiters gab es 22 Linkshänder und 2 Beidhänder.
- 204 Teilnehmer (89,9%) gaben an, als höchste abgeschlossene Bildung die Matura bzw. das Abitur absolviert zu haben und zum Zeitpunkt der Testung zu studieren, während 20 Personen (8,8%) bereits einen Universitäts- bzw. Fachhochschulabschluss besaßen. Die restlichen 3 Probanden (1,3%) hatten entweder als höchste abgeschlossene Bildung eine berufsbildende Schule oder die Pflichtschule beendet.
- Die angegebenen Studien wurden zur einfacheren Darstellung zu sechs Untergruppen zusammengefasst. Gruppe 1 umfasst Studiengänge wie Humanmedizin, Veterinärmedizin, Genetik, usw., während in der Gruppe 2 technische Studien vertreten sind (z.B. Elektrotechnik, Maschinenbau, Technische Mathematik,...). Psychologiestudenten wurden in Gruppe 3 subsumiert. In der Gruppe 4 wurden wirtschaftliche Studienlehrgänge wie z.b Betriebswirtschaftslehre miteinander vereint. Lehramt-, Philosophie- und Politikwissenschaftsstudenten wurden unter Gruppe 5 resumiert und in Gruppe 6 fanden sich schließlich all jene wieder, die

angaben, Angestellte bzw. arbeitslos zu sein. Tabelle 2 veranschaulicht die Verteilung der einzelnen Untergruppen.

Tabelle 2 Deskriptivstatistik zu den Untergruppen an Studienrichtungen

	männlich	weiblich	Häufigkeit	Prozent	Kumulierte Prozente
Medizin, Genetik, u.ä.	14	9	23	10,1	10,1
technisches Studium	25	1	26	11,5	21,6
Psychologie	10	68	78	34,4	55,9
Wirtschaft	13	16	29	12,8	68,7
Sprachen/Lehramt/ Philosophie/Politik	32	29	61	26,9	95,6
Angestellter, Arbeitslos, u.ä.	6	4	10	4,4	100,0
Gesamt	100	127	227	100,0	

Aufgrund des hohen Anteils an Studenten in dieser Stichprobe, muss darauf hingewiesen werden, dass die hieraus gewonnenen Daten und Ergebnisse nicht zwingend auf die gesamte Bevölkerung zutreffen müssen.

6 Zielsetzung und Forschungsfragen

6.1 Zielsetzung

Wie dem Titel der Arbeit zu entnehmen ist, liegt das übergeordnete Ziel darin, einen Raumvorstellungstest zu entwickeln, der Rasch-homogen ist, dessen Aufgaben also zu ein und derselben Fähigkeitsdimension gehören, und der das Konstrukt „mentale Rotation“ alltagsnahe (stereoskopisch) testet.

6.2 Forschungsfragen

Da sich der Großteil der Literatur lediglich mit solchen Tests befasst, die 3-dimensionales Itemmaterial in 2-dimensionaler Form vorgeben, und mögliche Unterschiede zu anderen 2-dimensionalen Aufgabentypen untersuchen, lassen sich hierfür wenig hypothesenleitete Forschungsfragen für diese Arbeit ableiten. Deshalb stützen sich die hier angeführten Hypothesen größtenteils auf Forschungsergebnisse aus der Vorgabe des Endlosschleifen-Test 2d/3d (Gittler & Arendasy, 2003), die im Fachbereich Differentielle Psychologie an der Universität Wien gewonnen wurden:

- Unterscheiden sich die 2-dimensionalen und die 3-dimensionalen Aufgabentypen quantitativ oder qualitativ voneinander? Aufgrund der Ergebnisse der probabilistischen Testtheorie sollen die Itemschwierigkeiten der beiden Aufgabentypen untersucht werden, wobei erwartet wird, dass die herkömmlich zweidimensionalen Aufgaben durchwegs schwieriger zu bearbeiten sind, als die stereoskopisch-dreidimensionalen. In diesem Zusammenhang sollen die Fragen beantwortet werden, ob
 - beide Testversionen miteinander Rasch-homogen sind, wovon ausgegangen wird, und ob
 - der Test nach Martin-Löf (1973) nicht signifikant wird, da dies bedeuten würde, dass die Nullhypothese angenommen werden kann, also beide Präsentationstypen dieselbe latente Dimension messen. Auch hier wird erwartet, dass dies der Fall ist.
- Bleibt der, in der Scientific Community anerkannte, Geschlechtsunterschied bestehen, selbst wenn die Darbietung stereoskopisch ist? Bezug nehmend auf die Dissertation von Arendasy (2000) wird erwartet, dass sich der Leistungsunter-

schied zwischen Männern und Frauen unter der 3d Bedingung zwar deutlich verringert, allerdings trotz allem signifikant bestehen bleibt.

- Liegen die Leistungen im Test für die Aufgabentypen die in stereoskopisch 3-dimensionaler Form vorgegeben wurden, höher als bei jenen Aufgabentypen die in herkömmlicher 2-dimensionaler Weise präsentiert wurden? Erwartet wird, dass es einen signifikanten Leistungsanstieg unter der 3-dimensionalen Bedingung geben wird, da die Darbietung alltagsnäher ist, und weniger kognitive Verarbeitungsschritte notwendig sind, um solche Aufgaben zu lösen.
- Unterscheiden sich Männer und Frauen hinsichtlich ihrer Bearbeitungsgeschwindigkeit voneinander? Aufgrund der unterschiedlichen Verarbeitungsstrategien zwischen Männern und Frauen wird erwartet, dass Frauen signifikant längere Bearbeitungszeiten benötigen als Männer. Allerdings wird erwartet, dass sich die Bearbeitungszeiten für die dreidimensionalen Aufgaben für beide Geschlechter verkürzen werden, da ein Teil der Enkodierungsphase in der stereoskopischen Fassung wegfällt.
- Gibt es externe Kriterien (wie beispielsweise die Wahl des Studiums, Selbsteinschätzung bezüglich der Orientierungsfähigkeit im Raum,...), die die Leistungen der Probanden im Test vorhersagen können? Zu erwarten ist, dass Personen, die ein technisches Studium gewählt haben grundsätzlich bessere Leistungen zeigen werden, sowie jene Probanden, die angeben über einen guten Orientierungssinn zu verfügen.

7 Ergebnisse

7.1 Modelltest nach Rasch

Da das erklärte Ziel dieser Arbeit die Konstruktion eines Rasch-homogenen Testes war, soll die Gültigkeit dieses Modells geprüft werden. Sollte das Rasch-Modell für den neu konstruierten Test „SPACE 2d/3d“ empirisch gelten, so würde dies nach Gittler (1994) folgende Schlussfolgerungen zulassen:

1. Der Test misst für alle Personengruppen (begabt/unbegabt, jung/alt, weiblich/männlich, etc.) essentiell dieselbe Fähigkeitsdimension.
2. Der Testrohwert, also die Anzahl der gelösten Aufgaben im Test, stellt ein faires Leistungsmaß dar, und
3. die auf Basis des Rasch-Modells ermittelten Personenparameter sind intervall-skaliert, und der Zusammenhang zwischen den Rohwerten im Test (rangskaliert) und der latenten Fähigkeitsdimension ist empirisch abgesichert.

Um zu Überprüfen, ob die neu konstruierten Testaufgaben im Sinne des Rasch-Modells homogen sind, wurden folgende parametrischen Modelltests auf Basis der χ^2 -Verteilung (sog. Conditional Likelihood Ratio (CLR) – Tests) durchgeführt:

- Modelltest 1 – internes Kriterium: Teilung der Stichprobe anhand niedriger und hoher Vpn-Rohscores.
- Modelltest 2 – externes Kriterium: Teilung der Stichprobe nach dem Geschlecht der Probanden.
- Modelltest 3 – externes Kriterium: Aufteilung der Stichprobe anhand der beiden unterschiedlichen Testversionen A und B.
- Modelltest 4 – externes Kriterium: Stichprobenteilung nach dem Alter der Versuchspersonen (1. Gruppe ≤ 21 Jahre; 2. Gruppe > 21 Jahre).
- Modelltest 5 – externes Kriterium: Teilung der Stichprobe anhand des Warming-up-Items (gelöst/nicht gelöst).
- Modelltest 6 – externes Kriterium: Aufteilung der Stichprobe aufgrund der Selbst-einschätzung der Probanden bezüglich ihrer Fähigkeit zur räumlichen Orientierung auf einer Skala von -1 (sehr gut) bis 99 (sehr schlecht) (1. Gruppe ≤ 23 ; 2. Gruppe > 23)

Die Ergebnisse der durchgeführten Tests sind der nachfolgenden Tabelle 3 zu entnehmen, wobei der kritische χ^2 -Wert für $df = 19$ für $\alpha = 1\%$ bei 36,2165 und für $\alpha = 5\%$ bei 30,1348 liegt.

Tabelle 3 Ergebnisse der Modelltests im vollständigen Rasch-Modell (Total ln L = -2068.113541)

Modelltest	CLR-Testwert	df	Abweichung
1	27,6246	19	nicht sign.
2	26,0296	19	nicht sign.
3	23,2522	19	nicht sign.
4	28,0651	19	nicht sign.
5	17,3147	19	nicht sign.
6	29,0818	19	nicht sign.

Da alle Modelltests insignifikant waren, kann davon ausgegangen werden, dass die Items beider Versionen miteinander Rasch-homogen sind und weitere Analysen auf Ebene der probabilistischen Testtheorie vorgenommen werden dürfen.

Die Überprüfung, ob die Darbietung von Aufgaben in 2-dimensionaler oder stereoskopisch 3-dimensionaler Weise dieselbe latente Dimension der Raumvorstellung erhebt, lässt sich unter anderem mit dem Modelltest nach Martin-Löf klären. Beim Martin-Löf Test handelt es sich nämlich um einen Signifikanztest für die Itemhomogenität dichotomer Daten, wobei zwei möglichst heterogene Itemgruppen (in diesem Fall 2d und 3d) gebildet und „für beide Testteile getrennt die Parameter unter Maximierung der bedingten Likelihood geschätzt“ (Carstens & Rost, 2000:22) werden. Die Berechnung des ML-LRT für die Link-Items aus beiden Versionen ($k_1 = 6$ (Item 0, 3, 4, 15, 16, 17); $k_2 = 5$ (Item 1, 2, 18, 19, 20)) gemeinsam, erzielt ein nicht signifikantes Ergebnis ($\chi^2 = 49,6096$, $df = 29$, $p = ,8645$, $n = 227$). In den Unterteilungen nach Testversion ergibt sich ebenfalls für Version A ($\chi^2 = 146,2689$, $df = 109$, $p = ,6362$, $n = 116$) bei $k_{1A} = 11$ (Item 0, 3, 4, 5, 6, 7, 8, 9, 15, 16, 17) und $k_{2A} = 10$ (Item 1, 2, 10, 11, 12, 13, 14, 18, 19, 20) sowie Version B ($\chi^2 = 146,2689$, $df = 109$, $p = ,8254$, $n = 111$) bei $k_{1B} = 11$ (Item 0, 3, 4, 10, 11, 12, 13, 14, 15, 16, 17) und $k_{2B} = 10$ (Item 1, 2, 5, 6, 7, 8, 9, 18, 19, 20) ein nicht signifikantes Ergebnis, weshalb die Aufgabentypen 2d und 3d miteinander verrechnet werden dürfen, da sie offensichtlich dieselbe latente Dimension messen.

Im nächsten Schritt wurden die logarithmierten Itemschwierigkeiten für die beiden Testversionen A ($\hat{\sigma}_A$, n = 116) und B ($\hat{\sigma}_B$, n = 111) getrennt, sowie für die Link-Items über beide Versionen hinweg ermittelt. Die Berechnungen wurden jeweils ohne Warming-up Item durchgeführt.

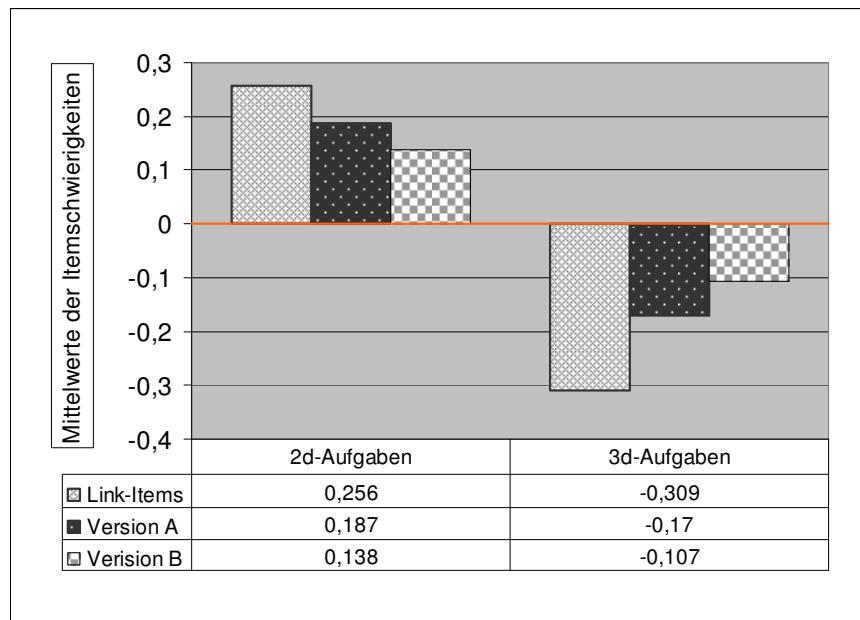


Abbildung 13 Mittelwerte der Itemschwierigkeiten für 2d- und 3d-Aufgaben

Wie in Abbildung 13 ersichtlich, weisen die zweidimensionalen Aufgaben erwartungsgemäß durchschnittlich höhere Itemschwierigkeiten auf als die stereoskopisch dreidimensionalen. Besonders deutlich wird die Differenz für die Itemschwierigkeiten der Link-Items über beide Versionen hinweg (n = 227, $k_{2d} = 5$ (Item 3, 4, 15, 16, 17); $k_{3d} = 5$ (1, 2, 18, 19, 20)), bei denen $\bar{\sigma}_{2d} = ,256$ und $\bar{\sigma}_{3d} = -,309$ beträgt.

Da es sich um einen neu entwickelten Test handelt, wurden die anschließenden Personenparameter ξ_v aus der Gesamtstichprobe erhoben. Der Personenparameter wurde über beide Versionen getrennt ($k = 20$) berechnet, wobei für beide Testversionen dieselben Personenparameter ermittelt wurden. Tabelle 4 zeigt welcher Personenparameter einer Person, je nach Anzahl der von ihr gelösten Aufgaben im Test, zugeschrieben wurde. Löste beispielsweise eine Person im Test insgesamt 6 Aufgaben, so lautet ihr Personenparameter $\xi_v = -,7029$. Je größer der Wert, desto fähiger ist die Person.

Tabelle 4 Gegenüberstellung der Personenparameter nach Warm (1989) mit der Anzahl der im Test gelösten Items und ihrer Antworthäufigkeiten (exkl. Warming-up-Item)

Anzahl gelöster Item	Häufigkeit des Antwortpattern	Personenparameter ξ_v
0	2	-3,6372
1	11	-2,4855
2	7	-1,9187
3	11	-1,5231
4	15	-1,2089
5	20	-0,9414
6	19	-0,7029
7	14	-0,4830
8	21	-0,2749
9	20	-0,0737
10	10	0,1249
11	15	0,3246
12	14	0,5291
13	17	0,7427
14	6	0,9709
15	8	1,2210
16	6	1,5041
17	5	1,8391
18	4	2,2627
19	2	2,8672
20	0	4,0693

Die graphische Veranschaulichung der Personenparameter im Vergleich zu den Rohscore-Werten in Abbildung 14 beschreibt deren empirisch bestimmten Zusammenhang auf der latenten Dimension. Der Vorteil dieser Graphik besteht darin, dass sie den Leistungszuwachs präziser diagnostizieren kann, als es auf Rohscore-Niveau möglich wäre, da eine „[...] numerisch gleiche Leistungssteigerung auf der manifesten Skala der Rohwerte [...] einen unterschiedlichen Fähigkeitszuwachs auf der latenten Dimension, je nachdem, in welchem Abschnitt der Rohwert-Skala diese Leistungssteigerung beobachtet wurde [...]“ (Gittler, 1994:110) bedeutet.

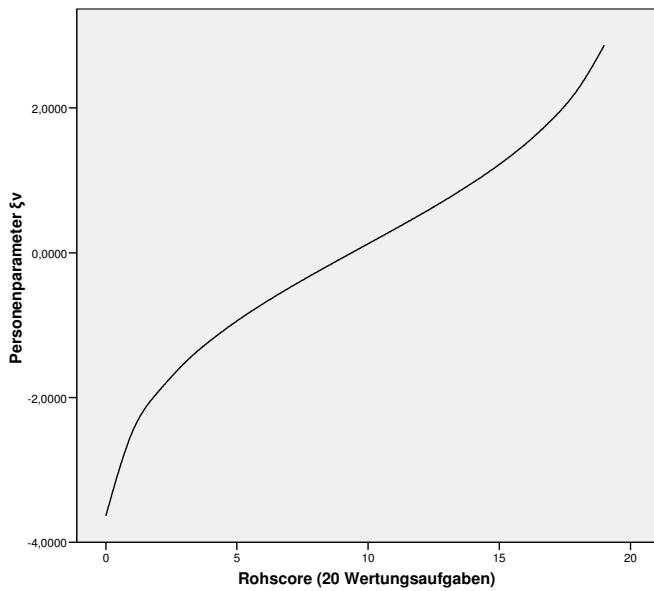


Abbildung 14 Der empirisch bestimmte Zusammenhang zwischen Rohwerten und Fähigkeiten (auf der latenten Dimension)

Die Personenparameter wurden im Anschluss an eine Gesamtschätzung für die 2d- und 3d-Aufgaben in den beiden Testversionen getrennt berechnet, und können der Tabelle 5 entnommen werden.

Tabelle 5 Gegenüberstellung der Personenparameter getrennt nach Version und 2d/3d in Bezug auf den Rohscore

Anzahl gelöster Items	Version A		Version B	
	Personen-parameter 2d	Personen-parameter 3d	Personen-parameter 2d	Personen-parameter 3d
0	-2,888	-3,0445	-2,9578	-2,981
1	-1,676	-1,8458	-1,7508	-1,778
2	-1,0382	-1,2238	-1,1187	-1,151
3	-0,5577	-0,7621	-0,6447	-0,6837
4	-0,1401	-0,3677	-0,2345	-0,2826
5	0,2567	0	0,1538	0,0929
6	0,6616	0,3677	0,5496	0,4699
7	1,1052	0,7621	0,9844	0,8755
8	1,6336	1,2238	1,5073	1,3508
9	2,3477	1,8458	2,2264	1,9895
10	3,6677	3,0445	3,5702	3,2087

Wie zu erwarten war, erhielten Personen die beispielsweise 5 2d Items richtig lösten höhere Personenparameter als für 5 gelöste 3d Aufgaben. Aus Tabelle 5 geht zusätzlich hervor, dass eine Person, die in Version A eine bestimmte Anzahl an Aufgaben gelöst hat, in etwa genauso befähigt ist, wie eine andere Person, die in Version B dieselbe Anzahl an Items gelöst hat.

Zusammenfassend lässt sich für die Rasch-Analysen sagen, dass aufgrund der nicht signifikanten Modelltests, die Rasch-Homogenität als gegeben angenommen werden darf. Die unterschiedlichen Itemschwierigkeiten für 2d- und 3d-Aufgaben stützen zusätzlich die Hypothese, dass die alltagsnahe stereoskopische Darbietungsweise einen erleichternden Effekt auf die Bearbeitung der Aufgaben haben dürfte. Die abschließende Berechnung der Personenparameter soll es im Weiteren ermöglichen, bei der Auswertung mittels klassischer Testtheorie auf intervallskalierte Daten zurückgreifen zu können.

7.2 Auswertung mittels klassischer Testtheorie

Test auf interne Konsistenz (Reliabilität)

Obwohl die interne Konsistenz aufgrund der bereits bewiesenen Rasch-Homogenität gegeben ist, wurde zusätzlich auf testtheoretischer Ebene ein Cronbach Alpha über beide Versionen des „SPACE 2d/3d“ hinweg berechnet. Mit $\alpha = ,802$ konnte auf Anhieb eine hohe interne Konsistenz erzielt werden. Für die einzelnen Versionen erhöht sich die interne Konsistenz sogar geringfügig auf $\alpha_{\text{Version A}} = ,805$ und $\alpha_{\text{Version B}} = ,814$.

Vergleichbarkeit der Testleistung von Testversionen A und B

Um zu Überprüfen, ob die beiden Testversionen A und B grundsätzlich miteinander vergleichbar sind, wurde in einem ersten Analyseschritt ein t-Test für beide Gruppen gerechnet, wobei nur jene Aufgaben berücksichtigt wurden, die in beiden Testversionen vorgekommen sind (ohne Warming-up Item sind dies also Item 1-4 und Item 15-20). Da dieser Test, wie erwartet nicht signifikant wurde ($t(225) = -,537$, $p = ,592$), kann davon ausgegangen werden, dass beide Testversionen gleich schwierig waren. Weiters zeigte sich auch, dass Männer, die Testversion A vorgegeben bekommen haben, sich nicht von jenen unterschieden die Testversion B bearbeitet haben. Dasselbe gilt für die Frauen aus Version A und B wie aus Tabelle 6 ersichtlich ist.

Tabelle 6 t-Test für unabhängige Stichproben für die Link-Item 1-4 und 15-20 getrennt nach Version und Geschlecht

			Levene-Test der Vari- anzgleichheit		T-Test für Mittelwertsvergleiche			
			F	Signifikanz	T	df	Sig. (2- seitig)	SD
sex	männlich	sco_10	,028	,867	-1,047	98	,298	,482
	weiblich	sco_10	,073	,787	,176	125	,860	,374

Geschlechtsunterschied auf Fähigkeitsparameterniveau

Da sichergestellt wurde, dass Männer und Frauen sich grundsätzlich in beiden Versionen auf einem vergleichbaren Leistungsniveau befinden, soll im Folgenden mittels klassischer Testtheorie die Frage nach dem Geschlechtsunterschied sowohl auf Ebene der Leistungsfähigkeit als auch auf Ebene der Bearbeitungsgeschwindigkeit untersucht werden

- für den Gesamttest (da der Fähigkeitsparameter über die beiden Testversionen für die Aufgaben 1- 20 gleich war)
- für die Bearbeitung von zweidimensionalen Aufgaben
- für die Bearbeitung von dreidimensionalen Aufgaben.

Die Berechnung eines t-Tests für unabhängige Stichproben ergab, dass es auf Basis der Personenparameter einen signifikanten Unterschied bei einem Signifikanzniveau von $\alpha=1\%$ zwischen den Geschlechtern gibt, sowohl in Bezug auf die Gesamttestleistung als auch auf Ebene der 2d- und 3d-Darbietungen. Der Geschlechtsunterschied bleibt also auch in der 3d-Bedingung signifikant bestehen.

Tabelle 7 t-Test für unabhängige Stichproben mit den abhängigen Variablen par_20, par_2d und par_3d in Bezug auf das Geschlecht

			Levene-Test der Vari- anzgleichheit		T-Test für Mittelwertsvergleiche			
			F	Signifi- kanz	T	df	Sig. (2- seitig)	SD
par_20			2,826	,094	4,971	225	<,001**	,1432523
par_2d			4,201	,042	4,183	225	<,001**	,1453904
par_3d			2,023	,156	4,389	225	<,001**	,1670667

Der Geschlechtsunterschied ist mit einer Effektstärke d nach Cohen als von mittlerer Stärke ($d = 0,66$; $\bar{x}_{\text{par20/m}} = 0,159$; $SD_{\text{par20/m}} = 1,133$; $\bar{x}_{\text{par20/w}} = -0,553$; $SD_{\text{par20/w}} = 1,02$)

für den Gesamttest zu bewerten. Dieser mittlere Effekt zeigt sich auch bei der Berechnung getrennt nach 2d und 3d Darbietung ($d_{2d} = 0,56$; $d_{3d} = 0,59$). Abbildung 15 verdeutlicht nochmals grafisch den signifikanten Geschlechtsunterschied auf Basis von 2d und 3d Personenparametern.

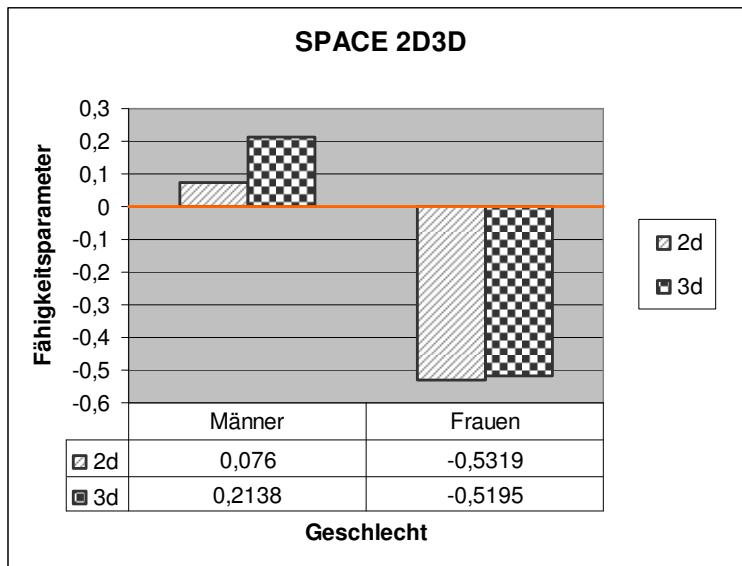


Abbildung 15 Darstellung des Geschlechtsunterschied in 2d und 3d Parameter

3d-Effekt als erleichternde Komponente

Im Weiteren soll nun die Frage geklärt werden, ob der eingesetzte 3d-Effekt eine signifikante Erleichterung bei der Bearbeitung der Aufgaben gebracht hat. Aus diesem Grunde wurde ein t-Test für gepaarte Stichproben auf Parameterniveau in 2d und 3d durchgeführt. Da das Ergebnis nicht signifikant wurde ($t(226) = - ,99$; $p = ,323$) kann davon ausgegangen werden, dass die Probanden nicht bzw. nur wenig vom 3-d Effekt profitiert haben. Ein ähnliches Resultat erhält man, wenn man denselben t-Test für gepaarte Stichproben für Männer und Frauen getrennt rechnet ($t_{\text{Männer}}(99) = -1,287$; $p = ,201$; $t_{\text{Frauen}}(126) = - ,14$; $p = ,889$). Allerdings zeigt sich, dass Männer anscheinend tendenziell mehr von der stereoskopischen Darbietungsweise profitieren als Frauen. Aufgrund einiger Rückmeldungen von Probanden, könnte dieser mangelnde 3d-Effekt auf eine zu intensive Darstellungsweise der stereoskopischen Aufgaben zurück zu führen sein.

Geschlechtsunterschied in der Bearbeitung von 2d und 3d Aufgaben

Zur Beantwortung der Frage, ob die Geschlechtsunterschiede bei Aufgabentypen, die in zweidimensionaler Weise präsentiert wurden, stärker vorhanden sind, wurde erneut ein t-Test gerechnet, allerdings für unabhängige Stichproben. Verrechnet wurden hierfür die 2d- und 3d- Fähigkeitsparameter getrennt nach Version A und B. Auf Basis eines Signifikanzniveaus von $\alpha = 5\%$ zeigte sich, dass in Version A der Geschlechtsunterschied in der 3d-Bedingung stärker vorhanden ist ($p < .001$), während in Version B der Geschlechtsunterschied in beiden Bedingungen konstant bleibt.

Tabelle 8 t-Test für unabhängige Stichproben für die 2d und 3d Parameter

		Levene-Test der Vari- anzgleichheit		T-Test für Mittelwertvergleiche				
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	SD
Version A	par_2d	1,112	,294	2,190	114	,031*	,4348106	,1985482
	par_3d	,086	,770	2,967	114	,004**	,7083260	,2387231
Version B	par_2d	3,014	,085	3,708	109	,000**	,7915359	,2134895
	par_3d	3,271	,073	3,240	109	,002**	,7625547	,2353523

Für Version A ergab die Effektstärke d für die 2d und 3d-dargebotenen Aufgaben einen kleinen bis mittleren Effekt ($d_{A/2d} = 0,41$; $d_{A/3d} = 0,55$) während das Cohen's d für Version B im mittleren bis starken Bereich angesiedelt ist ($d_{B/2d} = 0,71$; $d_{B/3d} = 0,62$).

Tabelle 9 Deskriptive Statistik für die 2d und 3d Parameter aufgeteilt nach den beiden Versionen A und B

			N	Mittelwert	SD
Version A	Par_2d	m	52	-,031927	1,1517370
		w	64	-,466738	,9862627
	Par_3d	m	52	,157063	1,2648251
		w	64	-,551263	1,2897616
Version B	Par_2d	m	48	,193504	1,2103244
		w	63	-,598032	1,0356091
	Par_3d	m	48	,275223	1,3923811
		w	63	-,487332	1,0877869

Bearbeitungszeit und Geschlecht

Bezüglich der Bearbeitungszeiten von Männern und Frauen, ergab sich, entgegen der Erwartung aufgrund möglicher unterschiedlicher Bearbeitungsstrategien und bisheriger

Resultate aus anderen Studien, kein signifikanter Unterschied ($t(225) = ,111; p = ,912$). Männer benötigten durchschnittlich für die Bearbeitung des Tests $\bar{x}_{time} = 12,82\text{min}$ ($SD_{time}=3,123$), während Frauen mit $\bar{x}_{time} = 12,78\text{min}$ ($SD_{time} = 2.523$) bei geringerer Varianz unbedeutend schneller gearbeitet haben (vgl. Abbildung 16). Hinsichtlich der Bearbeitung von zweidimensionalen Aufgaben und stereoskopischen Aufgaben finden sich ebenfalls keine signifikanten Unterschiede zwischen den Geschlechtern ($t_{2d}(225) = ,447; p_{2d} = ,655; t_{3d}(225) = -,26; p_{3d} = ,795$). Frauen wie Männer benötigen im Durchschnitt pro Darbietungsmodus etwa 6 Minuten.

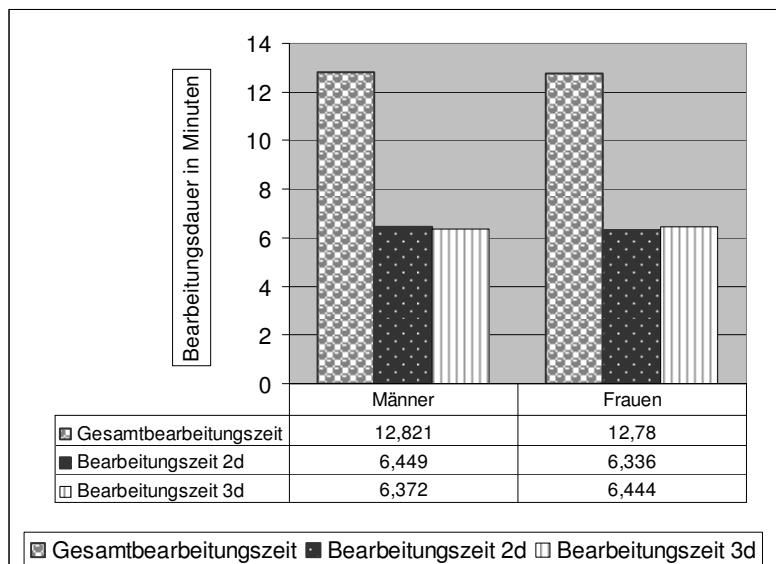


Abbildung 16 Darstellung der Bearbeitungszeiten für a) den Gesamttest, b) die 2d Aufgaben, c) die 3d Aufgaben

In einem nächsten Schritt wurden die unterschiedlichen Bearbeitungsgeschwindigkeiten der nicht miteinander verlinkten Aufgaben im Mittelteil des Tests, miteinander verglichen. Solche Items die in Version A zweidimensional präsentiert wurden, wurden mit den gleichen, allerdings stereoskopisch dargeboten, Aufgaben, verglichen ($t_{Aufgaben 5-9}(225) = -,88; p = ,38$). Dasselbe wurde anschließend mit den Aufgaben 10-14 für beide Versionen gerechnet mit einem ebenfalls insignifikanten Ergebnis ($t_{Aufgaben 10-14}(225) = -,076; p = ,940$). Die insignifikanten Resultate lassen darauf schließen, dass es bezüglich der Bearbeitungsgeschwindigkeit keinen Unterschied macht, ob die Aufgaben in zweidimensionaler oder dreidimensionaler Weise dargeboten werden.

Tabelle 10 Deskriptive Statistik für die mittleren Bearbeitungszeiten pro Version getrennt nach den Aufgabengruppen 5-9 und 10-14 im Test (in min.)

		Version	N	Mittelwert	SD
Männer	Time_5-9	A	52	4,1833	1,32997
		B	48	4,1076	1,58436
	Time_10-14	A	52	1,9502	,66666
		B	48	2,1096	,65870
Frauen	Time_5-9	A	64	3,9776	1,22026
		B	63	4,3302	1,50123
	Time_10-14	A	64	2,1634	,73557
		B	63	2,0488	,78843

Betrachtet man nun den Verlauf der Bearbeitungszeiten, wird anhand der Mittelwertvergleiche deutlich, dass beide Geschlechter für die Aufgaben 5 – 9 (unabhängig vom Vorgabemodus) länger brauchen als für die nachfolgenden Aufgaben 10 – 14 (siehe Tabelle 10), was auf einen möglichen Trainingseffekt hindeuten könnte.

Externe Kriterien und Testleistung

1. Planlesefähigkeit

Vor Beginn der Testung wurden die Probanden gebeten einige Aussagen über sich selbst zu treffen. Unter anderem bekamen Sie folgende Fragestellung, bei der Sie sich selber anhand einer Analogskala einschätzen mussten:

Stellen Sie sich vor, Sie sind in einer fremden Stadt und Sie wollen einen ganz bestimmten Platz finden. Um diesen zu finden benutzen Sie einen Stadtplan. Schätzen Sie bitte ein, wie gut Sie einen Weg mittels Stadtplan bzw. Landkarte finden können.

Die Analogskala reichte von der Schulnote „Sehr gut“ (-1) bis zur Note „Nicht genügend“ (99) und konnte mittels Schieberegler an jede beliebige Position bewegt werden. Personen, die sich selber eine sehr gute räumliche Orientierungsfähigkeit zuschrieben lösten im Durchschnitt 9,98 Aufgaben im Test während jene Personen, die geringes Vertrauen in ihre Orientierungsleistungen hatten, durchschnittlich nur 6,78 Aufgaben richtig bearbeiten. Wie zu erwarten war, gibt es einen starken Zusammenhang zwischen der selbst gewählten Note und der erzielten Leistung ($r_{\text{gesamt}} = - ,304$, $p < ,001$, $n = 227$; $r_{\text{Männer}} = - ,21$, $p = ,036$, $n = 100$; $r_{\text{Frauen}} = - ,273$, $p = ,002$, $n = 127$).

Tabelle 11 Deskriptivstatistik der ANOVA bezüglich Selbsteinschätzung "Stadtplanlesen" und dem Rohscore

	N	Mittelwert	Standard-abweichung	Standard-fehler	95%-Konfidenzintervall für den Mittelwert	
					Untergrenze	Obergrenze
Note 1	102	9,98	4,296	,425	9,14	10,82
Note 2	69	7,91	4,553	,548	6,82	9,01
Note 3	20	6,95	3,332	,745	5,39	8,51
Note 4	27	6,04	3,621	,697	4,60	7,47
Note 5	9	6,78	5,263	1,754	2,73	10,82
Gesamt	227	8,49	4,478	,297	7,90	9,07

Auch die einfaktorielle Varianzanalyse zeigt einen signifikanten Unterschied zwischen guten und schlechten „Stadtplanlesern“ ($F(4,222) = 6,663$, $p < ,001$). Vor allem jene, die sich selber die Note 1 gaben unterschieden sich signifikant von jenen, die sich „nur“ die Note 2 zuschrieben ($p = ,05$) sowie von jenen die sich lediglich „genügend“ orientieren konnten ($p = ,001$).

2. Studium

Wie bereits in der Beschreibung der Stichprobe (Kapitel 5.3) erwähnt, wurden die unterschiedlichen Angaben bezüglich Studium/Arbeit zu insgesamt sechs Untergruppen zusammengefasst. Diese Untergruppen sollen im Folgenden die Frage klären, ob die Wahl des Studiums als möglicher Prädiktor für die Leistung im „SPACE 2d/3d“ dienen kann. Die einfaktorielle ANOVA wird zwar signifikant ($F(5,221) = 4,625$; $p < ,001$) und weist in erwarteter Form Mittelwertsunterschiede zwischen jenen mit technischem Studium und Psychologiestudenten auf ($p = ,005$) allerdings lässt sich diese Differenz eher durch den generell gegebenen Geschlechtsunterschied als durch die unterschiedliche Präferenz in der Studienwahl erklären, da in der vorliegenden Stichprobe 25 männliche Technikstudenten lediglich einer weiblichen gegenüberstanden, während die Verteilung für die Studienrichtung Psychologie sich konträr verhielt (10 Männer, 68 Frauen). Für die genaue Verteilung von Männern und Frauen auf die einzelnen Studien sei an dieser Stelle nochmals auf Tabelle 2 in Kapitel 5.3 verwiesen.

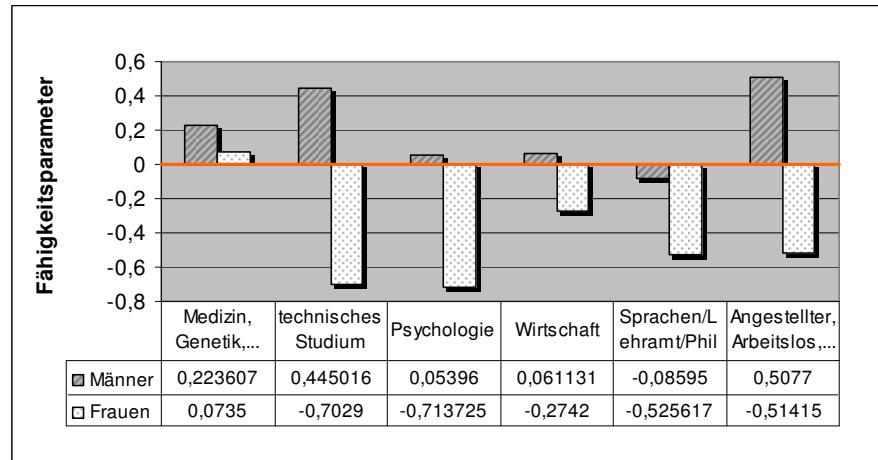


Abbildung 17 Grafische Darstellung der Mittelwertsverteilung der Fähigkeitsparameter getrennt nach Studium und Geschlecht

3. Alter

Ein weiterer möglicher Prädiktor für die Leistungen im „SPACE 2d/3d“ stellt das Alter dar, das in der erhobenen Stichprobe zwischen 18 und 30 Jahren ($\bar{x}_{\text{Alter}} = 21,45$; $SD = 2,57$) lag. Bei einer Unterteilung der Teilnehmer in fünf Gruppen (18/19J; 20J; 21J; 22/23J; 24J+) kam die einfaktorielle Varianzanalyse zu einem nicht signifikanten Ergebnis ($F(4,222) = 1,4$; $p = ,235$).

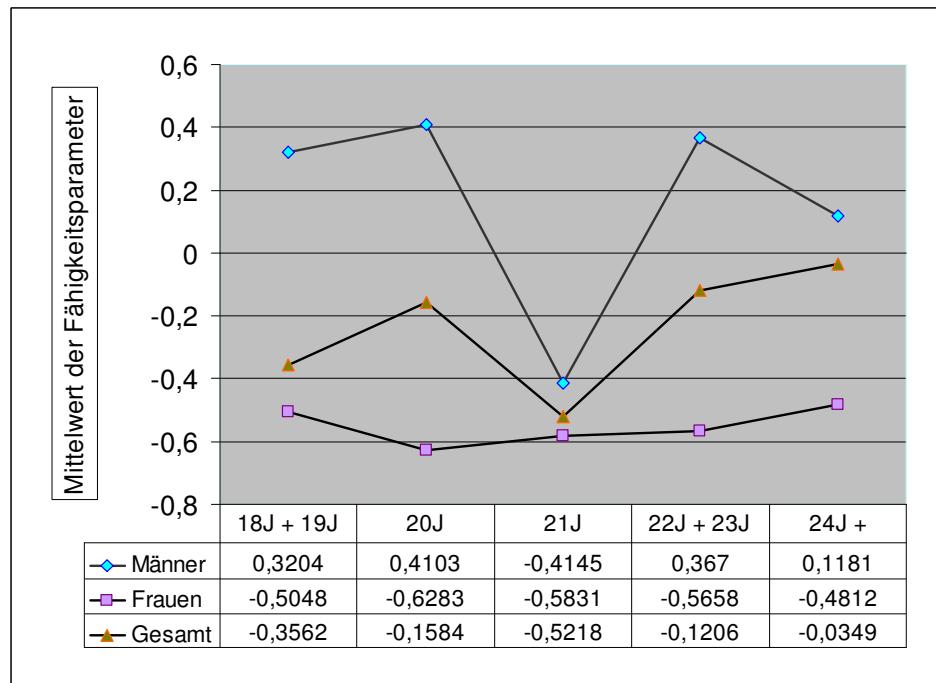


Abbildung 18 Grafische Darstellung der Parametermittelwerte für die einzelnen Altersgruppierungen nach Geschlecht getrennt

Wie in Abbildung 18 ersichtlich, liegt das Fähigkeitsniveau der Männer in allen Altersgruppen über dem der Frauen. Aber auch die nach Geschlecht getrennt gerechnete ANOVA blieb insignifikant (Männer: $F(4,95) = 1,524$; $p = ,201$; Frauen: $F(4,122) = ,074$; $p = ,99$).

Tabelle 12 Deskriptiv Statistik für die mittleren Personenfähigkeitsparameter getrennt nach Geschlecht

		N	Mittelwert	SD
männlich	18J + 19J	9	,320367	,7177967
	20J	19	,410311	,9642137
	21J	16	-,414519	1,0726789
	22J + 23J	21	,366995	1,2734998
	24J +	35	,118063	1,1991189
weiblich	18J + 19J	41	-,504795	,9568665
	20J	23	-,628335	1,1110639
	21J	28	-,583075	1,1106667
	22J + 23J	23	-,565791	1,1580255
	24J +	12	-,481150	,6260102

8 Zusammenfassung und Kritik

8.1 Zusammenfassung

Eine der zentralen Zielsetzungen für die vorliegende Arbeit war die Entwicklung eines Rasch-homogenen Raumvorstellungstests, der das Konstrukt „Mentale Rotation“ sowohl in herkömmlich zweidimensionaler als auch in stereoskopisch dreidimensionaler Weise erfassen soll. Anhand der insignifikanten Modelltests (nieder-hoch; etc.) kann dieses Unterfangen als erfolgreich bezeichnet werden. Zusätzlich konnte anhand der Martin-Löf Statistik nachgewiesen werden, dass die beiden Aufgabentypen (2d/3d) dieselbe latente Dimension erfassen.

Die Analysen auf Basis der klassischen Testtheorie ergaben auf Niveau von Fähigkeitsparametern, dass der vielfach zitierte und anerkannte Geschlechtsunterschied in dieser Studie repliziert werden konnte, und selbst bei alltagsnaher dreidimensionaler Darbietung signifikant bestehen blieb.

Bezüglich der Bearbeitungszeit konnte die These, dass Frauen grundsätzlich einen langsameren Arbeitsstil wählen als ihre männlichen Kollegen, nicht bestätigt werden, da beide Geschlechter für die Bearbeitung des Tests im Durchschnitt etwa gleich viel Zeit in Anspruch nahmen. Beide Geschlechter dürften zusätzlich im Verlauf des Tests von einem Trainingseffekt profitieren, da sie in der zweiten Hälfte im Mittel um 2 Minuten schneller wurden.

Für die externen Kriterien konnte gezeigt werden, dass lediglich die Selbsteinschätzung in Bezug auf die eigene Orientierungsleistung als Prädiktor für die zukünftige Leistung im SPACE 2d/3d herangezogen werden kann, da der signifikante Gruppenmittelwertsunterschied zwischen den Studienrichtungen hauptsächlich auf den Geschlechtsunterschied zurückzuführen ist, und das Alter in der vorliegenden Stichprobe zu gering streut, als dass es zu einem aussagekräftigen Unterschied kommen könnte.

8.2 Kritik und Anregung für künftige Forschung

Auf Grund der Tatsache, dass das Ziel dieser Arbeit die Entwicklung eines Rasch-homogenen Raumvorstellungstest war, fanden einige wichtige Aspekte bei der Niederschrift verhältnismäßig wenig Platz, unter anderem eine genaue Itemanalyse bzw. Vergleichsstudien zur Darbietungsmethode von Vandenberg und Kuse (1978) die für ihre Aufgaben noch ein Gitternetz einblenden mit jener Vorgabe ohne Gitternetz.

Da die Stichprobe zudem größtenteils aus Studenten besteht, lassen sich die Ergebnisse nicht ohne weiterer Überprüfung auf die allgemeine Bevölkerung übernehmen. Weiters wäre eine Kriteriumsvalidierung des Tests anhand einer parallelisierten Stichprobe von Technikern und Nicht-Technikern nicht von der Hand zu weisen. In der vorliegenden Studie konnte hierauf keine Rücksicht genommen werden, da es sich lediglich um eine erste Analyse des neukonstruierten Test handelte, weshalb die Techniker mehrheitlich durch Männer und Psychologen zumeist durch Frauen repräsentiert wurden.

Bezug nehmend auf das verwendete Itemmaterial wäre es außerdem von Interesse, den Stereoeffekt je nach subjektiven Empfinden anzupassen, da einige Probanden angaben, dass sie die Bearbeitung unter 3d-Bedingung teilweise zu schmerhaft für die Augen empfanden, während andere Probanden bestätigten, dass ihnen die Aufgaben unter stereoskopischer Darbietung leichter fielen. Analog wäre die Testung im „interlaced“ Modus (Bilder werden durch ein Zeilensprungverfahren und geeigneter Hardware dreidimensional wiedergegeben) eine sinnvolle Ergänzung zu den vorliegenden Ergebnissen.

Für die Itemanalyse bietet sich mittels Linear Logistischem Testmodell noch ein weites Forschungsfeld, das für sich genommen eine eigenständige wissenschaftliche Arbeit füllen kann. Von Interesse wären hierbei die Faktoren, die den Schwierigkeitsgrad eines Item maßgeblich bestimmen, wie beispielsweise die Anzahl an Würfeln, aus denen eine Figur besteht, ob die Figur mit zwei gleichlangen „Armen“ abschließt oder wie oft die Würfelfigur eine Abzweigung im Raum macht. Natürlich spielt auch der Rotationswinkel in diesem Zusammenhang eine wichtige Rolle.

Ebenfalls in zukünftigen Studien zu klären, ist die Frage, inwiefern sich die Antwortmuster von Männer und Frauen voneinander unterscheiden, denn in der vorliegenden

Studie wurden nur jene Aufgaben als richtig gewertet, in denen alle Teilaufgaben richtig waren. Lässt man jedoch Teilaufgaben zu, könnte sich der Geschlechtsunterschied womöglich auf die „impossible problems“ reduzieren – also auf jene Antworten, bei denen die Figuren gespiegelt dargeboten wurden (Kerkman, 2001).

9 Abstract – Deutsch

Der kognitive Geschlechtsunterschied zugunsten der Männer in mentaler Rotationsfähigkeit gilt in der Scientific Community mittlerweile als anerkannt und gesichert (Voyer, Voyer & Bryden, 1995). Ziel der vorliegenden Arbeit war es eine Rasch-homogene Version eines Raumvorstellungstests zu entwickeln, der sowohl zweidimensional als auch dreidimensional dargebotene Aufgabenstellungen enthält. Hierfür wurde der Mental Rotation Tests nach Vandenberg und Kuse (1978) modifiziert. Nach Überprüfung der Gültigkeit des Rach-Modells für das Aufgabenmaterial wurde untersucht, ob Frauen von einer stereoskopisch-dreidimensionalen Präsentation des Itemmaterials stärker profitieren als Männer. Insgesamt nahmen 227 Studierende (Männer = 100; Frauen = 127) unterschiedlicher Fachrichtungen an der Testung teil. Sie bearbeiteten jeweils 10 zweidimensionale und 10 stereoskopisch dreidimensionale Aufgaben. In der oberen Bildschirmhälfte bekamen die TestteilnehmerInnen eine Ausgangsfigur zu sehen, unter der vier Antwortfiguren in der gleichen Größe abgebildet waren. Die TestteilnehmerInnen hatten die Aufgabe für jedes Item zu entscheiden, ob es sich bei den Antwortalternativen um gedrehte Ansichten der Ausgangsfigur handelte, oder um gedrehte und gespiegelte Ansichten. Die Ergebnisse zeigen signifikante Geschlechtsunterschiede in beiden Bedingung (2d vs. 3d), wodurch sich zusammenfassend sagen lässt, der in der Literatur beschriebene Vorteil zugunsten der Männer auch in dieser Untersuchung unter Verwendung dreidimensionaler Aufgabendarbietung erhalten blieb.

Key Words: Mentale Rotation, Rasch-Modell, Geschlechtsunterschied

10 Abstract – English

Previous studies have demonstrated that men outperform women in cognitive tasks which require mental rotation (Voyer, Voyer & Bryden, 1995). There are different reasons for this sex difference discussed in the context of the nature-nurture problem. In the present study a modified Rasch-homogeneous Version of the Vandenberg and Kuse Mental Rotation Test (1978) was used to examine whether women benefit from a stereoscopic three dimensional presentation of the items. 227 undergraduate students of various branches of studies (100 males; 127 females) explored 10 two dimensional and 10 stereoscopic objects in a random assignment. The subjects view a target figure on the top of the page and four stimulus figures of equal size presented below. The subjects task was to decide for each of the four response alternatives whether they were solely rotated or a rotated mirror image of the target. Results show that the sex difference remains significant for both conditions (2d vs. 3d). In summary men and women obtain to the same extend from the stereoscopic presentation mode. Once more, the well known sex difference concerning mental rotation tasks could be validated.

Key Words: Mental Rotation, Rasch Model, Sex difference

11 Literaturverzeichnis

- Aloisi, A.M. (2007). Geschlecht und Hormone. In S. Lautenbacher, O. Güntürkün & M. Hausmann (Hrsg.), *Gehirn und Geschlecht* (S. 3–18). Heidelberg: Springer.
- Arendasy, M. (2000). *Psychometrischer Vergleich computergestützter Vorgabeformen bei Raumvorstellungsaufgaben: Stereoskopisch-dreidimensionale und herkömmlich-zweidimensionale Darbietung* (Unveröffentlichte Dissertation) Universität Wien, Wien.
- Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, 20, 327–344,
- Bjorklund, D.F., & Brown, R.D. (1998). Physical Play and Cognitive Development: Integrating Activity, Cognition, and Education. *Child Development*, 69(3), 604–606.
- Bugelski, B.R. (1970). Words and things and images. *American Psychologist*, 25, 1002–1012.
- Caplan, P.J., MacPherson, G.M., & Tobin, P. (1985). Do Sex-Related Differences in Spatial Abilities Exist?: A Multilevel Critique With New Data. *American Psychologist*, 40(7), 786–799.
- Carroll, J.B. (1993). *Human cognitive abilities: A survey of factoranalytic studies*. Cambridge: Cambridge University Press.
- Carstensen, C.H., & Rost, J. (2000). *MULTIRA ein Programmsystem zur Analyse mehrdimensionaler Rasch-Modelle. Handbuch zum Computerprogramm MULTIRA Einführung – Theorie & Praxis – Referenz*. Kiel: IPN. abgerufen auf http://www.ipn.uni-kiel.de/abt_ppm/multira/mmanual_de.pdf
- Cherney, I.D. (2008). Mom, let me play more computer games: they improve my mental rotation skills. *Sex Roles*, 59, 776–786.
- Cooke-Simpson, A., & Voyer, D. (2007). Confidence and gender differences on the Mental Rotations Test. *Learning and Individual Differences*, 17, 181 – 186.
- Dabbs, J.M. (1990). Age and seasonal variation in serum testosterone concentration among men. *Chronobiology International*, 7(3), 245–249.
- De Vries, G.J. (2004). Minireview: Sex Differences in Adult and Developing Brains: Compensation, Compensation, Compensation. *Endocrinology*, 145(3), 1063–1068.

- Dietrich, T., Krings, T., Neulen, J., Willmes, K., Erberich, S., Thron, A., & Wurm, S. (2001). Effects of Blood Estrogen Level on Cortical Activation Patterns during Cognitive Activation as Measured by Functional MRI. *Neuroimage*, 13, 425 – 432.
- El Koussy, A.A.H. (1935). The Visual Perception of Space. *British Journal of Psychology*, 20, 1–80.
- Eliot, J., & Smith, I.M. (1983). An International Directory of Spatial Tests. Windsor: Nfer-Nelson Publishing Company.
- Farah, M.J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, 95, 307-317.
- Fischer, G.H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I.W. Molenaar (Hrsg.), *Rasch Models. Foundations, Recent Developments, and Applications* (S. 131–155). New York: Springer.
- Fröhlich, W.D. (2005). *Wörterbuch Psychologie*. München: Deutscher Taschenbuch Verlag.
- Geiser, C., Lehmann, W., & Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence*, 36, 556–563.
- Gittler, G. (1994). Intelligenzförderung durch Schulunterricht: Darstellende Geometrie und räumliches Vorstellungsvermögen. In G. Gittler, M. Jirasko, U. Kastner-Koller, C. Korunka & A. Al-Roubaie (Hrsg.), *Die Seele ist ein weites Land – Aktuelle Forschung am Wieder Institut für Psychologie* (S. 105–122). Wien: WUV – Wiener Universitätsverlag.
- Gittler, G., & Arendasy, M. (2003). Endlosschleifen: Psychometrische Grundlagen des Aufgabentyps E^P. *Diagnostica*, 49 (4), 164–175.
- Guilford, J.P. (1967). *Nature of Human Intelligence*. New York: McGraw-Hill.
- Halpern, D.F., & Tan, U. (2001). Stereotypes and Steroids: Using a Psychobiosocial Model to Understand Cognitive Sex Differences. *Brain and Cognition*, 45, 392–414.
- Hausmann, M., Schoofs, D., Rosenthal, H.E.S., & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences – A psychobiosocial approach. *Psychoneuroendocrinology*, 34, 389–401.
- 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.

- Heider, C., Adlmann, E.M., & Gittler, G. (2010a, April). *Beeinflussen Hormone unsere kognitiven Fähigkeiten? Zum Zusammenhang von Östrogen und Raumvorstellung.* Poster präsentiert auf der 9. Tagung der Österreichischen Gesellschaft für Psychologie, Salzburg, Österreich. Abstract abgerufen auf <http://www.events.sbg.ac.at/oegp2010/Abstractband.pdf>.
- Heider, C., Adlmann, E.M., & Gittler, G. (2010b, Juli). *Sind die Hormone an allem Schuld? Zum Zusammenhang von Östrogen und Raumvorstellung.* Poster präsentiert auf der 2. Posterausstellung an der Fakultät für Psychologie
- Heil, M., & Jansen-Osmann, P. (2008). Sex differences in mental rotation with polygons of different complexity: Do men utilize holistic processes whereas women prefer piecemeal ones? *Quarterly Journal of Experimental Psychology*, 61(5), 683–689.
- Helleday, J., Bartfai, A., Ritzen, E.M., & Forsman, M. (1994). General intelligence and cognitive profile in women with Congenital Adrenal Hyperplasia (CAH), *Psychoneuroendocrinology*, 19, 343–356.
- Hines, M., Fane, B.A., Pasterski, V.L., Mathews, G.A., Conway, G.S., & Brook, C. (2003). Spatial abilities following prenatal androgen abnormality: targeting and mental rotations performance in individuals with congenital adrenal hyperplasia, *Psychoneuroendocrinology*, 28, 1010–1026.
- Hirnstein, M., Bayer, U., & Hausmann, M. (2009). Sex-Specific Response Strategies in Mental Rotation. *Learning and Individual Differences*, 19(2), 225–229.
- Holzinger, K.J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2, 41-54.
- Hornke; L. (2002). Item-Generation Models for Higher Order Cognitive Function. In S.H. Irvine & P.C. Kyllonen (Hrsg.), *Item Generation for Test Development* (S. 159–178). New Jersey: Lawrence Erlbaum Association Inc.
- Johnson, E.S., & Meade, A.C. (1987). Developmental Patterns of Spatial Ability: An Early Sex Difference. *Child Development*, 58, 725–740.
- 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.
- Kerkman, D.D., Wise, J.C, Harwood, A.E. (2001). Impossible ‘mental rotation’ problems: A mismeasure of women’s spatial abilities? *Learning and Individual Differences*, 12, 253–269.

- Kimura, D. (1997). Sex, sexual orientation and sex hormones influence human cognitive function. *Current Opinion in Neurobiology*, 6, 259–263.
- Kimura, D. (2002). Sex differences in the brain. *Scientific American*, 267(3), 32 – 37.
- Kimura, D., & Hampson, E. (1994). Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Current Directions of Psychological Science*, 3 (2), 57–61.
- Kosslyn, S.M. (1988). Aspects of cognitive neuroscience of mental imagery. *Science*, 240, 1621-1626.
- Linn, M.C. & Petersen, A.C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development*, 56, 1479-1498.
- Lohman, D.F. (1979). *Spatial ability: A review and re-analysis of the correlational literature* (Technical Report No. 8). Stanford: Stanford University, School of Education, Aptitude Research Project.
- Martin-Löf, P. (1973). *Statistika modeller*. Ansteckningar från seminarier läsaret 1969-70 utarbetade av Rolf Sundberg. 2: a uppl. Stockholm: Institutet för Försäkringsmatematik och Matematisk Statistik vid Stockholms Universitet.
- Massa, L.J., Mayer, R.E., & Bohon, L.M. (2005). Individual differences in gender role beliefs influence spatial ability test performance. *Learning and Individual Differences*, 15, 99–111.
- McGee, M.G. (1979). Human spatial abilities: psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86, 889–918.
- Moffat, S.D., & Hampson, E. (1996). A curvilinear relationship between testosterone and spatial cognition in humans: Possible influence of hand preference. *Psychoneuroendocrinology*, 21(3), 323–337.
- Newcombe, N., Bandura, M.M., & Taylor, D.G. (1983). Sex Differences in Spatial Ability and Spatial Activities. *Sex Roles*, 9(3), 377–286.
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, 76/3, 241–263.
- Peters, M. (2005). Sex differences and the factor of time in solving Vandenberg and Kuse mental rotation problems. *Brain and Cognition*, 57, 176–184.
- Peters, M., & Battista C. (2008). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition*, 66, 260–264.

- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg & Kuse mental rotation test: different versions and factors that affect performance. *Brain and Cognition*, 28(1), 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.
- Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences*, 40, 609–619.
- Rahman, Q., & Wilson, G.D. (2003). Large Sexual-Orientation-Related Differences in Performance on Mental Rotation and Judgment of Line Orientation Tasks. *Neuropsychology*, 17(1), 25–31.
- Resnick, S.M., Berenbaum, S.A., Gottesman, I.I., & Bouchard, T. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology*, 22/2, 191–198.
- Roberts, J.E., & Bell, M.A. (2003). Two- and three-dimensional mental rotation tasks lead to different parietal laterality for men and women. *International Journal of Psychophysiology*, 50, 235–246.
- Saucier, D.M., McCreary, D.R., & Saxberg J.K.J. (2002). Does gender role socialization mediate sex differences in mental rotations? *Personality and Individual Differences*, 32, 1101–1111.
- Schroeder, W., Martin, K., & Lorensen, B. (1998). *The Visualization Toolkit: An Object Oriented Approach to 3D Graphics*. New Jersey: Prentice Hall.
- Seidentopf, C. (18. August 2010). fMRI easy [Internetseite, die hilft Ergebnisse aus fMRI-Studien nachvollziehen zu können] abgerufen auf <http://www.fmri-easy.de>.
- Sheehan, P.W. (1967). A shortened form of Betts' questionnaire upon mental imagery. *Journal of Clinical Psychology*, 23/3, 386–389.
- Simpson, E.R. (2003). Sources of estrogen and their importance. *Journal of Steroid Biochemistry and Molecular Biology*, 86, 225–230.
- Solso, R.L. (2001). *Cognitive Psychology*. Boston: Allyn and Bacon.
- Shepard, R.N. (1975). Form, formation, and transformation of internal representations. In R.L. Solso (Hrsg.), *Information processing and cognition: The Loyola Symposium* (S. 103–117). Hillsdale, NJ: Erlbaum.

- Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.
- Silverman, I., Choi, J., & Peters, M. (2007). The Hunter-Gatherer Theory of Sex Differences in Spatial Abilities: Data from 40 Countries. *Archives of Sexual Behaviour, 36*, 261–268.
- Spearman, C.E. (1927). *The ability of man: their nature and measurement*. London: Macmillan.
- Spohn, F., & Buse, L. (2010). Freies Testosteron und Tageszeit als Determinanten individueller Unterschiede in der Raumvorstellung: Ergebnisse einer Feldstudie. Poster präsentiert auf dem 47. Kongress der Deutschen Gesellschaft für Psychologie, Bremen, Deutschland.
- Terlecki, M.S., Newcombe, N.S., & Little, M. (2008). Durable and Generalized Effects of Spatial Experience on Mental Rotation: Gender Differences in Growth Patterns. *Applied Cognitive Psychology, 22*, 996–1013.
- Thomsen, T., Hugdahl, K., Ersland, L., Barndon, R., Lundervold, A., Smievoll, A.I., Roscher, B.E., & Sundberg H. (2000). Functional magnetic resonance imaging (fMRI) study of sex differences in a mental rotation task. *Medical Science Monitor, 6*, 1186 – 1196.
- Thurstone, L.L. (1931). Multiple-factor analysis of vocational interests. *Personnel Journal, 10*, 198–205.
- Vandenberg, S.G., & Kuse, A.R. (1978). Mental Rotation: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599–604.
- Van Goozen, S.H.M., Cohen-Kettenis, P.T., Gooren,L.J.G., Frijda, N.H., & Van De Poll, A. (1995). Gender differences in behaviour: Activation effects of cross-sex hormones. *Psychoneuroendocrinology, 20*(4), 343–363.
- Vogel, J. J., Bowers, C. A., & Vogel, D. S. (2003). Cerebral lateralization of spatial abilities: A meta-analysis. *Brain and Cognition, 52*(2), 197–204.
- Voyer, D., Voyer, S., & Bryden, M.P. (1995). Magnitude of Sex Differences in Spatial Abilities: A Meta-Analysis and Consideration of Critical Variables, *Psychological Bulletin, 117*(2), 250–270.
- Warm, T.A. (1989). Weighted likelihood estimation of ability in item response models. *Psychometrika, 54*, 427–450.

12 Abbildungsverzeichnis

ABBILDUNG 1 HYPOTHETISCHE VERARBEITUNGSPROZESSE ZUR LÖSUNG VON MENTALEN ROTATIONSAUFGABEN BEI FRAUEN (LI.) UND MÄNNERN (RE.): (1) PRIMÄR VISUELLES AREAL, (2) EXTRASTRIATER KORTEX, (3) TEMPORAL INFERIORE /OCCIPITAL- INFERIORE REGION, (4) PARIETALE REGION (INTRAPARIETALER SULCUS), (5) PRÄMOTORISCHER KORTEX (JORDAN ET AL., 2002:2406)	20
ABBILDUNG 2 HORMONELLE SCHWANKUNGEN WÄHREND EINES MENSTRUATIONSZYKLUS	23
ABBILDUNG 3 DARSTELLUNG DER KLASSISCHEN UND MODERNS SICHTWEISE BEIM TESTEN KOGNITIVER FÄHIGKEITEN (MASSA ET AL., 2005:101)	26
ABBILDUNG 4 "TETRIS©"	27
ABBILDUNG 5 SCHEMATISCHE DARSTELLUNG DES PSYCHOBIOSOZIALEN MODELS BEI DEM ANLAGE UND UMWELT STETIG UND UNTRENNBAR SIND (HALPERN UND TAN, 2001:396)	30
ABBILDUNG 6 DIE DREI UNTERSCHIEDLICHEN AUFGABENTYPEN VON SHEPARD UND METZLER (1971)	31
ABBILDUNG 7 MENTAL ROTATION Test NACH VANDENBERG UND KUSE (1978).	32
ABBILDUNG 8 UNTERSCHIEDLICHE KOMPLEXITÄT DES RASTERAUFBAUS	33
ABBILDUNG 9 FIGUR MIT 1 BZW. 2 ABZWEIGUNGEN IM RAUM.....	33
ABBILDUNG 10 BEISPIEL FÜR EIN NEUGENERIERTES ITEM.....	35
ABBILDUNG 11 BEISPIEL FÜR EIN STEREOSKOPISCH GENERIERTES ITEM (ZU SEHEN MIT ANAGLYPHENBRILLE).....	35
ABBILDUNG 12 VERSUCHSDESIGN DES SPACE 2D/3D	44
ABBILDUNG 13 MITTELWERTE DER ITEMSCHWIERIGKEITEN FÜR 2D- UND 3D-AUFGABEN.....	51
ABBILDUNG 14 DER EMPIRISCH BESTIMMTE ZUSAMMENHANG ZWISCHEN ROHWERTEN UND FÄHIGKEITEN (AUF DER LATENTEN DIMENSION)	53
ABBILDUNG 15 DARSTELLUNG DES GESCHLECHTSUNTERSCHIED IN 2D UND 3D PARAMETER	56
ABBILDUNG 16 DARSTELLUNG DER BEARBEITUNGSZEITEN FÜR A) DEN GESAMTTEST, B) DIE 2D AUFGABEN, C) DIE 3D AUFGABEN	58
ABBILDUNG 17 GRAFISCHE DARSTELLUNG DER MITTELWERTSVERTEILUNG DER FÄHIGKEITSPARAMETER GETRENNT NACH STUDIUM UND GESCHLECHT.....	61
ABBILDUNG 18 GRAFISCHE DARSTELLUNG DER PARAMETERMITTELWERTE FÜR DIE EINZELNEN ALTERSGRUPPIERUNGEN NACH GESCHLECHT GETRENNT	61

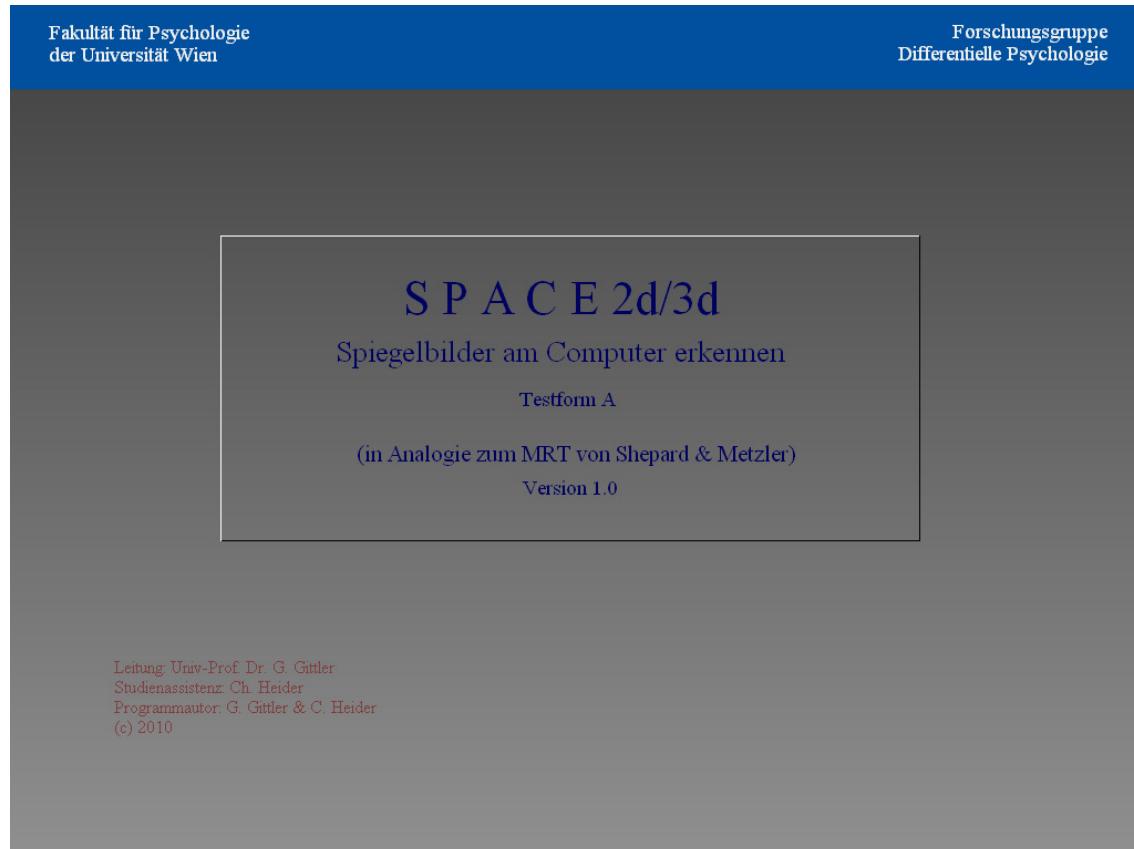
Ich habe mich bemüht sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit einzuholen. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

13 Tabellenverzeichnis

TABELLE 1 KREUZTABELLE FÜR GESCHLECHT UND TESTVERSION.....	45
TABELLE 2 DESKRIPTIVSTATISTIK ZU DEN UNTERGRUPPEN AN STUDIENRICHTUNGEN.....	46
TABELLE 3 ERGEBNISSE DER MODELLTESTS IM VOLLSTÄNDIGEN RASCH-MODELL (TOTAL LN L = - 2068.113541).....	50
TABELLE 4 GEGENÜBERSTELLUNG DER PERSONENPARAMETER NACH WARM (1989) MIT DER ANZAHL DER IM TEST GELÖSTEN ITEMS UND IHRER ANTWORTHÄUFIGKEITEN (EXKL. WARMING-UP-ITEM).....	52
TABELLE 5 GEGENÜBERSTELLUNG DER PERSONENPARAMETER GETRENNT NACH VERSION UND 2D/3D IN BEZUG AUF DEN ROHSCORE	53
TABELLE 6 T-TEST FÜR UNABHÄNGIE STICHPROBEN FÜR DIE LINK-ITEM 1-4 UND 15-20 GETRENNT NACH VERSION UND GESCHLECHT	55
TABELLE 7 T-TEST FÜR UNABHÄNGIGE STICHPROBEN MIT DEN ABHÄNGIGEN VARIABLEN PAR_20, PAR_2D UND PAR_3D IN BEZUG AUF DAS GESCHLECHT.....	55
TABELLE 8 T-TEST FÜR UNABHÄNGIGE STICHPROBEN FÜR DIE 2D UND 3D PARAMETER.....	57
TABELLE 9 DESKRIPTIVE STATISTIK FÜR DIE 2D UND 3D PARAMETER AUFGETEILT NACH DEN BEIDEN VERSIONEN A UND B.....	57
TABELLE 10 DESKRIPTIVE STATISTIK FÜR DIE MITTLEREN BEARBEITUNGSZEITEN PRO VERSION GETRENNT NACH DEN AUFGABENGRUPPEN 5-9 UND 10-14 IM TEST (IN MIN.)	59
TABELLE 11 DESKRIPTIVSTATISTIK DER ANOVA BEZÜGLICH SELBSTEINSCHÄTZUNG "STADTPLANESEN" UND DEM ROHSCORE	60
TABELLE 12 DESKRIPTIV STATISTIK FÜR DIE MITTLEREN PERSONENFÄHIGKEITSPARAMETER GETRENNT NACH GESCHLECHT	62

14 Anhang

14.1 Bilderdokumentation des „SPACE 2d/3d“ Vorlaufs



Herzlich willkommen!

Bitte Namen oder Probandencode eingeben:

Text eingeben und weiter mit der Enter-Taste.

Aktuelles Datum (tt/mm/jjjj) eingeben:

Weiter mit der Enter-Taste.

Geschlecht eingeben: (m = männlich, w = weiblich)

Buchstabe eingeben und weiter mit Enter-Taste.

Alter (in Jahren) eingeben:

Händigkeit eingeben: (r = Rechtshänder, b = Beidhänder, l = Linkshänder)

Ist Ihre Muttersprache Deutsch? (j = ja, n = nein)

Buchstabe eingeben und weiter mit Enter-Taste.

Höchste abgeschlossene Ausbildung:

- 1 = Sonderschule
- 2 = Volksschule
- 3 = Neunte Schulstufe (Pflichtschule)
- 4 = Berufsbildende Mittlere Schule / Fachschule / Lehre
- 5 = Matura (AHS, BHS) / Abitur
- 6 = Fachhochschule / Akademie
- 7 = Universität

Ziffer eingeben und weiter mit Enter-Taste.

Bitte geben Sie Ihre Studienrichtung / Ihren Beruf an:

Studium / Beruf angeben und weiter mit Enter-Taste.

Stellen Sie sich vor Sie sind in einer fremden Stadt und Sie wollen einen ganz bestimmten Platz finden. Um diesen zu finden benutzen Sie einen Stadtplan.

Schätzen Sie bitte ein, wie gut Sie einen Weg mittels Stadtplan bzw. Landkarte finden können.



Schieberegler mit Pfeiltasten zur gewünschten Position bewegen, dann weiter mit Enter.

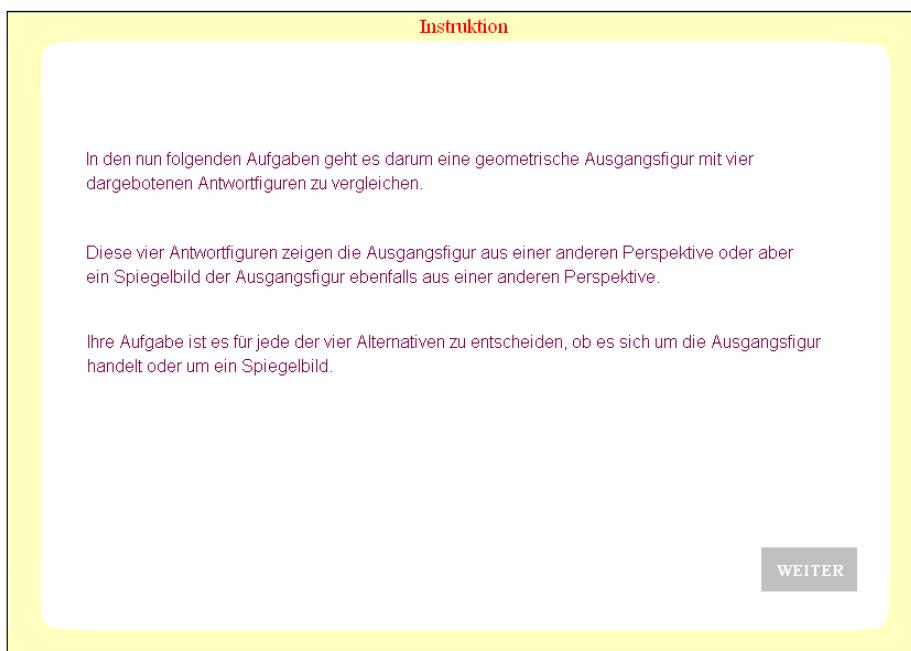
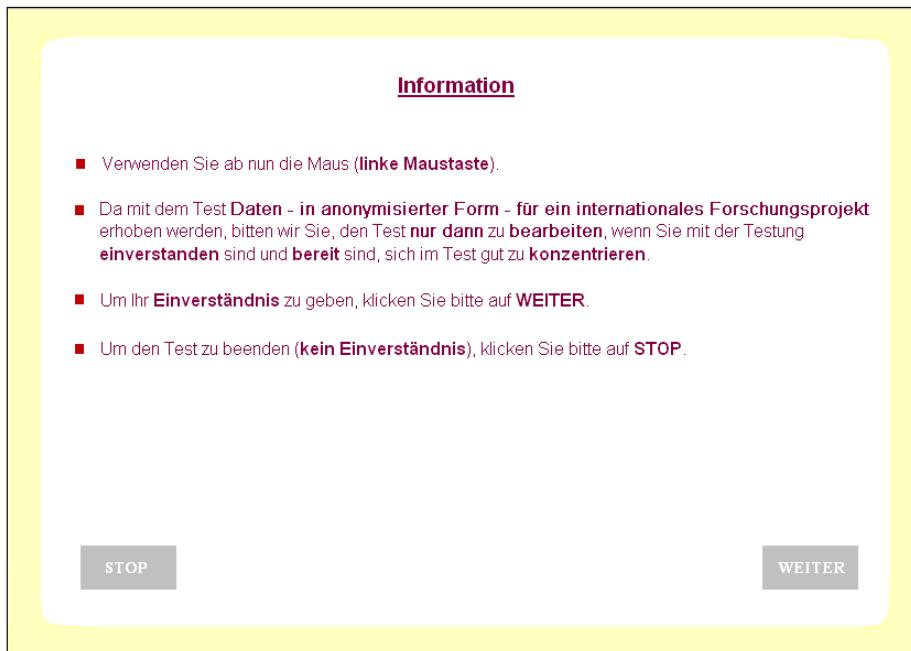
Ist bei Ihnen eine Farbschwäche bekannt?

1 = Ja (z.b. Rot-Grün-Farbschwäche)

2 = Nein

3 = andere Sehschwäche, die Tiefensehen nicht zulässt

Ziffer eingeben und weiter mit Enter-Taste.



Instruktion

Im unten dargestellten Fenster sehen Sie in der ersten Zeile die Ausgangsfigur, darunter befinden sich die vier Antwortfiguren.

Ihre Aufgabe besteht darin für jede der vier Antwortfiguren zu entscheiden, ob sie der Ausgangsfigur gleicht oder nicht.

Im Beispiel entsprechen die Antworten eins und vier der Ausgangsfigur, deshalb ist "J" für JA markiert. Die Antworten zwei und drei entsprechen nicht der Ausgangsfigur (Spiegelbilder), deshalb ist "N" für NEIN markiert.



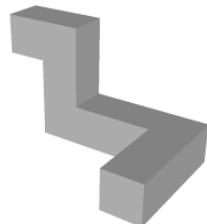
 J N J N J N J N

WEITER

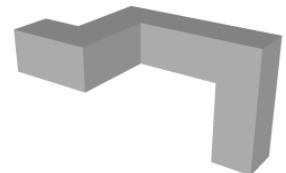
Beispiel 1

Bearbeiten Sie dieses Beispiel, indem Sie die richtigen Antworten anklicken.

Korrigieren ist möglich. Sobald Sie auf WEITER klicken, werden Ihre Antworten gespeichert.

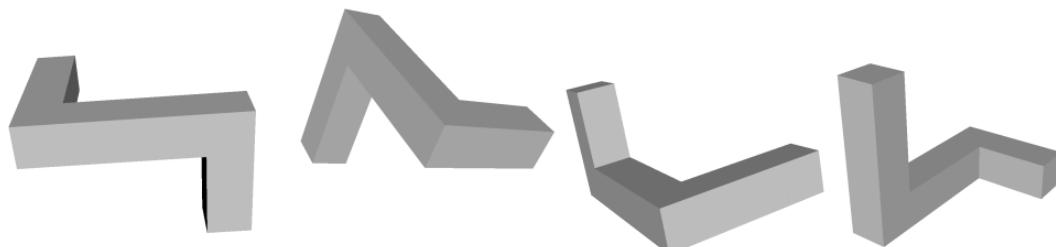

 J N

 J N

 J N

 J N

WEITER

Beispiel 2

**J N****J N****J N****J N****WEITER****Wichtige Informationen vor Testbeginn**

Es folgen nun 21 Testaufgaben. Sie sollen möglichst genau arbeiten, und trotzdem so rasch wie es Ihnen möglich ist. (Eine Stoppuhr rechts oben im Bild soll Sie daran erinnern).

Als richtige Lösungen können alle Antwortkombinationen in Frage kommen (z.B. 4xJ, 4xN, sowie alle weiteren Kombinationen).

Im Laufe der Testbearbeitung werden Sie gebeten, eine so genannte 3D-Brille aufzusetzen. Mit dieser Brille erscheinen die Bilder räumlich-dreidimensional. (Die Brillenträger unter Ihnen setzen die 3D-Brille über die eigene Brille.)

Nützen Sie die Pausen beim Brillenwechsel, um Ihre Augen auszuruhen. Danach sollen Sie wieder mit voller Konzentration weiterarbeiten.

Im Test wird Ihnen nicht mehr rückgemeldet, ob Sie eine Aufgabe richtig oder falsch bearbeitet haben.

Um mit dem Test zu beginnen, klicken Sie WEITER.
Um die Instruktion noch einmal zu lesen, klicken Sie ZURÜCK.

ZURÜCK**WEITER**

14.2 Rasch-Modell-Test über beide Versionen

```
*****
*****
c:\temp\heider\rm.lis on Saturday 21. August 2010 at 17:18:39
Results for data file c:\temp\heider\rm.dat
obtained by means of program LpcM-WiN 1.0

*****
*****
Calculation of a Rasch Model
Total Sample
*****
*****
227 Persons
 1 Treatment Group
 21 Items (out of 25 Items)
 21 Item Parameter Estimates (sum-normalized)
 16 Iterations
Convergence reached in 0.0000 seconds

Total Log-Likelihood      -2186.390100

      Item          Standard
      Parameters    Errors

      1   0.1603      0.1405
      2   0.1822      0.0747
      3   0.2696      0.0699
      4   0.0057      0.0731
      5  -0.0614      0.0745
      6  -0.0389      0.0732
      7   0.4000      0.0708
      8  -0.0389      0.0732
      9  -1.1596      0.1692
     10   0.1383      0.0760
     11  -0.6819      0.1551
     12  -0.1291      0.0847
     13  -0.1291      0.0847
     14   0.5304      0.0862
     15   0.6833      0.0939
     16  -2.2440      0.2203
     17   0.9278      0.1173
     18   0.0943      0.0758
     19   0.3783      0.0693
     20   0.1603      0.0755
     21   0.5522      0.0886
     22 Not used!
     23 Not used!
     24 Not used!
     25 Not used!
Mean per Subgroup: 8.94

Median per Subgroup: 9.00

*****
      Conditional Likelihood-Ratio-Test
      (Mean)
*****
```

```

Andersen chi_2 =          29.6675
Degrees of Freedom =      20
chi_2 at alpha(5 percent) = 31.4021      (Wilson-Hilferty approxima-
tion)
chi_2 at alpha(1 percent) = 37.5910      (Wilson-Hilferty approxima-
tion)

```

```

*****
Conditional Likelihood-Ratio-Test
(Median)
*****
```

```

Andersen chi_2 =          29.0829
Degrees of Freedom =      20
chi_2 at alpha(5 percent) = 31.4021      (Wilson-Hilferty approxima-
tion)
chi_2 at alpha(1 percent) = 37.5910      (Wilson-Hilferty approxima-
tion)
```

14.3 Martin-Löf – Test für die 2d und 3d präsentierten Aufgaben

```

*****
C:\temp\heider\rm_link.lis on Tuesday 10. August 2010 at 18:10:37
Results for data file C:\temp\heider\rm_link.dat
obtained by means of program LpcM-WiN 1.0
```

```

*****
Calculation of a Rasch Model
Total Sample
*****

```

```

227 Persons
 1 Treatment Group
 11 Items (out of 15 Items)
 11 Item Parameter Estimates (sum-normalized)
 12 Iterations
Convergence reached in 0.0000 seconds
```

```
Total Log-Likelihood      -1010.438688
```

	Item Parameters	Standard Errors
1	0.1155	0.1383
2	0.1370	0.0718
3	0.2226	0.0681
4	-0.0359	0.0811
5	-0.1015	0.0854
6	-2.2032	0.2174
7	0.8699	0.1381
8	0.0509	0.0784
9	0.3292	0.0824
10	0.1155	0.0736
11	0.5000	0.1235

```

12 Not used!
13 Not used!
14 Not used!
15 Not used!
Mean per Subgroup: 4.81

```

Median per Subgroup: 5.00

```
*****
Martin-Löf Test
```

```
*****
```

```

Group 1: 1, 4, 5, 6, 7, 8,
Group 2: 2, 3, 9, 10, 11,
LogLikelihood (Group 1): -425.068952
LogLikelihood (Group 2): -357.582000
Martin-Löf-Statistic: 26.9330 n.s. (5% und 1%)
(chi_2(5%)= 42.5509, chi_2(1%)= 49.6096, and df=29)

```

14.4 WLE-Parameter nach Warm (1989)

RightCnt	StartLE	MLE	SE (MLE)	WLE	SE (WLE)
0	999.0000	999.0000	999.0000	-3.6372	1.7924
1	-2.9444	-2.8635	1.0276	-2.4855	0.9290
2	-2.1972	-2.1129	0.7477	-1.9187	0.7170
3	-1.7346	-1.6465	0.6293	-1.5231	0.6154
4	-1.3863	-1.2942	0.5628	-1.2089	0.5557
5	-1.0986	-1.0021	0.5209	-0.9414	0.5171
6	-0.8473	-0.7458	0.4932	-0.7029	0.4912
7	-0.6190	-0.5122	0.4749	-0.4830	0.4739
8	-0.4055	-0.2925	0.4634	-0.2749	0.4630
9	-0.2007	-0.0810	0.4574	-0.0737	0.4573
10	0.0000	0.1273	0.4562	0.1249	0.4562
11	0.2007	0.3366	0.4596	0.3246	0.4595
12	0.4055	0.5514	0.4680	0.5291	0.4674
13	0.6190	0.7764	0.4818	0.7427	0.4806
14	0.8473	1.0181	0.5027	0.9709	0.5003
15	1.0986	1.2854	0.5330	1.2210	0.5288
16	1.3863	1.5922	0.5775	1.5041	0.5702
17	1.7346	1.9639	0.6466	1.8391	0.6329
18	2.1972	2.4554	0.7667	2.2627	0.7376
19	2.9444	3.2389	1.0456	2.8672	0.9535
20	999.0000	999.0000	999.0000	4.0693	1.8351
		2.7950		3.2063	

14.5 Curriculum Vitae

Persönliche Angaben

Name: Christiane Heider, bakk. phil.
Geburtstag: 12. November 1983
Geburtsort: Wien
Eltern: Dr. Manfred und Dr. Ingeborg Heider

Ausbildung

1993 – 2002 Neusprachliches Gymnasium Theresianische Akademie
Seit 10/2002 Diplomstudium der Psychologie
Universität Wien
10/2005 – 01/2009 Bakkalaureatsstudium der Japanologie
Universität Wien
Seit 02/2009 Masterstudium für Wirtschaft und Gesellschaft in Ostasien
Universität Wien

Berufserfahrung

Seit 10/2004 Ehrenamtliche Rettungssanitäterin bei der Johanniter Unfall Hilfe Wien (JUH)
24.09. – 26.09.2007 Persönliche Assistentin der Tagungsleiterin der 9. Arbeitstagung der Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik 2007 (DPPD 2007)
10/2007 – 01/2008 Studentmentor für Studienanfängerinnen und Studienanfängern im Rahmen des PS Bildungspsychologie (CBM)
12/2007 – 01/2009 Praktikantin am Institut für Japanologie Wien

<u>02/2008 – 03/2008</u>	6-Wochen-Praktikum am Allgemeinen Krankenhaus der Stadt Wien – Klinische Abteilung für Neurologische Rehabilitation
<u>Seit 09/2008</u>	Studienassistentin am Institut für Klinische, Biologische und Differentielle Psychologie an der Fakultät für Psychologie der Universität Wien

Publikationen

- Adlmann, E., Gittler, G., & Heider, C. (2010a, April). *Geschlechtsunterschiede in mentaler Rotation: Sex versus Gender?* Poster präsentiert auf der 9. Tagung der Österreichischen Gesellschaft für Psychologie, Salzburg, Österreich. Abstract abgerufen auf <http://www.events.sbg.ac.at/oegp2010/Abstractband.pdf>.
- Adlmann, E.M., Heider, C., & Gittler, G. (2010b, Juli). *Geschlechtsunterschiede in der Raumvorstellung: Zum Einfluss der Geschlechtsrollenidentität.* Poster präsentiert auf der 2. Posterausstellung an der Fakultät für Psychologie.
- Gittler, G., Adlmann, E., & Heider, C. (2010). *Raumvorstellung alltagsnah getestet: Ein Vorteil für Frauen?* Vortrag auf der 9. Tagung der Österreichischen Gesellschaft für Psychologie, Salzburg, Österreich. Abstract abgerufen auf <http://www.events.sbg.ac.at/oegp2010/Abstractband.pdf>.
- Heider, C., Adlmann, E.M., & Gittler, G. (2010a, April). *Beeinflussen Hormone unsere kognitiven Fähigkeiten? Zum Zusammenhang von Östrogen und Raumvorstellung.* Poster präsentiert auf der 9. Tagung der Österreichischen Gesellschaft für Psychologie, Salzburg, Österreich. Abstract abgerufen auf <http://www.events.sbg.ac.at/oegp2010/Abstractband.pdf>.
- Heider, C., Adlmann, E.M., & Gittler, G. (2010b, Juli). *Sind die Hormone an allem Schuld? Zum Zusammenhang von Östrogen und Raumvorstellung.* Poster präsentiert auf der 2. Posterausstellung an der Fakultät für Psychologie